THE SEASONALITY OF SURFACE TEMPERATURE WARMING: A GLOBAL COMPARISON OF CLIMATE MODEL ENSEMBLES AND OBSERVATIONS

A Thesis

by

ALEXA ZABASKE

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	John Nielsen-Gammon
Committee Members,	Andrew Dessler
	Ping Chang
Head of Department,	R. Saravanan

December 2021

Major Subject: Atmospheric Sciences

ABSTRACT

The annual cycle of surface temperature is altered over time because the annual mean surface temperature trend is not constant throughout the year, but instead exhibits distinct seasonality. Over the last century, the observed annual Global Mean Surface Temperature (GMST) trend peaks in March, however historical simulations by General Circulation Models (GCMs) peak a few months earlier. This model to observation mismatch has been studied several times in the last 20 years, but the model discrepancy is still present in the latest generation of GCMs. This study quantifies the observed seasonal trends at individual grid points, and in the form of zonally averaged latitude bands, using surface air temperature (SAT) data over land and sea surface temperature (SST) for ocean regions for all regions of the globe. Three ensembles of coupled GCMs are compared to observations: the MPI Grand Ensemble, the CMIP5 and CMIP6 multi-model ensembles. The use of large climate model ensembles enables the quantification of forced trends and effects due to natural variability in the seasonality of long-term surface temperature trends. Long-term seasonal temperature trends are calculated as the annual harmonic of surface temperature trends. The distributions of the simulation ensemble members are compared to the observations using the Mahalanobis distance statistic.

The largest mismatch between models and observations stems from the GCMs' gross underestimation of the forced seasonal warming trends that occurs over Northern Hemisphere (NH) mid to high latitude regions. Large seasonal warming is observed in Southern Hemisphere (SH) mid-latitude SSTs with peak warming in March, thereby

reinforcing the observed seasonality in the NH and GMST trends. The observed large boreal spring peak warming trend in NH land regions suggests the snow albedo feedback could be the primary mechanism that is altering the seasonal cycle of surface temperature, according to a conceptual model of energy balance. Using the same conceptual model, the simulated seasonality of warming over NH land in GCMs suggests that the sea-ice albedo is the dominant forcing mechanism driving changes to the seasonal cycle of surface temperatures.

ACKNOWLEDGEMENTS

I'd like to give the greatest thanks to Dr. Nielsen-Gammon (N-G), for working with me since I was an undergraduate student. It's been a privilege to work as his student throughout my graduate studies, and with his guidance I have grown tremendously as a scientist. Additionally, I'd like to thank the members of my thesis committee, Dr. Dessler, for introducing Dr. N-G and I to the idea of using the MPI Grand Ensemble for this project, and Dr. Chang, for suggesting a spatial grid-box analysis to support the latitude band analysis I had already started.

Thank you to my research group, for their support, positive attitudes, and being wonderful friends. To my friends - those who I've met via our studies of atmospheric sciences, and those who've always been around - thank you for your support and encouragement from near and far.

And lastly, I am so thankful for my fiancé and my family, for their endless love, support, and patience as I worked on my thesis.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Professor John Nielsen-Gammon and Professor Andrew Dessler of the Department of Atmospheric Sciences and Professor Ping Chang of the Department of Oceanography.

Funding Sources

Graduate study was supported by the Office of the Texas State Climatologist.

NOMENCLATURE

NH	Northern Hemisphere
SH	Southern Hemisphere
GMST	Global Mean Surface Temperature
SAT	Near-Surface Air Temperature
SST	Sea Surface Temperature
tas	Model Standard Variable: Near-Surface Air Temperature
ts	Model Standard Variable: Surface Temperature
GCM	General Circulation Model, Global Climate Models
CMIP	Coupled Model Inter-comparison Project
RMS	Root Mean Squared
EOF	Empirical Orthogonal Function

TABLE OF CONTENTS

Page

ABSTRACT II

ACKNOWLEDGEMENTS IV

CONTRIBUTORS AND FUNDING SOURCES V

NOMENCLATURE VI

TABLE OF CONTENTS VII

LIST OF FIGURES IX

LIST OF TABLES XI

1. INTRODUCTION 1

1.1. Background 1
1.2. Motivation and Hypothesis 6

2. DATA 9

2.1. Observations 9
2.2. GCM Output 11

3. ANALYSIS METHODS 16

3.1. Data Masking 16
3.2. Data Aggregation and Trend Calculation 21
3.3. Fourier Analysis 22
3.4. Use and Interpretation of Ensembles 25
3.5. Ensemble Calculations and Statistics 26

4. CONCEPTUAL MODEL 28

4.1. Conceptual Model 28

5. GLOBAL ANALYSIS 34

5.1. Global Mean Analysis 34
5.2. Spatial Global Analysis 37

6. ZONAL BAND ANALYSIS 46

6.1. HadCRUT4.6 46
6.2. MPI-GE 49
6.3. CMIP5 and CMIP6 56

7. CONCLUSIONS 70

7.1. Results and Discussion 70
7.2. Summary and Conclusion 76

REFERENCES 80

LIST OF FIGURES

Page

Figure 1.1.1: *Monthly Global Surface Temperature* 4

Figure 1.1.2: *Seasonal SAT Linear Trends (Nigam et al., 2017)* 5

Figure 1.1.3: *Lag and Gain Fields (Stine et al., 2009)* 6

Figure 3.1.1: *HadCRUT4.6, CRUTEM4, and HadSST3* 19

Figure 3.1.2: *HadCRUT4.6 Land and Sea Classification* 20

Figure 3.1.3: *MPI-GE Land and Sea Classification Map* 21

Figure 3.3.2: *Fourier Analysis on Global Mean Trends* 24

Figure 4.1.1: *Conceptual Model Diagram* 33

Figure 5.1.1: *Global Mean Seasonal Warming Trend, Century* 36

Figure 5.1.2: *Global Mean Seasonal Warming Trend, Half Century* 36

Figure 5.1.3: *Global Mean Seasonal Warming Trends on Polar Plot* 37

Figure 5.2.1: *HadCRUT4.6 Amplitude* 42

Figure 5.2.2: *HadCRUT4.6 Phase* 42

Figure 5.2.3: *HadCRUT4.6 Ensemble Spread* 43

Figure 5.2.4: *HadCRUT4.6 Amplitude to Spread Ratio* 43

Figure 5.2.5: *MPI-GE Amplitude* 44

Figure 5.2.6: *MPI-GE Phase* 44

Figure 5.2.7: *MPI-GE Ensemble Spread* 45

Figure 5.2.8: *MPI-GE Amplitude to Spread Ratio* 45

Figure 6.1.1: *HadCRUT4.6, CRUTEM4, and HADSST3* 48

Figure 6.1.2: *HadCRUT4.6 Land and Sea* 49

Figure 6.2.1: *MPI-GE Sea Variables* 53

Figure 6.2.2: *MPI-GE vs. HadCRUT4.6 SST* 54

Figure 6.2.3: *MPI-GE Land and Sea* 55

Figure 6.2.4: *MPI-GE vs. HadCRUT4.6 SAT Land* 56

Figure 6.3.1: *Models vs. Observations SAT Land* 61

Figure 6.3.2: *Models vs. Observations SST* 62

Figure 6.3.3: *Models vs. Observations Combined Surface Temperature* 63

Figure 6.3.4: *Polar Plot, SAT 75° to 90°* 64

Figure 6.3.5: *Polar Plot, SAT 60° to 75°* 64

Figure 6.3.6: *Polar Plot, SAT 45° to 60°* 65

Figure 6.3.7: *Polar Plot, SAT 30° to 45°* 65

Figure 6.3.8: *Polar Plot, SAT 15° to 30°* 66

Figure 6.3.9: *Polar Plot, SAT 0° to 15°* 66

Figure 6.3.10: *Polar Plot, SST 0° to 15°* 67

Figure 6.3.11: *Polar Plot, SST 15° to 30°* 67

Figure 6.3.12: *Polar Plot, SST 30° to 45°* 68

Figure 6.3.13: *Polar Plot, SST 45° to 60°* 68

Figure 6.3.14: *Polar Plot, SST 60° to 75°* 69

Figure 6.3.15: *Polar Plot, SST 75° to 90°* 69

Figure 7.1.1(a-c): *Summary of Latitudinal SAT and SST Trends* 79

LIST OF TABLES

Page

Table 2.1: *Globally Averaged Observation Datasets*.....10

Table 2.2: *Gridded Global Observation Ensembles* 10

Table 2.3: *CMIP5 Model Output*.....13

Table 2.4: *CMIP6 Model Output*.....14

Table 2.5: *MPI-GE Model Output*.....15

1. **INTRODUCTION**

1.1. Background

Anthropogenic greenhouse gas emissions have raised the annual global mean surface temperature by approximately 1.1°C relative to pre-industrial levels. Warming of the global mean surface temperature has long-term effects such as sea-ice loss, sea-level rise, severe drought, stronger hurricanes, and devastating effects on the biosphere. Natural variability of the Earth's climate can temporarily dampen or amplify warming trends observed at the surface. For example, extreme El Niño or La Niña events influence seasonal surface temperature and precipitation among the northern mid-latitudes, and the Pacific Decadal Oscillation can modify hemispheric temperature trends over several years. Since the natural range of atmospheric-oceanic variability can be large, it may take many decades for the global warming signal to emerge in many climate-related phenomena. It is important to identify climate change trends as early as possible, both for our understanding of the climate system and to evaluate the fidelity of changes simulated by global climate models (GCMs).

The global mean warming trend is not uniform throughout the year. Previous studies have found that some seasons have faster rates of warming on decadal time scales. The timing of the annual maximum rate of warming, and its amplitude relative to the annual mean rate of warming, affect the timing of seasonal transitions of surface temperature. Figure 1.1.1 illustrates the observed and simulated global-scale monthly surface temperature trends over the last 130 years. There is a noticeable difference in the timing of the largest and smallest temperature trends simulated by climate models

compared to what has globally been observed. Not only is the global annual mean warming higher in observations than CMIP5 climate models, the amplitude of seasonal warming is also larger.

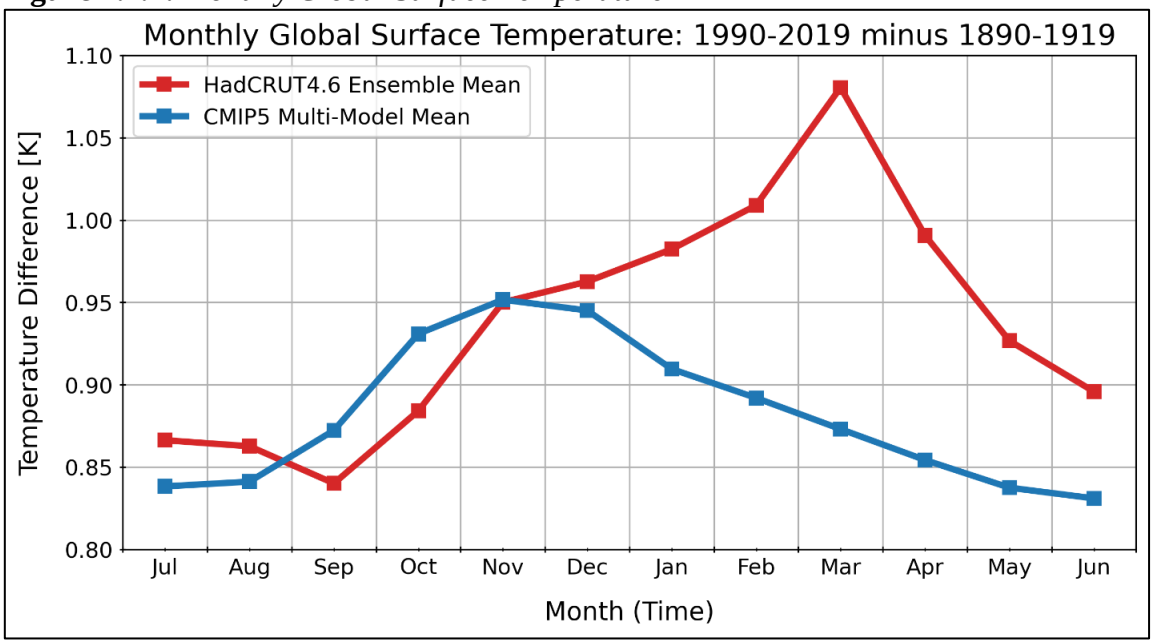
The timing and magnitude of seasonal warming is not spatially uniform. Northern Hemisphere (NH) observations of near-surface air and sea surface temperature in the mid and high latitudes exhibit the strongest differential warming signal (Mann and Park 1996, Wallace and Osborn 2002, Dwyer et al. 2012, Qian and Zhang 2015). Observations show that the fastest warming over land has occurred between late winter and early spring (see also Wang et al. 2009, Nigam et al. 2017). Figure 1.1.2 from the Nigam et al. 2017 study shows the linear trends of seasonal surface temperature across the Northern Hemisphere. Of the four seasons, NH spring has the fastest rate of warming, seen in the second of the four panels in Figure 1.1.2. Seasonal surface temperature trends in the Southern Hemisphere show a shift towards later seasons, however this shift is relatively smaller and less statistically significant than NH trends (Stine et al. 2009, Qian and Zhang 2015).

Figure 1.1.3, produced by Stine et al. 2009, analyzes the first harmonic of monthly surface temperature over 54 years, using both land and ocean data. The regions of lag towards earlier seasons on the NH continents are located in the same regions as spring peak warming trends shown in figure 1.1.2. Additionally, these same regions show a decrease in overall seasonal temperature amplitude (where Gain is negative in figure 1.1.3). This is due to the fact that a shift towards earlier seasons in the NH would result in a warmer NH winter, thus decreasing the total range of temperature in a given

year. The opposite trends can be noted over ocean surfaces; there is an increase in seasonal amplitude (gain) and a seasonal phase delay. While climate models generally agree with the NH average amplitude trend, they demonstrate phase delays, which is opposite of the observed phase trend towards earlier seasonal transitions. The majority of climate models simulate a maximum rate of warming that occurs in late fall to early winter (Mann and Park 1996, Stine et al. 2009, Nigam et al. 2017).

Models tend to overestimate the magnitude of seasonal warming in higher northern latitudes compared to observations, and underestimate seasonal warming trends in lower latitudes (Wallace and Osborn 2002, Dwyer et al. 2012). The phase lead (earlier seasons) seen in the NH land average and phase lag (later seasons) seen in the SH ocean average may be due to the natural variability that occur in observations (Mann and Park 1996, Stine and Huybers 2012). The primary driver of phase trends in models is driven by the sea-ice albedo effect at higher latitudes, and trends in the lower latitudes are driven by changes in surface heat fluxes (Dwyer et al. 2012, Donahoe and Battisti 2013).

Figure 1.1.1: *Monthly Global Surface Temperature*



The average monthly surface temperatures for the period 1890-1919 is subtracted from those in 1990-2019. This data is plotted for the HadCRUT4.6 (red) and CMIP5 (blue).

Figure 1.1.2: *Seasonal SAT Linear Trends (Nigam et al., 2017)*

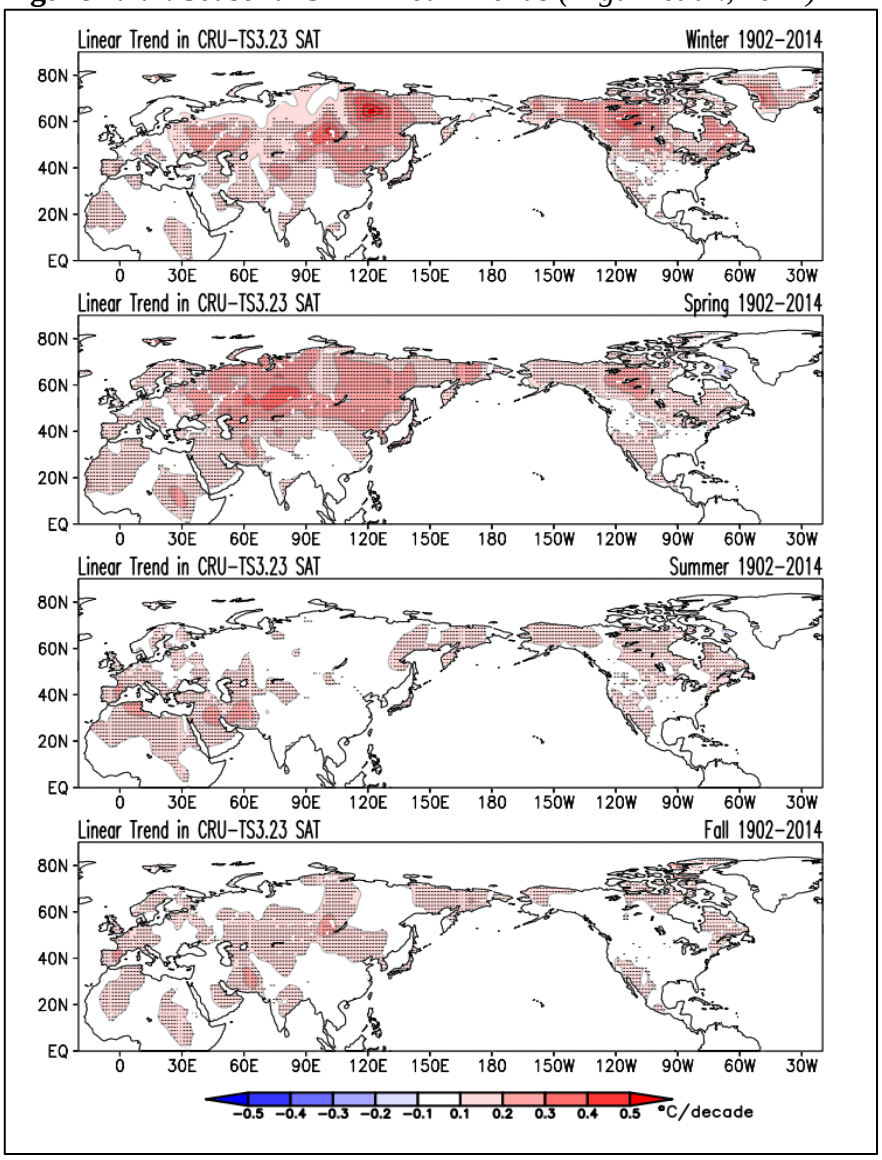


Figure 1 from Nigam et al. 2017: "The linear trend in seasonal near-surface air temperature (SAT) over the northern continents during 1902–2014. The 0.58° resolution CRU TS3.23 SAT is analyzed in boreal seasons: winter [December–February (DJF)], spring [March–May (MAM)], summer [June–August (JJA), and fall [September–November (SON)]. Contour interval and shading threshold is 0.18°C decade, with warm colors showing a positive trend. The fields are shown after nine applications of the nine-point smoother (smth9) in GrADS. Trends significant at the 95% level are stippled." Reprinted from Nigam et al. 2017.

Figure 1.1.3: *Lag and Gain Fields (Stine et al., 2009)*

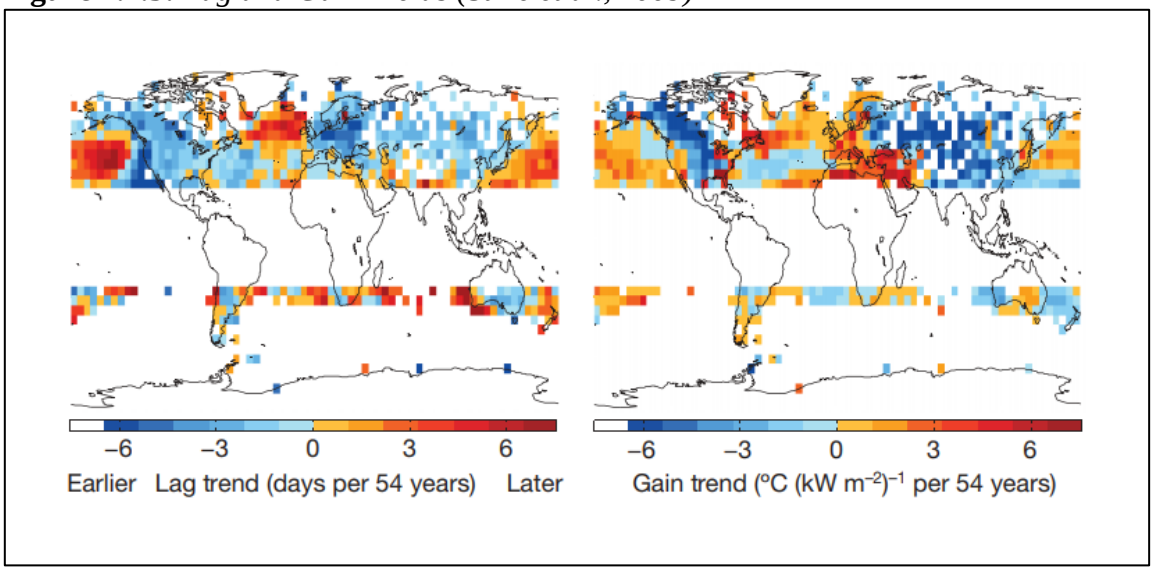


Figure 1e and 1f from Stine et al. 2009: "Lag and gain fields… 1e: Phase (lag, λ), trend in days per 54 years and 1f: Amplitude (gain, *G*), °C (kW m^{-2})$^{-1}$ per 54 years. Both variability and trend maps are plotted on the 'dense network' (1954–2007), without land and ocean masks applied. Results have been excluded in the tropics, where data availability is poor, and where less than 85% of the variance in an average year is explained by the yearly component." Reprinted from Stine et al. 2009.

1.2. Motivation and Hypothesis

A recent analysis of seasonal warming trends that compares climate models to observations across all regions of the globe, analyzing surface temperature over both land and ocean, does not currently exist. Most previous studies have focused on observed temperature trends, while relatively few have combined their analysis with climate model output from a multi-model ensemble, such as from the Coupled Model Intercomparison Projects (CMIPs). Moreover, the methods to quantify seasonal surface temperature trends in previous studies are typically done by linear analyses of the standard seasons (DJF, MAM, JJA, SON, see Figure 1.1.2), using EOF analysis, or by

analyzing trends in the first Fourier harmonic of monthly temperatures (see Figure 1.1.3). The research presented in this study uses Fourier analysis of monthly surface temperatures, but in a different fashion. Instead of fitting a Fourier series to the monthly average temperatures, the Fourier series will be fit to monthly average temperature differences between two distinct climatological time periods. By doing the Fourier analysis on the temperature differences, the timing and magnitude of the annual maximum warming rate can be computed. In other words, it reveals the annual harmonic of surface temperature trends.

This study seeks to determine the robustness of the annual harmonic of observed surface warming trends over the last several decades and the extent to which climate models reproduce this harmonic. A Fourier analysis of monthly mean surface temperature trends will be analyzed globally, locally, and in zonally-averaged latitude bands. Sensitivity to the extent of zonal averaging and selection of the land/ocean mask is tested and accounted for. Recent versions of existing global observation datasets will be analyzed, including the set of 100 surface temperature realizations from HadCRUT4.6 (Morice et al. 2012). The purpose of comparing multiple global observation datasets is to investigate the sensitivity of seasonal trends to the varying analysis schemes used by each dataset. The same methods will be applied to CMIP5 and CMIP6, in addition the MPI-GE 100-member grand ensemble (Taylor et al. 2012, Eyring et al. 2015, Maher et al. 2019). The analysis of the MPI ensemble members will evaluate the robustness of seasonal warming trends to natural variability in a single climate model. The comparison of CMIP5 to CMIP6 multi-model output will reveal

whether correction of structural deficiencies in the newest generation of climate models has led to better agreement with the observed seasonal warming trends. If the discrepancy in monthly surface warming trends between climate models and observations is determined to be larger than can be explained by natural climate variability or analysis scheme error, then, assuming observations are largely correct, model structural errors are yielding incompatible seasonal warming trends to what has historically been observed.

The datasets used in the study will be introduced and described in Section 2. The analysis methods applied to the data, such as land-sea masking and the Fourier analysis, will be outlined in Section 3. Section 4 will discuss a conceptual model for the seasonality of annual surface temperature warming. Section 5 will present the global analysis of seasonal warming trends with surface air and sea temperature together. In Section 6, the seasonal trends will be analyzed across zonal bands for surface air temperature and sea surface temperature individually. Section 7 will provide discussion of results and conclusions.

2. DATA

2.1. Observations

All of the observation data in this study are in the form of gridded datasets of surface temperature. Grid points over land are from measurements of near-surface 1.5 m to 2 m air temperature (abbreviated as SAT), while grid points over the ocean are measurements of sea surface temperature (abbreviated SST). Grid points along the coastal regions will generally include measurements from both SAT and SST. Tables 2.1 and 2.2 display all global observation datasets that will be used in this study. Each dataset has a unique way of combining discrete SAT and SST measurements into a continuous grid that spans the globe. The datasets have varying resolutions, and they differ in how they handle regions with sparse or no data, known and potential biases, urban heat island effect, and missing or erroneous data. Most of the datasets use the Global Historical Climatology Network (GHCN) for land measurements and International Comprehensive Ocean-Atmosphere Data Set (ICOADS) for sea measurements, but some include additional measurements from other sources. The observation datasets in Table 2.1 will be analyzed globally as is, with no land/sea mask applied. These data are downloaded from and globally averaged using the KNMI climate explorer database (Trouet and van Oldenborgh, 2013). They are only used for the globally averaged analysis section (5.1).

The HadCRUT4.6 ensemble listed in Table 2.2 will be used in the globally averaged analysis in addition to spatial and latitudinal analyses. The HadCRUT4.6 dataset is an integration of two datasets at 5° latitude-longitude resolution. CRUTEM4

(Jones et al., 2012) provides the SAT component, and HadSST3 (Kennedy et al., 2011b-c) provides SST. CRUTEM points are located over land, along coasts, and on some islands. HadSST3 data are over the open ocean, large bodies of water, and coasts. Therefore, there are some grid points, specifically islands and coasts, that are a blend of SAT and SST. The HadCRUT4.6 ensemble members are generated by sampling various error perturbations that represent characteristics of possible measurement bias and bias adjustments at each grid point (Morris et al. 2012).

Table 2.1: Globally Averaged Observation Datasets			
Dataset Name	Analysis	Download date	Reference
HadCRUT.5.1.0.0 median	Global mean	04/01/2021	Morice et al. 2021
HadCRUT.4.6.0.0 median	Global mean	04/01/2021	Morice et al. 2012
HadCRUT.4.6.0.0, filled in with kriging	Global mean	04/01/2021	Cowtan and Way 2014
GISTEMP v4 1200 km	Global mean	04/01/2021	Lenssen et al. 2019
GISTEMP v4 250 km	Global mean	04/01/2021	Zhang et al. 2020
NOAAGlobalTemp v5	Global mean	04/01/2021	Lenssen et al. 2019

Table 2.1: *Globally Averaged Observation Datasets*
The globally averaged observation datasets used in this study. The data in this table are downloaded through the KNMI climate explorer website, already globally averaged.

Table 2.2: Gridded Global Observation Ensembles			
Dataset Name	Analysis	Download date	Reference
HadCRUT.4.6.0.0 100 ensemble members	Global spatial analysis, Global mean, zonal mean, land masked, sea masked	04/01/2021	Morice et al. 2012
HadSST.3.1.1.0 median	Used as a land-sea mask	04/01/2021	Kennedy et al. 2011b-c
CRUTEM.4.6.0.0 median	Used as a land-sea mask	04/01/2021	Jones et al. 2012

Table 2.2: *Gridded Global Observation Ensembles*
The gridded global observation datasets used in this study. The data in this table are downloaded directly from the Met office website. All 100 ensemble members of the HadCRUT.4.6.0.0 as used in various analyses. The HadSST.3.1.1.0 median and CRUTEM.4.6.0.0 median are used to mask out SAT and SST components of the full HadCRUT.4.6.0.0 dataset. Each data set is on a 5° latitude by 5° longitude grid.

2.2. GCM Output

Output from three sets of general circulation model (GCM) runs will be analyzed in this study: the Max Planck Institute for Meteorology Grand Ensemble (MPI-GE) (Maher et al. 2019), the Coupled Model Intercomparison Project version 5 (CMIP5) (Taylor et al. 2012), and the Coupled Model Intercomparison Project version 6 (CMIP6) (Eyring et al. 2015). CMIP5 and CMIP6 are multi-model ensembles (Table 2.3 and 2.4). One ensemble member, the r1i1p1 and the r1i1p1f1, will be used from each individual model within the CMIP5 and CMIP6, respectively. The MPI-GE consists of 100 ensemble members run on a single coupled model (Table 2.5). The MPI-GE ensemble members are generated by initializing the historical simulation using output from successive years throughout its piControl experiment. Monthly output for the MPI-GE and CMIP5 will be from the historical simulations for the years 1890-2005 and from the RCP8.5 scenario for 2006 to 2019. The CMIP6 monthly output will be from the historical simulation for 1890-2014 and from the ssp585 scenario for 2015-2019. There are two standard climate model variables used in this study: tas and ts. Each model ensemble member used in this study contains both variables. The variable tas is the 2-meter near-surface air temperature, and ts is the surface temperature (sometimes called skin temperature). Both variables are outputted from the atmospheric realm of the coupled model over the entire globe. In order to isolate the near-surface air temperature over land, tas will be analyzed at grid points containing no ocean. To evaluate sea surface temperature over the ocean, ts will be used in grid points containing no land. For comparison purposes, tas will also be analyzed on the same ocean grid points as ts.

Throughout the study, SAT and tas will be referred to as synonyms, as well SST and ts. The only exception is for analysis when no masking is applied to the observations, then the model variable tas is used with no masking.

Table 2.3: CMIP5 Model Output	
Modeling Center / Modeling Group	CMIP5 Model Name
Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia	ACCESS1-0 ACCESS1-3
Beijing Climate Center (BCC), China Meteorological Administration	bcc-csm1-1
College of Global Change and Earth System Science, Beijing Normal University (BNU)	BNU-ESM
Canadian (Can) Centre for Climate Modelling and Analysis	CanESM2
National Center for Atmospheric Research (NCAR)	CCSM4
Community Earth System Model Contributors (CESM)	CESM1-CAM5
Centro Euro-Mediterraneo per I Cambiamenti Climatici (CMCC)	CMCC-CM CMCC-CMS CMCC-CESM
Centre National de Recherches Météorologiques (CNRM) / Centre Européen de Recherche et Formation Avancée en Calcul Scientifique	CNRM-CM5
Commonwealth Scientific and Industrial Research Organization (CSIRO) in collaboration with Queensland Climate Change Centre of Excellence	CSIRO-Mk3.6.0
EC-EARTH consortium	EC-EARTH
LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences and CESS,Tsinghua University	FGOALS_g2 FGOALS_s2
The First Institute of Oceanography (FIO), SOA, China	FIO-ESM
NOAA Geophysical Fluid Dynamics Laboratory (GFDL)	GFDL-CM3 GFDL-ESM2G GFDL-ESM2M
NASA Goddard Institute for Space Studies (NASA GISS)	GISS-E2-H GISS-E2-H-CC GISS-E2-R GISS-E2-R-CC
National Institute of Meteorological Research/Korea Meteorological Administration	HadGEM2-AO
Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)	HadGEM2-CC HadGEM2-ES
Institute for Numerical Mathematics	inmcm4
Institut Pierre-Simon Laplace (IPSL)	IPSL-CM5A-LR IPSL-CM5A-MR IPSL-CM5B-LR

Table 2.3: CMIP5 Model Output Continued	
Modeling Center / Modeling Group	CMIP5 Model Name
Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	MIROC5
Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	MIROC-ESM MIROC-ESM-CHEM
Max-Planck-Institut für Meteorologie (Max Planck Institute for Meteorology) (MPI)	MPI-ESM-LR MPI-ESM-MR
Meteorological Research Institute, Japan	MRI-CGCM3
Norwegian Climate Centre	NorESM1-M NorESM1-ME

Table 2.3: *CMIP5 Model Output*

The historical + rcp8.5 forcing experiment is used for CMIP5 models. One ensemble member from each model is selected; the variant IDs are r1i1p1 for CMIP5. The model variables tas (2 meter Near-Surface Air Temperature) and ts (Surface Temperature) will be used for each single ensemble member for all models listed. Ensemble members have various global grid sizes (not listed here). Data masking, such as land and sea grid points, are generated based on the model resolution, then applied before computing global and latitudinal averages.

Tale 2.4: CMIP6 Model Output	
Modeling Center / Modeling Group	CMIP6 Model Name
Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia	ACCESS-CM2 ACCESS-ESM1-5
Alfred Wegener Institute (AWI), Helmholtz Centre for Polar and Marine Research	AWI-CM-1-1-MR
Beijing Climate Center (BCC), China Meteorological Administration	BCC-CSM2-MR
Chinese Academy of Meteorological Sciences (CAMS)	CAMS-CSM1-0
Chinese Academy of Sciences (CAS)	CAS-ESM2-0
Community Earth System Model Contributors (CESM) - Whole Atmosphere Community Climate Model (WACCM)	CESM2-WACCM
Community Integrated Earth System Model Contributors	CIESM
Centro Euro-Mediterraneo per I Cambiamenti Climatici (CMCC)	CMCC-CM2-SR5 CMCC-ESM2
Canadian (Can) Centre for Climate Modelling and Analysis	CanESM5

Tale 2.4: CMIP6 Model Output Continued	
Modeling Center / Modeling Group	CMIP6 Model Name
Energy Exascale Earth System Model (from CESM)	E3SM-1-1
EC-EARTH consortium	EC-Earth3 EC-Earth3-CC EC-Earth3-Veg EC-Earth3-Veg-LR
LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences and CESS,Tsinghua University	FGOALS-f3-L FGOALS-g3
The First Institute of Oceanography (FIO), SOA, China	FIO-ESM-2-0
NOAA Geophysical Fluid Dynamics Laboratory (GFDL)	GFDL-ESM4
Indian Institute of Tropical Meteorology Pune	IITM-ESM
Institute for Numerical Mathematics	INM-CM4-8 INM-CM5-0
Institut Pierre-Simon Laplace (IPSL)	IPSL-CM6A-LR
Korea Institute of Ocean Science and Technology, Republic of Korea	KIOST-ESM
Atmosphere and Ocean Research Institute (University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science & Technology	MIROC6
Max-Planck-Institut für Meteorologie (Max Planck Institute for Meteorology) (MPI)	MPI-ESM1-2-HR MPI-ESM1-2-LR
Meteorological Research Institute, Japan	MRI-ESM2-0
Nanjing University of Information Science and Technology Earth System Model v3	NESM3
Norwegian Climate Centre	NorESM2-LM NorESM2-MM
Taiwan Earth System Model	TaiESM1

Table 2.4: *CMIP6 Model Output*

The historical + ssp585 forcing experiment is used for CMIP6 models. One ensemble member from each model is selected; the variant IDs are r1i1f1p1 for CMIP6. The model variables tas (2 meter Near-Surface Air Temperature) and ts (Surface Temperature) will be used for each single ensemble member for all models listed. Ensemble members have various global grid sizes (not listed here). Data masking, such as land and sea grid points, are generated based on the model resolution, then applied before computing global and latitudinal averages.

Table 2.5: MPI-GE Model Output		
Modeling Center / Modeling Group	Model Name	Reference
Max-Planck-Institut für Meteorologie (Max Planck Institute for Meteorology) (MPI)	MPI Grand Ensemble (MPI-GE)	Maher et al., 2019

Table 2.5: *MPI-GE Model Output*

The historical + rcp8.5 forcing experiment is used for the MPI-GE ensemble members. There are 100 ensemble members available, each initialized with different years from the MPI-GE piControl run. The model variables tas (2 meter Near-Surface Air Temperature) and ts (Surface Temperature) will be used for each of the 100 ensemble members. MPI-GE is on a 1.875° latitude by 1.875° longitude grid.

3. ANALYSIS METHODS

3.1. Data Masking

There are three types of data masking done in this study. The first method is specific to the HadCRUT4.6 dataset and its components. The HadCRUT4.6 dataset consists of 100 ensemble members. The second uses the Python library "global_land_mask", which is generated for and applied to all data types in section 6. A third level of masking is used in a subsection of section 6, which is the process of masking MPI-GE large ensemble dataset to only contain points that are available from HadCRUT4.6.

The HadCRUT4.6 dataset contains some grid points of SAT, some of SST, and some of blended temperature value of SAT and SST measurements (Jones et al. 2012). SAT data comes from the CRUTEM4 dataset, and SST data comes from the HadSST3 dataset. One-hundred ensembles of CRUTEM4 and HadSST3 are created and merged together to create the HadCRUT4.6 100 ensemble members merged dataset. Figure 3.1.1 show the spatial contribution of SAT and SST to the full HadCRUT4.6 dataset; the points labeled "Overlap" denote points in which the temperature at that grid box is a blend of SAT and SST. The 100 individual members from CRUTEM4 are not available for public use. In order to assess the seasonal warming trend of SAT over land, any grid points with SST data or a blend of SAT and SST (overlap) data will be excluded (labeled "No-HadSST3"). Similarly, to investigate the seasonal warming trend of SST over ocean, any points containing SAT individually or SAT data blended with SST will be excluded (labeled "No-CRUTEM4"). In other words, the entire HadSST3 spatial

coverage will be masked out to assess SAT over land, and the entire CRUTEM4 spatial coverage will be masked out to assess SST over ocean. The resulting data that is analyzed for SAT and SST is displayed in Figure 3.1.1 as the red region and the blue region, respectively.

The method of using CRUTEM4 and HadSST3 to mask out land and sea grid points to the HadCRUT4.6 dataset will be compared to the primary method of land-sea masking done in this study, which uses the GLOBE dataset (Global Land One-kilometer Base Elevation) (Globe Task Team, 1999). The information provided by the GLOBE dataset is accessed using the Python library "global_land_mask" (Karin, 2020). Although using CRUTEM4 and HadSST3 as a land-sea mask for observations classifies almost all the same points as using GLOBE to create a land-sea mask, the GLOBE generated mask will be used for consistency with models, and it is more accurate from a geographic standpoint as to what is considered land and what is considered sea. To generate a land-sea mask for a gridded surface temperature dataset (either observations or model output), each grid box is subdivided into 100 sub-grid boxes. For example, this produces sub-grid boxes of 0.5° by 0.5° for the HadCRUT4.6, which is on a 5° by 5° grid. Using geographic information from GLOBE, each of the 100 sub-grid boxes is marked as land or ocean. For the grid box at the original resolution to be marked as 100% land, all the sub-grid boxes must be labeled as land. Similarly, if all the sub-grid boxes are labeled ocean, then that grid box is considered 100% ocean on the native resolution. The notation for a dataset that has been masked to only include 100% land grid boxes uses the suffix "_Land", and when masked to include only 100% ocean

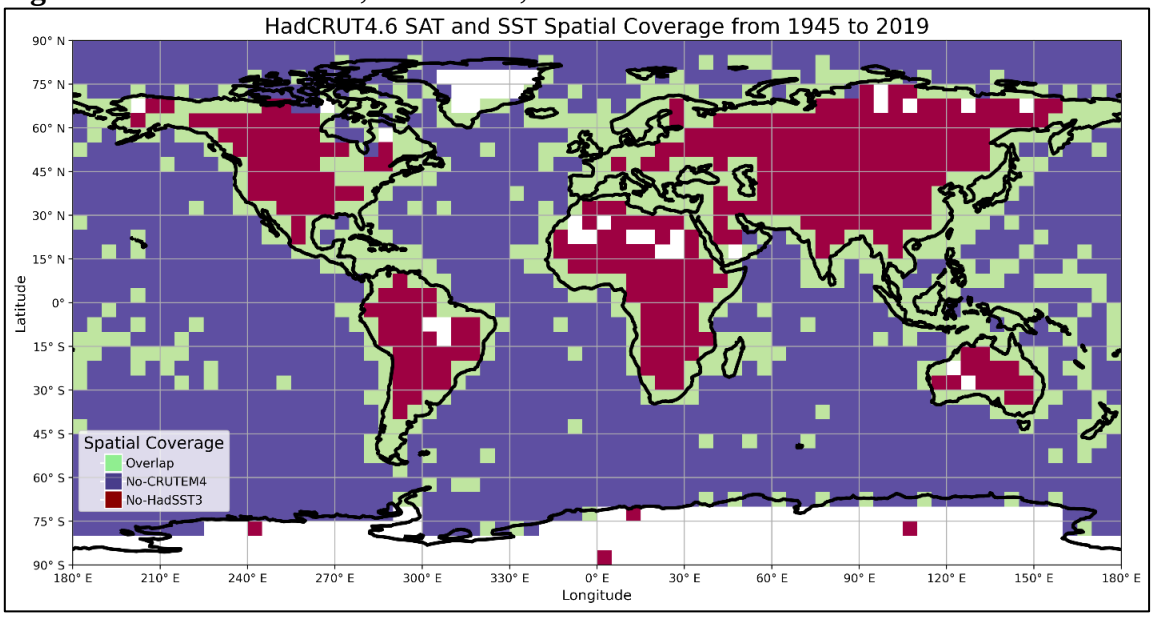
points, the suffix is "_Sea". Figure 3.1.2 demonstrates which of the HadCRUT4.6 grid points are added or removed to the collection of grid points classified as land-only and sea-only points by using the 100% land and 100% sea masks generated from GLOBE compared to using HadSST3 and CRUTEM4 as the land-sea mask. Many of the points located on islands that are considered "Overlap" points of SAT and SST are added to the realm of 100% ocean (added by HadCRUT4.6_Sea), and points along ice regions are removed (subtracted by HadCRUT4.6_Sea). The change in spatial coverage of grid points labeled as land-only by the 100% land mask from GLOBE (aka HadCRUT4.6_Land) compared to using CRUTEM4 (aka "No-HadSST3") overlap is most noticeable in the mid and high latitude land regions.

Land-sea masking based on the GLOBE dataset via the Python library will be generated for the MPI-GE, and each model dataset in CMIP5 and CMIP6's unique grid resolution in the same sub-grid box partitioning as described above. Shown in Figure 3.1.3 is the land-sea mask spatial coverage at the MPI-GE spatial resolution (1.875° by 1.875°). The notations "MPI-GE_Land" and "MPI-GE_Sea" represents the grid boxes that are considered 100% land or 100% sea based on the MPI-GE's native resolution.

Lastly, an additional level of masking will be applied to the MPI-GE; the models' grid boxes will be masked to match only where there is available observation data from the HadCRUT4.6 data. This "available observations" mask that is applied to the model is time dependent, based on the available observations over time. The suffix "_ObsMask" indicates that the available observation mask has been applied, using the

spatial coverage of HadCRUT4.6_Land, HadCRUT4.6_Sea, or all available HadCRUT4.6 grid points over time.

Figure 3.1.1: *HadCRUT4.6, CRUTEM4, and HadSST3*



Grid boxes with available data in the HadCRUT4.6 dataset from 1945 to 2019. HadSST3 provides SST, and CRUTEM4 provides SAT values. The spatial coverage of HadSST3 is of the blue region + the green region. The spatial coverage of CRUTEM4 is of the red region + the green region. Therefore, the green region is a blend of SAT and SST and is labeled “Overlap”. The red region is labeled “No-HadSST3”, as it is the CRUTEM4 region minus the overlapping points from HadSST3. The blue region is label “No-CRUTEM4”, and is the HadSST3 region without any overlapping points from CRUTEM4.

Figure 3.1.2: *HadCRUT4.6 Land and Sea Classification*

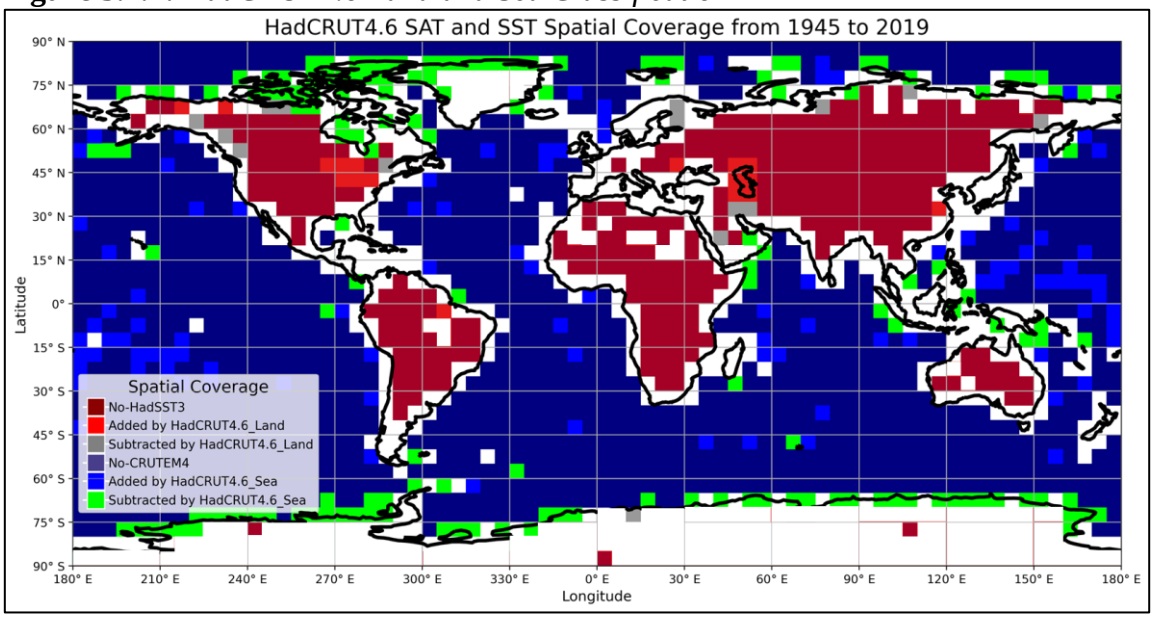
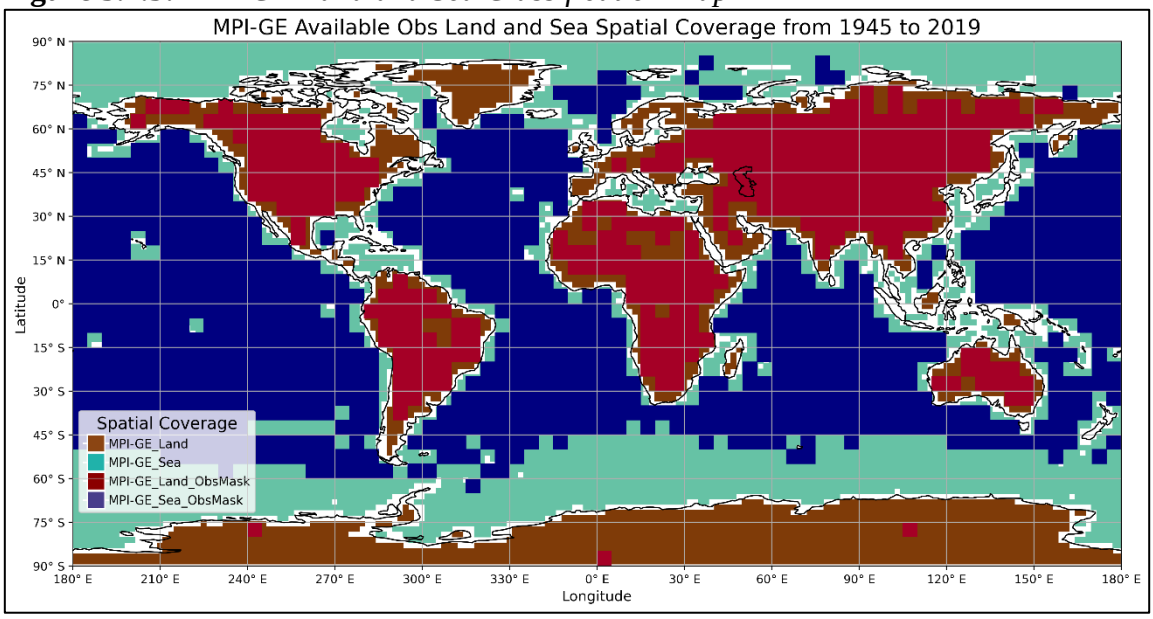


Figure 3.1.2: Two methods of land-sea masking are compared. No-HadSST3 denotes grid boxes of the HadCRUT4.6 that do not contain SST, and No-CRUTEM4 denotes grid boxes that do not contain SAT. A 100% land and 100% ocean mask are created specifically for the HadCRUT4.6 grid using the GLOBE geographic dataset. The light red and grey grid boxes are what is added and subtracted by the GLOBE 100% land points with respect to the No-HadSST3 spatial coverage. The light blue and green grid boxes are what is added and subtracted by the GLOBE 100% sea points with respect to the No-CRUTEM4 spatial coverage.

Figure 3.1.3: *MPI-GE Land and Sea Classification Map*



The spatial coverage of land-only and sea-only data points generated for the MPI-GE dataset using GLOBE to classify what points are 100% land and 100% sea. Additionally, the HadCRUT4.6_Land are HadCRUT4.6_Sea generated from GLOBE are overlaid on top of the MPI-GE grids. This is used when analyzing the model only where observations are available. This additional level of masking to the model is denoted by the suffix "_ObsMask".

3.2. Data Aggregation and Trend Calculation

Analysis of trends takes place at individual grid points, aggregated within zonal bands, or aggregated over the entire globe. Aggregation and trend calculation proceed as follows. First, if land or ocean data only is desired, a land-sea mask as described in section 3.1 is applied to the gridded dataset. The available observation mask is also applied at this stage to the MPI-GE. Then, all monthly temperature values in a given month and zonal band, or globally, are spatially averaged together, taking into account the latitudinal dependence of grid box size. Next, two 30-year climatologies are

computed from the monthly time series for the zonally averaged region. The 30-year climatologies used in this study are 1) 1890-1919, 2) 1945-1974, and 3) 1990-2019. Subtracting climatology (1) from (3) results in the monthly mean surface temperature trend for each month over 100 years. Similarly, climatology (3) minus (2) results in the monthly mean surface temperature trend for each month over 45 years. The resulting 12 monthly values of decadal surface temperature change follow an annual cycle, with the magnitude of temperature change between the two time periods varying by month (see the solid red line in figure 1.1.1, reproduced in 3.3.1).

3.3. Fourier Analysis

The phase and amplitude of the annual cycle of monthly surface temperature trends can be computed using a Fourier analysis. A Fourier analysis is the analysis of a periodic function into its simple sinusoidal components, whose sum forms a Fourier series. To isolate the seasonal cycle, a Fourier series of n=1 will be fitted to the cycle of monthly warming trends. The resulting equation is:

$$\Delta T = a_0 + a_1 \cos t + b_1 \sin t + C ,$$

where ΔT is the change in monthly temperature between the two time periods, a_0 is the mean annual warming between the two time periods, a_1 and b_1 are the amplitude of the annual cycle warming trend in sine and cosine components, and C represents the higher order terms that are not quantified on an annual frequency. The components of the amplitude of the annual cycle warming trend from the above equation are used to calculate the annual amplitude as follows:

$$r = \sqrt{{a_1}^2 + {b_1}^2} \ .$$

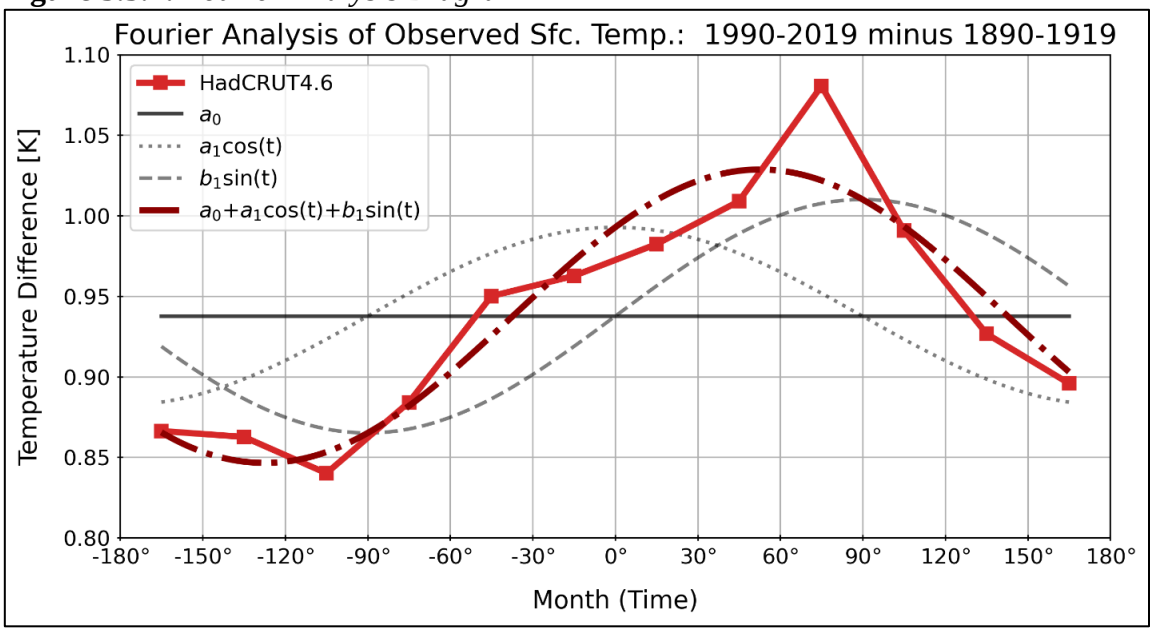
This amplitude *r* represents the magnitude of monthly temperature change relative to the average annual temperature change (a_0) between the two time periods. In other words, it is the amplitude of the equation for ΔT and it is seen plotted in figure 3.3.1. The annual phase (θ) of monthly temperature trends are calculated from ΔT equation by:

$$\theta = \cos^{-1}(a_1/r) \ \wedge \ 0 > \theta > \pi \text{ , then,}$$

$$b_1 > 0 \rightarrow 0 < \theta < \pi \ \vee \ b_1 < 0 \rightarrow 0 < \theta < -\pi \, .$$

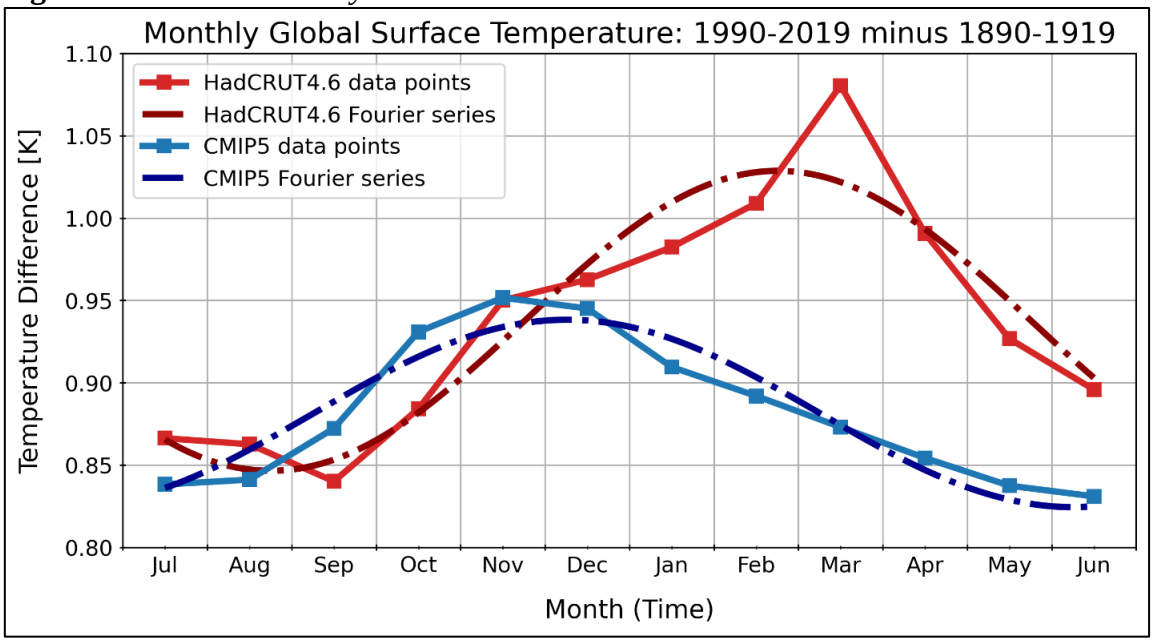
There are 12 data points per year that are used to calculate the equation, so each month corresponds to a π/6, or 30°, increment in phase. It is defined as such that a 0 to π/6 phase angle represents a peak warming month of January (see the x-axis in figure 3.3.1). Resolving the annual cycle of surface temperature warming with a Fourier series, instead of seasonal linear trends, identifies the timing and magnitude of warming variations between the two time periods with respect to an annual cycle in a continuous fashion rather than in discrete seasons or months (see an example of this using the data from the introduction in figure 3.3.2).

Figure 3.3.1: *Fourier Analysis Diagram*



The average monthly surface temperature for 1990-2019 minus 1890-1919 for HadCRUT4.6, along with its resulting Fourier series and individual components.

Figure 3.3.2: *Fourier Analysis on Global Mean Trends*



The average monthly surface temperature for 1990-2019 minus 1890-1919 for HadCRUT4.6 and CMIP5 model mean, along with each of their resulting Fourier series.

3.4. Use and Interpretation of Ensembles

The HadCRUT4.6 observation ensemble allows for quantification of structural uncertainty due to measurement error and bias, and choice of analysis scheme (Jones et al. 2012). It doesn't necessarily represent uncertainty that comes from the use of different sources for SAT or SST data, uncertainty due to incomplete observations, and doesn't cover every possible choice of analysis scheme. The observed record represents one of the many possible chains of atmospheric events that could have occurred within the bounds of natural climate variability. The spread of the observation ensemble around the ensemble mean represents much of the uncertainty in the ability to measure the actual temperature trends from the observed temperature record.

The ensemble spread of the MPI-GE is a measure of natural climate variability of the climate model. The ensemble members of the MPI-GE historical simulation are initialized by sampling the climate state in various decades throughout its pre-industrial control run (Maher et al. 2019). This allows the historical simulation to follow random chains of atmospheric events within the bounds of the MPI-GE climate model. The CMIP5 and CMIP6 multi-model ensembles sample natural variability from the various models that make up the entire ensemble. Additionally, CMIP multi-model ensembles sample climate model structural uncertainty due to the many ways these various models are constructed to simulate the Earth's climate system. Using a large ensemble of climate model simulations yields a robust estimate of the forced response (ensemble mean) and natural variability (ensemble spread) (Lehner et al., 2020). It also enables

quantitative estimation of structural uncertainty in the forced response, but does not reduce that uncertainty.

3.5. Ensemble Calculations and Statistics

To calculate an ensemble's mean annual cycle of surface temperature trends, the cosine $a_{1,i}$ and sine $b_{1,i}$ components for each ensemble member i for all ensemble members N are averaged together, such as the following:

$$a_{1,\mu} = \frac{\sum_{i=1}^{N} a_{1,i}}{N} \text{ and } b_{1,\mu} = \frac{\sum_{i=1}^{N} b_{1,i}}{N}.$$

This results in $a_{1,\mu}$ and $b_{1,\mu}$, which can be used to calculate ensemble mean phase and amplitude of temperature trends. To calculate the spread of an ensemble, the following root-mean-squared (RMS) Euclidian distance between ensemble members and ensemble mean is calculated:

$$RMS\ distance = \sqrt{\left|a_{1,i} - a_{1,\mu}\right|^2 + \left|b_{1,i} - b_{1,\mu}\right|^2}\ .$$

This RMS distance is used in section 5.2. In that section, the MPI-GE model ensemble and the HadCRUT4.6 observation ensemble are analyzed spatially on their native grid resolutions. Thus, the models and observations are not quantitively compared on a grid-point by grid-point basis. In sections 5.1, and section 6.1-6.3, the zonal averages of the models and observations are compared to one another using Mahalanobis distance.

The Mahalanobis distance is a statistic that is used in this study to determine whether the observed temperature trends are statistically indistinguishable from the distribution of model ensemble member trends. Rather than Euclidean distance, the

Mahalanobis distance takes into account covariance of the dimensions of the distribution. The following equation is a matrix calculation, in which the diagonal of $Z_{(1,2)}^2$ is the Mahalanobis distance between a datapoint x_1 taken from distribution *1*, and a second distribution *2* of datapoints that have a mean of μ_2 and an inverse covariance matrix Cov_2^{-1}:

$$Z_{(1,2)}{}^2 = (x_1 - \mu_2)' \cdot \mathrm{Cov}_2{}^{-1} \cdot (x_1 - \mu_2),$$ (Sarma and Vardhan, 2019).

The resulting Mahalanobis distance $Z_{(1,2)}^2$ will have dimensions equal to the degrees of freedom of the distributions *1* and *2*. In this study there are two variables, the cosine amplitude and sine amplitude (a_1 and b_1) of the annual cycle of surface temperature trends; therefore there is one degree of freedom. The Mahalanobis distance statistic follows a chi-squared distribution. To reject the null hypothesis that the datapoint x_1 is from distribution *2*, the Mahalanobis distance $Z_{(1,2)}^2$ needs to be greater than chi-squared value for one degree of freedom. This is converted to a p-value, and tested using three significance levels (0.1, 0.01, and 0.001).

Specifically in this study, the Mahalanobis distance from the HadCRUT4.6 ensemble datapoints (distribution *1* in the above equation) to the distribution of the model ensemble (distribution *2*) are calculated. Graphs in sections 5 and 6 that compare models to observations have one asterisk "*" to signify a p-value < 0.1, two asterisks "**" to signify a p-value < 0.01, and three asterisks "***" to signify a p-value of < 0.001.

4. CONCEPTUAL MODEL

4.1. Conceptual Model

There are many possible drivers of decadal-scale changes to the phase and amplitude of the seasonal cycle of surface temperature. Three types of drivers of have been identified by previous studies. These are: orbital effects, human-induced climate change, and natural climate variability (Figure 4.1). The annual cycle of surface temperature is directly modified by orbital changes such as axial precession while the other two drivers alter the seasonal cycle through a series of interconnected forcing mechanisms. In this section, the mechanisms of those drivers on the timing and spatial characteristics of the seasonal warming cycle are considered.

Axial precession is the cyclical movement of the orientation of earth's rotational axis. It takes about 26,000 years for the earth's rotational axis to rotate through one full cycle. Over the course of the cycle, the timing of the solstices steadily lags relative to the orbit's perihelion. This slow progression of the perihelion affects the timing of maximum incoming solar radiation each year, thereby affecting the timing of the annual maximum surface temperature (Thompson 1995). Presently, the perihelion occurs in the beginning of July, and the aphelion in the beginning of January. The perihelion and aphelion drift forward in time, which causes seasonal minima and maxima to occur later than would be defined by the solstices. This places the month of peak warming in October (between perihelion and aphelion), and unlike seasonal warming trends driven by other forcing mechanisms, the trend in the annual cycle of surface temperature occurs in the same month for both hemispheres. The relationship between axial precession and

changes to the seasonal cycle of surface temperature is indicated in the top of figure 4.1.1.

Thompson (1995) verified that in central England, for the last few centuries, the annual phase of surface temperature drifted earlier, as expected from precession. However, around the early part of the 20th century towards the beginning of the industrial revolution, the phase of the annual cycle no longer drifted according to precession of the solstices. Instead, the phase trend began to drift in the opposite direction, which Thompson attributed to the rising CO_2 levels.

Anthropogenic emissions of greenhouse gases such as CO_2 increase the global mean surface temperature by decreasing the outgoing longwave radiation from Earth to space. Direct emission of aerosols or their precursors affect the global mean surface temperature by decreasing the amount of incoming shortwave radiation to reach the surface, causing a relative decrease in temperature. Increased greenhouse gas emissions and aerosol emissions primarily alter the global mean surface temperature, but the various positive climate feedbacks induced by these emissions are what drives changes to the annual cycle of surface temperature (see right-hand side of Figure 4.1.1).

Two climate feedback mechanisms involving surface albedo are the sea-ice feedback and snow feedback. First, an increase in the annual average surface temperature would result in less sea-ice and less snow cover annually. An increase in annual temperature could be a long term forced trend due to anthropogenic emissions, but it could also happen from natural variability. Less area of sea-ice results in more ocean surface area exposed, which decreases surface albedo and drives surface air

temperatures to be at or above 0 °C. Sea-ice is present year-round, but the change in sea ice from preindustrial levels is greatest in early autumn, the annual minimum of sea-ice levels. Therefore, sea-ice feedback could moderate the seasonal cycle in two ways. Having more ocean area exposed in the autumn would result in a later transition to winter. This puts the month of peak warming trend around September for the Northern Hemisphere, March for the Southern Hemisphere. Alternatively, but not exclusively, having more ocean area exposed year-round could result in in the greatest warming rate in mid-winter (January in the Northern Hemisphere, July for the Southern Hemisphere), when the temperature contrast between air located over sea ice and air over exposed ocean is greatest. The sea-ice albedo feedback primarily alters the seasonal cycle of temperature over polar regions in either hemisphere.

The snow feedback manifests in a similar fashion as sea-ice feedback, specifically for land regions in the Northern Hemisphere midlatitudes, but the main difference is that snow cover is not changing year-round. Because the largest trends in Northern Hemisphere snow cover decline are in boreal spring, a maximum warming trend in March or April is expected in the mid to high latitudes. While the Southern Hemisphere sea-ice experiences seasonal variations similar (but opposite) to the Northern Hemisphere, there isn't much capacity for the snow feedback to occur because there is a lot less land for snow to accumulate in the SH mid-latitudes.

An increase in annual global temperature increases the rate of evaporation, which then adds more water vapor into the air. Water vapor is a greenhouse gas, thus contributing to further increase in temperature. This process is also known as the water

vapor feedback loop, a positive feedback between global mean temperature and specific humidity. On regional scales, humidity and evapotranspiration are paired together because changes in humidity alter the evapotranspiration rate, and vice versa. Increased precipitation, together with increased evapotranspiration, imply a moderated seasonal temperature cycle (Stine et al. 2009, Nigam et al. 2017). In non-arid areas, evapotranspiration is largest in early summer, when vegetation is active, vapor pressure deficits are large, and incoming solar energy is large. An increase in evapotranspiration would result in a cooler summer, which translates to a warming peak in winter over land. The relationship between global mean temperature, humidity/evapotranspiration, and the seasonal cycle of temperature is indicated in Figure 4.1.1.

Changes in humidity/evapotranspiration can lead to changes in vegetation. An increase would result in more vegetation, which would decrease the surface albedo because vegetation tends to be darker than bare soil. Changes in vegetation could also occur directly from changes in the global mean surface temperature, if the altered climate is outside of local plant life's needs. Reduction in vegetation may release stored greenhouse gasses such as CO_2, which would contribute to additional global temperature increase. An increase in vegetation would amplify the seasonal cycle of absorption of solar radiation, thereby producing maximum warming in summer over land, opposite to the direct effect of increases in evapotranspiration.

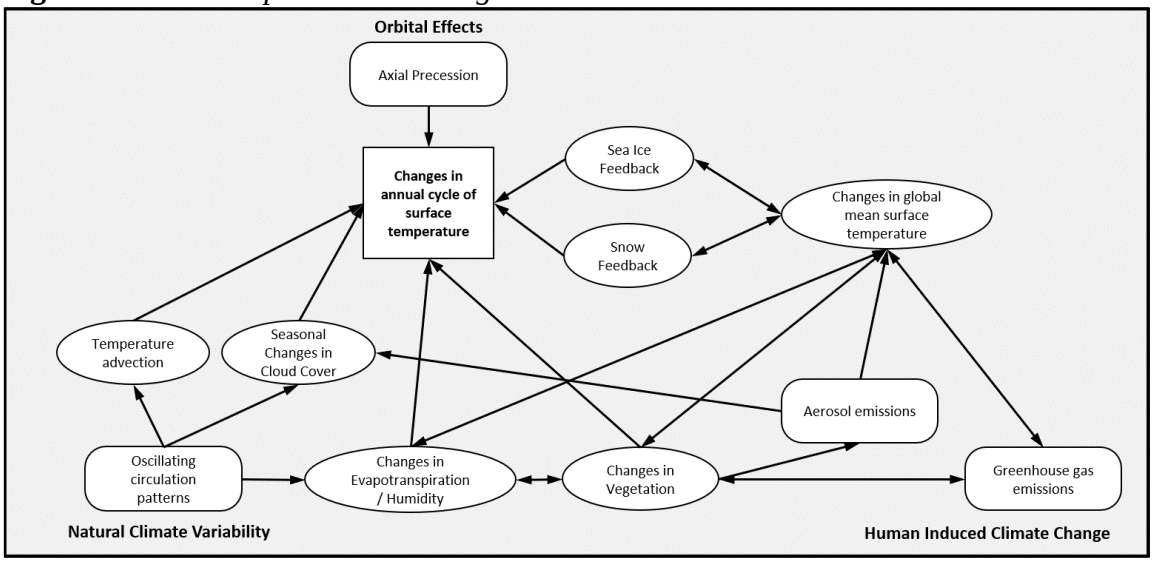
Changes in vegetation could also release aerosols, such as through biomass burning. Modified aerosol concentrations from any cause can indirectly affect regional cloud cover, thus affecting the seasonal cycle of surface temperature (see Fig. 4.1.1).

These changes in the seasonal cycle due to modified aerosol concentrations, though non-negligible, would be difficult to trace on a global, decadal-averaged scale.

Oscillating modes of atmospheric-oceanic states affect the seasonal cycle of surface temperature (Mann and Park 1996, Wang et al. 2009, Stine and Huybers 2012). Natural variability and internal variability are terms used to describe the atmospheric-oceanic circulation patterns that temporally oscillate within the bounds of the earth's climate. Natural variability drives changes in large-scale wind patterns, thereby altering regional temperatures by advection. For example, a winter that is regionally warmer due to a La Niña would decrease the amplitude of the surface temperature cycle for a given year. More than one La Niña or El Niño can happen in a row because ENSO is not perfectly periodic. Slight changes in the initial conditions of a model simulation can result in a different series of ENSO events. This makes it impossible for climate models to replicate the specific sequence of observed natural variability. Seasonal cloud cover and precipitation are also affected by oscillating modes of natural variability. Clouds alter the regional radiative balance by reflecting shortwave radiation and trapping surface longwave radiation, thus affecting the temperature at the surface. The direct and indirect pathways for how natural variability can modify the seasonal cycle of surface temperature are outlined on the left-hand side of Figure 4.1.1. Globally and zonally averaged changes in natural variability and how this affects the seasonal cycle of surface temperature are difficult to identify, as these changes may appear in various seasons, regions, and magnitudes, depending on the climatic process and corresponding teleconnections, and averaging allows locally amplified and reduced warming trends to

at least partially cancel. Nonetheless, it is important to note that generally, natural variability will affect ocean and land together, and display a similar seasonal signal, compared to the land-sea contrast seen by other forcing mechanisms.

Figure 4.1.1: *Conceptual Model Diagram*



The drivers of changes in the annual cycle of surface temperature (rounded rectangles) and their forcing mechanisms (ovals).

5. GLOBAL ANALYSIS

5.1. Global Mean Analysis

The phase and amplitude of monthly global surface temperature trends are compared over two different lengths. All observation datasets are included in the global average analysis, in addition to all ensemble members of the MPI-GE, and one ensemble per model for all the models within the CMIP5 and CMIP6. No masking is applied to any of the datasets yet, so all land and sea points are included.

The purpose of analyzing of the 100-year trend (Figure 5.1.1) is to capture the longest available observational record across all datasets. The amplitude of the annual cycle of warming over this time period is about 0.05K according to HadCRUT4.6 ensemble mean. The other observation datasets (outlined in table 2.1), except for one, fall within the range of phase and amplitude of the HadCRUT4.6 ensemble spread. All the GCMs have a similar mean amplitude as the observed mean amplitude, though the spread is large. The multi-model ensembles (CMIPs) have larger range of amplitude than the MPI-GE. This indicates there is added uncertainty of the simulated amplitude when using a multi-model ensemble. The average month of peak warming for the GCMs ranges from early December to early January, which is 1.5 to 2.5 months earlier than the observed month of peak warming, late February. Simulated seasonal temperature trends over 100 years are significantly different from what has been recorded and analyzed, regardless of choice of observation dataset.

The 45-year trend (Figure 5.1.2) indicates the robustness of the surface temperature trends with respect to using an alternate time period to calculate the decadal

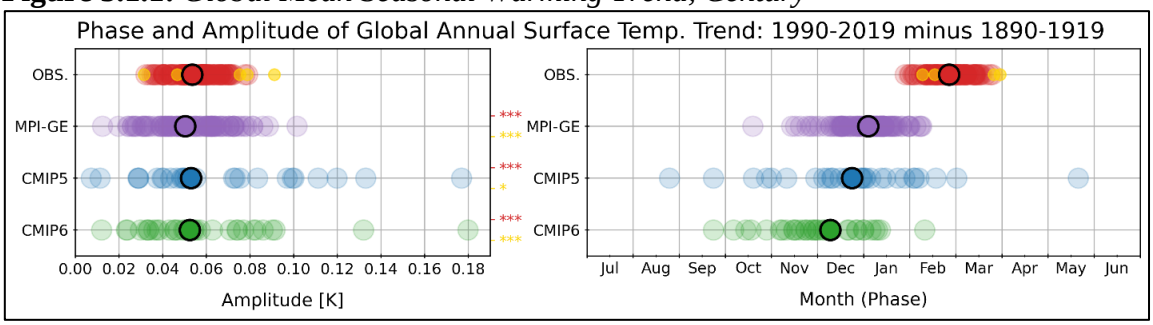
trend. The observed amplitude of seasonal warming for the 45-year trend is similar to the 100-year trend, but the spread across datasets is smaller. The other observation datasets veer towards a slightly earlier peak annual warming than is sampled by the HadCRUT4.6 ensemble. As for the GCMs, the spread is reduced in the 45-year trend, which slightly underestimates the observed monthly warming amplitude. All observation datasets, even those with earlier peak seasonal warming than the HadCRUT4.6 ensemble, are not within the same distribution as simulated by any of the model ensembles. The model simulated phase of the 45-year annual warming trend in this timespan is 3 to 4 months earlier than the observations and a month earlier than the 100-year trend.

Figure 5.1.3 is based on the same data as in 5.1.1 and 5.1.2, but the phase and amplitude components are plotted together on a polar diagram for each of the two seasonal trend analyses. The polar diagrams further emphasize the overlap, or lack thereof, between models and observations. On this globally averaged scale, there are only a few GCM ensemble members that simultaneously overlap the phase and amplitude of the observations for the 100-year trend. For the 45-year trend, the distance between observed and model monthly warming trends is greater than for the longer time period; virtually no models overlap the mean observed trend or the spread of observed trend estimates. The righthand side of figure 5.1.3 also demonstrates that, on a global scale, the latest generation of climate models (CMIP6) are phase shifted farther away from observations than the previous generation (CMIP5). For any of the model

ensembles over either timeframe, the observations from any dataset imply statistically different distributions than the model ensemble variability.

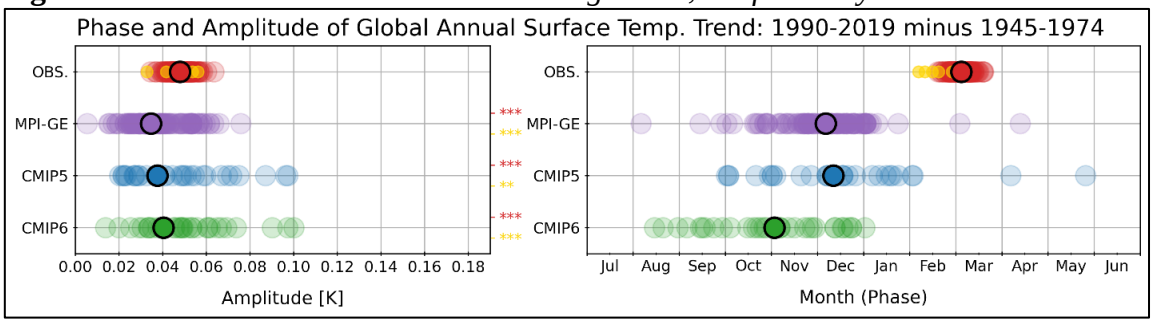
The results from this section give insight into the robustness of the global trend using two different time periods of record. The spatial Fourier analysis in the following sections uses the 1945-1974 and 1990-2019 (45-year) trends because it includes better spatial data coverage.

Figure 5.1.1: *Global Mean Seasonal Warming Trend, Century*



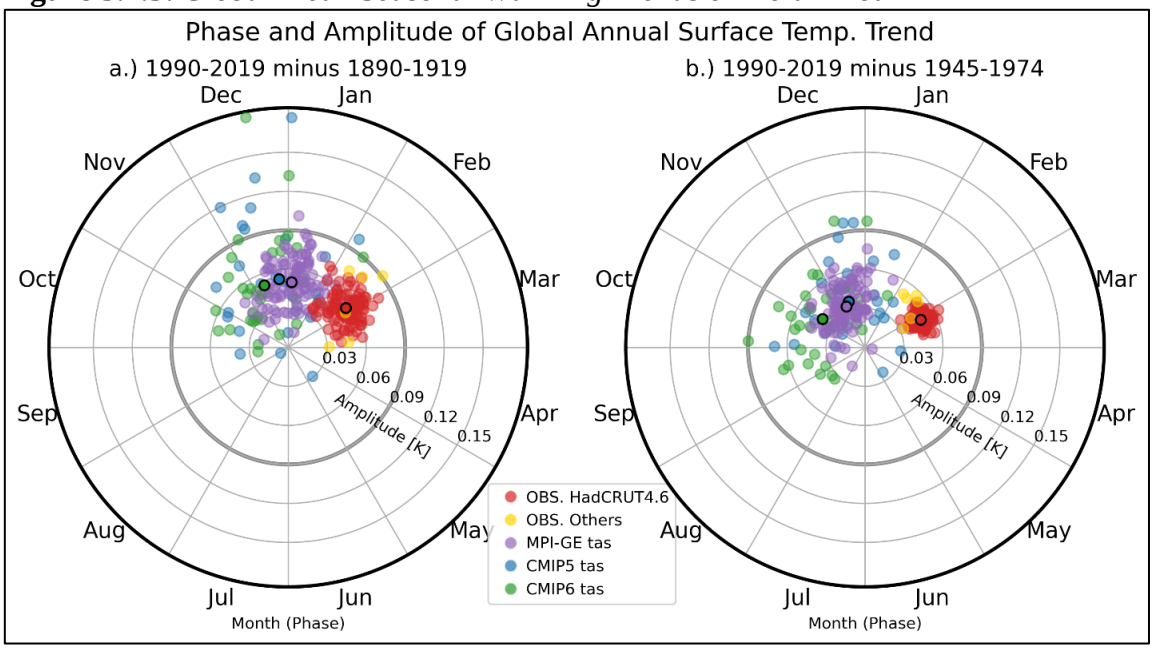
Amplitude (left) and phase (right) of the Fourier fit to the seasonal warming cycle from 1890-1919 to 1990-2019. Circles with the black outline are of the observation ensemble or model ensemble mean. OBS.: Yellow dots are observation datasets from Table 2.1, and red dots HadCRUT4.6 observation ensemble members. Model variable tas without masking is used for MPI-GE, CMIP5, and CMIP6. Red asterisks denote whether or not the HadCRUT4.6 ensemble is significantly different from the distribution of model ensemble members for each model ensemble, and the yellow asterisks are the same except the computation includes all the observations (Red and Yellow dots). The largest (least significant) p-value of these is shown.

Figure 5.1.2: *Global Mean Seasonal Warming Trend, Half Century*



Amplitude (left) and phase (right) of the Fourier fit to the seasonal warming cycle from 1945-1974 to 1990-2019. See further descriptive details in Figure 5.1.1's caption.

Figure 5.1.3: *Global Mean Seasonal Warming Trends on Polar Plot*



The left side (a) are the same data from figure 5.1.1 and the right side (b) are the same data from figure 5.1.2. The amplitude and phase of the Fourier fit to the seasonal warming cycle are represented by r and θ on the polar diagram. Circles with the black outline are of the observation ensemble or model ensemble mean.

5.2. Spatial Global Analysis

The previous section discussed the overall, global mean monthly surface temperature trend. This section investigates the spatial patterns of the monthly surface temperature trends across the globe, using the HadCRUT4.6 observation ensemble dataset and the MPI-GE gridded model output of surface air temperature (tas). The primary method is to analyze the seasonal cycle of surface temperature warming at individual grid points, starting with the phase and amplitude of HadCRUT4.6 ensemble mean in Figure 5.2.1 and Figure 5.2.2.

The most prominent seasonal warming feature is located over large land masses in the mid and high latitudes of the Northern Hemisphere (NH). Spanning 40°N to 80°N, located in inland regions of northern North America, Europe, and Asia, the amplitude of the monthly warming trend difference is approximately 1 K on average across these land regions (Fig 5.2.1). In these regions of high amplitude across NH land, the peak warming is primarily in late boreal winter, with some regions in the Asian high latitudes and North American mid latitudes reaching a seasonal maximum warming in early boreal spring (Fig 5.2.2). This phase of peak warming in NH spring aligns with the globally averaged trend displayed in Figure 5.1.2. Oceans in the NH extra-tropics have a smaller amplitude of monthly warming than surrounding land. However, there are a few local maxima of amplitude in this zone of about 0.5 K, with a phase of early boreal fall, which is essentially opposite the phase of observed warming over land.

In Southern Hemisphere (SH) non-tropical regions, the amplitude of seasonal warming is higher over the ocean than over land. There is more ocean surface area in the SH and more land area in the NH, so seasonal trends over land surfaces in the SH may be influenced (and thus dampened) by the ocean's seasonal temperatures and trends thereof. Regions of strong seasonal warming in the SH ocean peak at a similar time in the calendar as NH land.

Much of the tropics and SH land display little spatial pattern or zonal consistency about which season the maximum rate of warming occurs. This is primarily due to the data in this region exhibiting a low monthly warming amplitude, which yields highly variable and potentially unrepresentative grid point values of phase. The exception is

Northern Africa and the Middle Eastern land regions, where the amplitude is larger. The relative maximum in monthly warming in these subtropical regions exhibit a peak warming in boreal summer.

Figure 5.2.3 shows root mean squared distance at each grid point between the ensemble members and ensemble mean of HadCRUT4.6, which is a measure of the structural uncertainty in the HadCRUT analysis. The regions with large seasonality of warming amplitude pointed out in the previous paragraphs are co-located with the largest ensemble spread (largest RMS distance). However, it is important to note that the ensemble spread is at least one order of magnitude smaller than the ensemble mean amplitude in all regions of the globe. This is shown in figure 5.2.4 as the ratio of the amplitude of the ensemble mean to the ensemble spread. Other than a few scattered points, the ratio is greater than 2 almost everywhere, with globally averaged signal to noise ratio of approximately 20. This means that the ensemble mean amplitude of seasonal warming is consistent across all observation ensemble members and is not strongly affected by structural uncertainty.

The simulated amplitude and phase of seasonal surface temperature trends at each grid point of the MPI-GE ensemble mean are shown in figure 5.2.5 and 5.2.6. The MPI-GE displays higher amplitude over land than ocean in both hemispheres, with the strongest seasonal warming amplitudes at the poles. HadCRUT4 lacks data at very high latitudes, however the regions of large amplitude over NH continents do not extend as far southward into the mid-latitudes as it does in the observations. The season of peak warming in the Arctic according to the MPI-GE is mid-winter, which is a few months

sooner than in HadCRUT4.6 middle to high latitudes. In the SH, mid-latitude oceans do not exhibit the distinct seasonal warming trend that is featured in the observations. While there is a relative maximum over land in Africa and the Middle East, it is displaced southward by approximately 10 degrees latitudinally compared to observations. The ocean areas, both tropical and non-tropical, have a lot of spatial variability in the phase of the seasonal warming trend. This is primarily due to a very small seasonal warming amplitude over the ocean. It should be pointed out that this is the spatial plot of near-surface air temperature (tas model standard variable). In the following sections, after an appropriate land/sea mask has been applied, MPI-GE ocean regions will be analyzed using surface temperature (ts model standard name), to be more consistent with HadCRUT4.6.

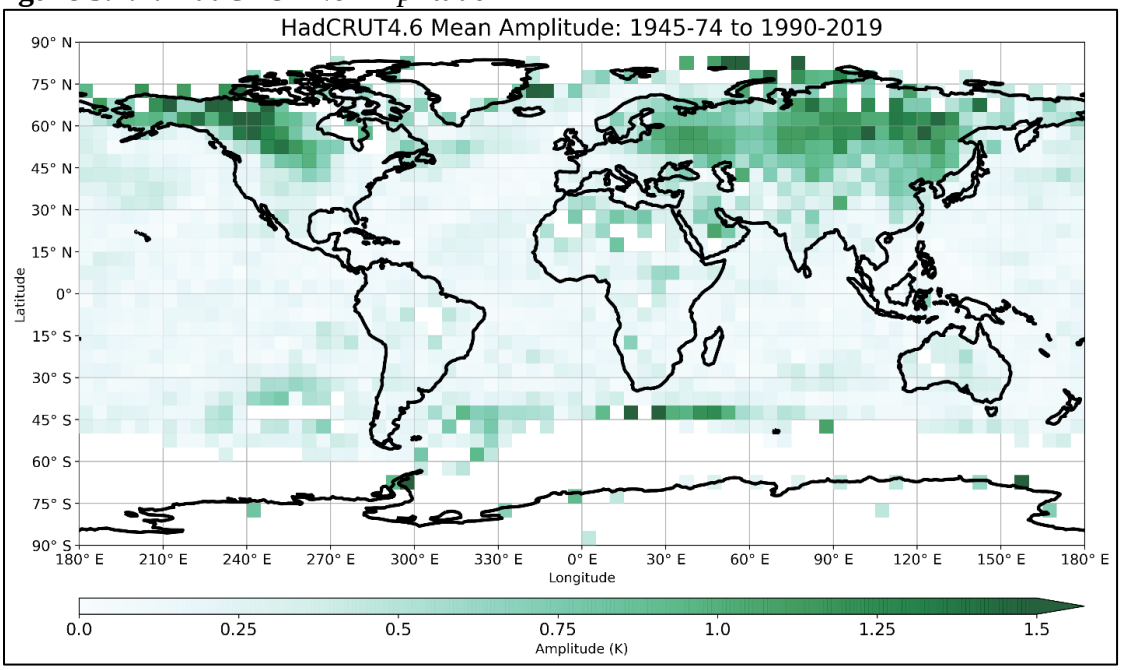
The ensemble spread in MPI-GE is on the same order of magnitude as the ensemble mean amplitude (Fig 5.2.7). The highest RMS distance of the model's ensemble members to the ensemble mean is in the high latitudes, particularly on the margins of sea ice. Figure 5.2.8 is the ratio of the model mean seasonal warming amplitude to the spread of the seasonal warming trends represented by the ensemble members. The ratio is less than one in a lot of areas of the globe, primarily due to annual warming trends having small or no seasonality in those areas while ensemble variability persists. There are a few notable areas where the model mean seasonal warming amplitude is greater than the magnitude of simulated natural variability, such as the Arctic and Antarctic oceans, regions of the southern tropical Pacific and northern

tropical Atlantic, as well as over land in parts of Africa, southern Asia, and western United States.

The spatial Fourier analysis across the globe appears to show a substantial climate change signal in many regions. Natural variability estimated with the MPI-GE ensemble is around three times smaller than the mean observed seasonal warming trend in regions where a strong seasonal trend has been observed (compare figure 5.2.1 and 5.2.7).

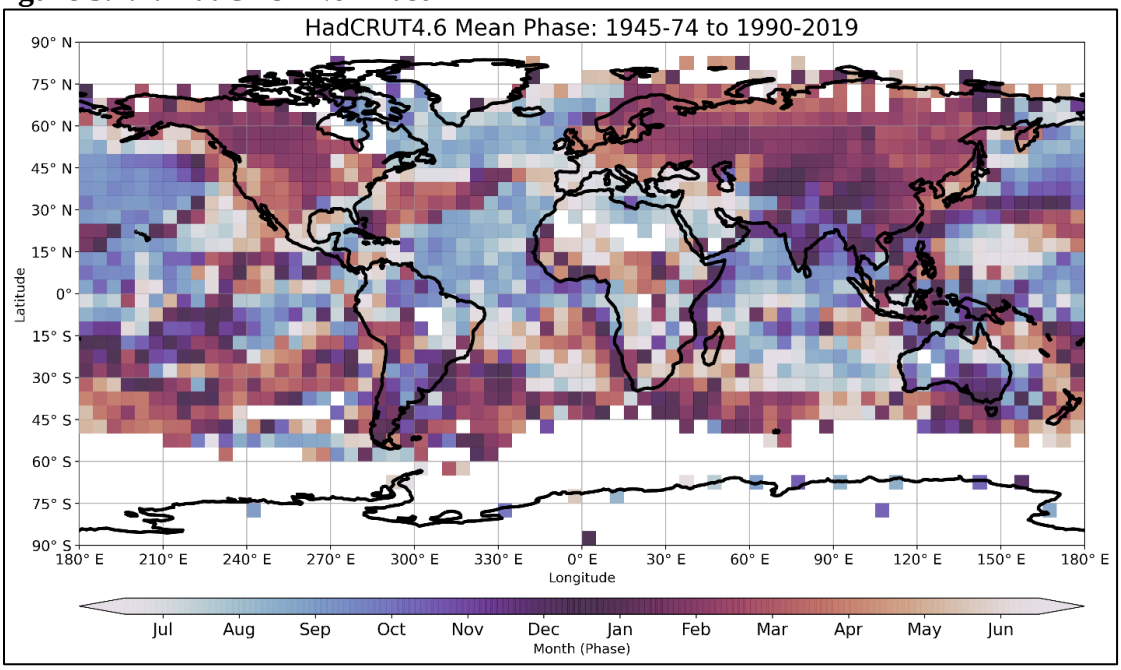
In order to reduce the noise from structural uncertainty and natural variability, the remaining analysis will aggregate the seasonal warming trends by zonally averaging surface temperature across 15-degree latitude bands. The zonal regions of distinct seasonal warming trends identified throughout this section range in area coverage size from 10° to 40° latitudinally. It is possible to aggregate some regional features together into 30° (or more) latitude bands, but 15° latitude bands were chosen so that smaller but distinct features were not lost while keeping the region large enough to reduce noise from natural variability in the models. Additionally, as has been done in previous studies, land and ocean surface temperature will be analyzed separately because the observed trends exhibit a substantial land vs. sea contrast, and are likely driven by different forcing mechanisms. The zonal aggregation of trends and separation of land and sea grid points will enable the identification of systematic differences across land and sea and between models and observations, and how these partitioned trends contribute to the global annual cycle of surface temperature warming.

Figure 5.2.1: *HadCRUT4.6 Amplitude*



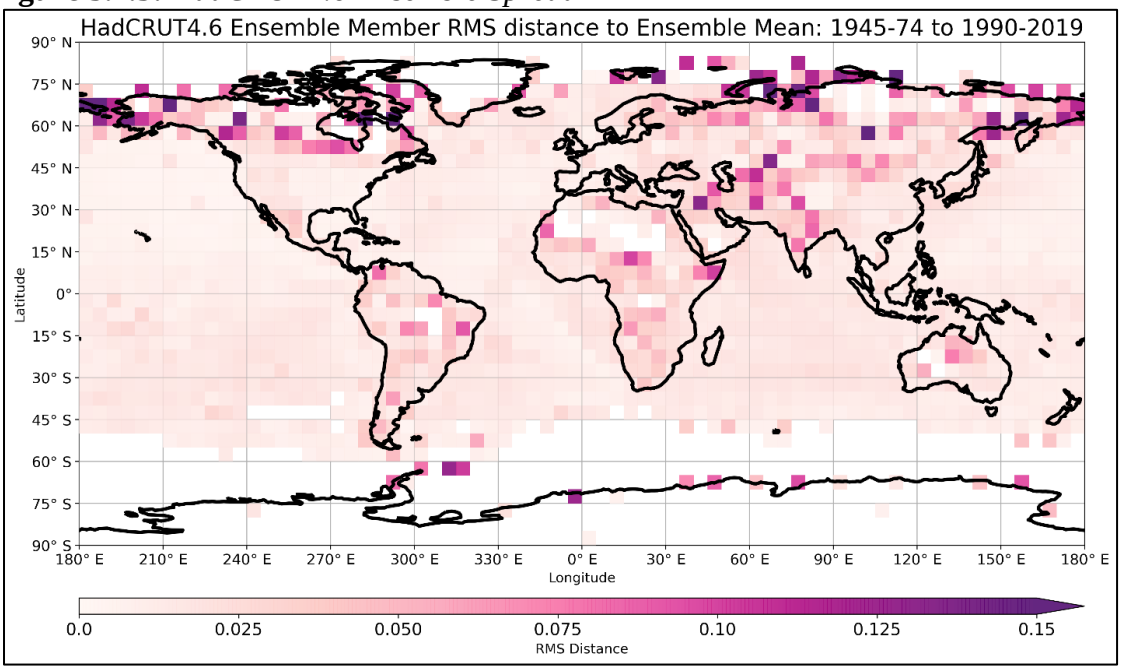
Amplitude of the Fourier fit to the seasonal warming cycle in the HadCRUT4.6 observation ensemble mean from 1945-1974 to 1990-2019.

Figure 5.2.2: *HadCRUT4.6 Phase*



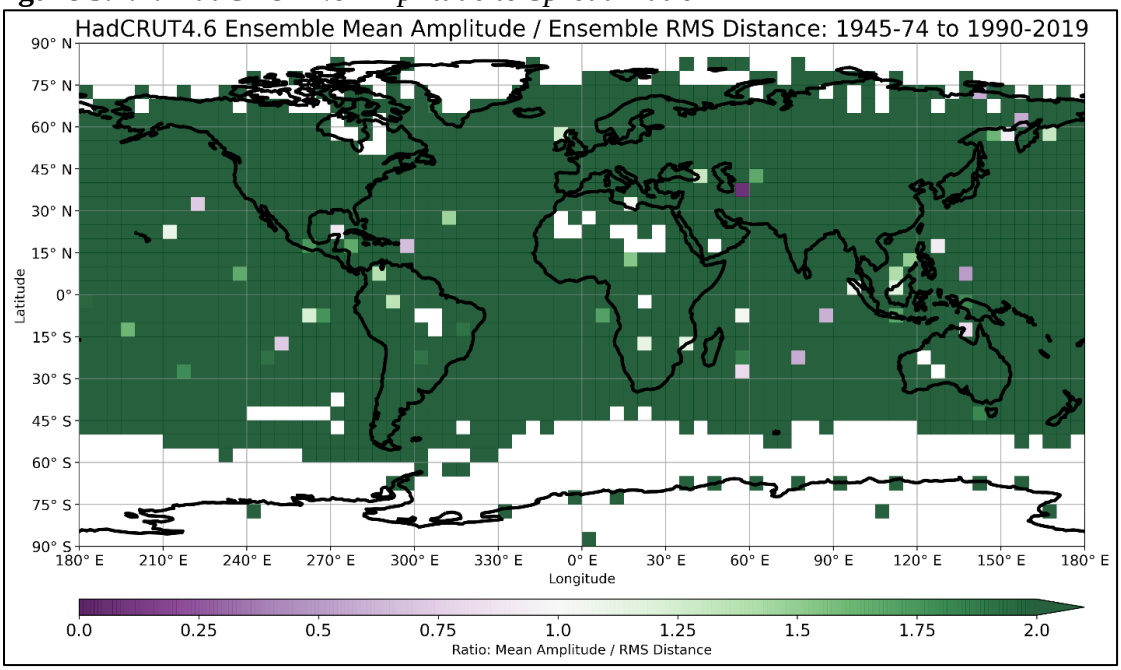
Phase of Fourier fit to the seasonal warming cycle in the HadCRUT4.6 observation ensemble mean from 1945-1974 to 1990-2019. The month of the peak of the Fourier fit is indicated by the color.

Figure 5.2.3: *HadCRUT4.6 Ensemble Spread*



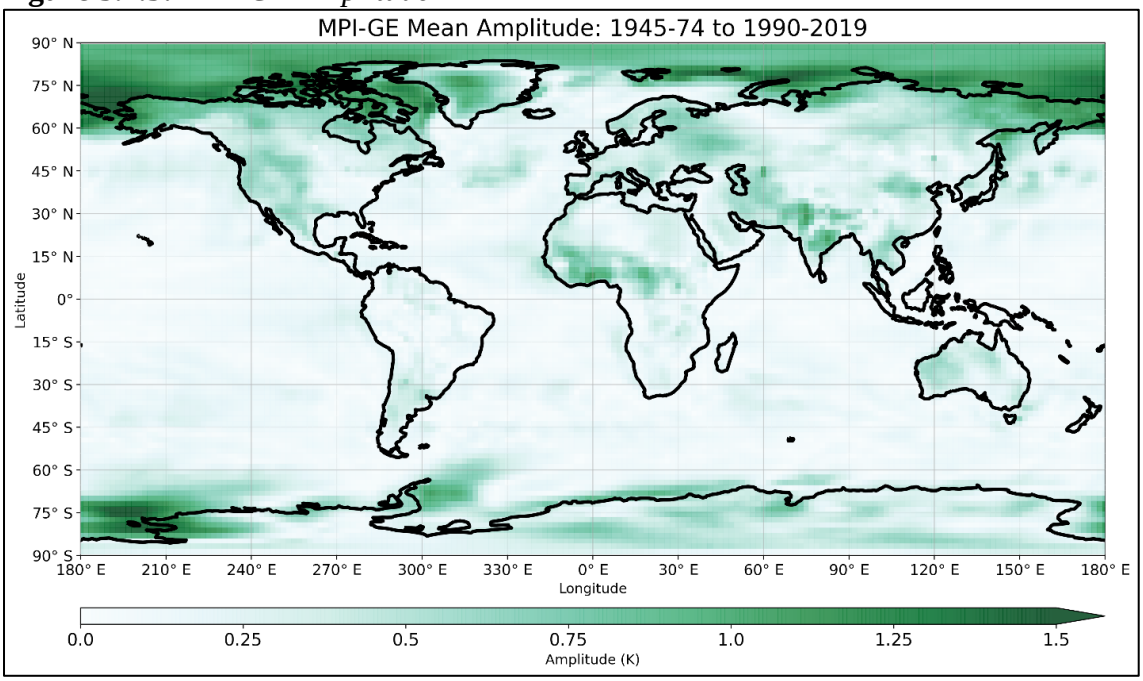
The root mean squared (RMS) distance between the Fourier fit to the seasonal warming cycle of the 100 ensemble members of HadCRUT4.6 and the ensemble mean, using monthly temperature trends from 1945-1974 and 1990-2019.

Figure 5.2.4: *HadCRUT4.6 Amplitude to Spread Ratio*



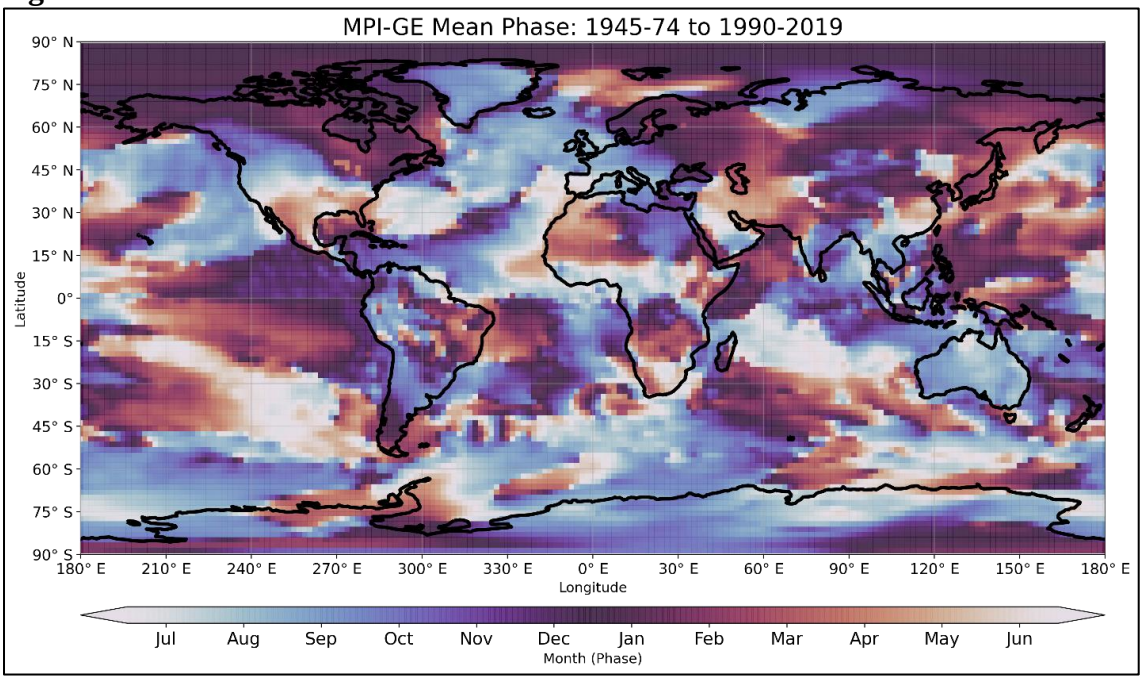
The ratio of HadCRUT4.6 ensemble mean amplitude (fig 5.2.1) to the RMS distance (fig 5.2.3). Monthly temperature trends from 1945-1974 and 1990-2019.

Figure 5.2.5: *MPI-GE Amplitude*



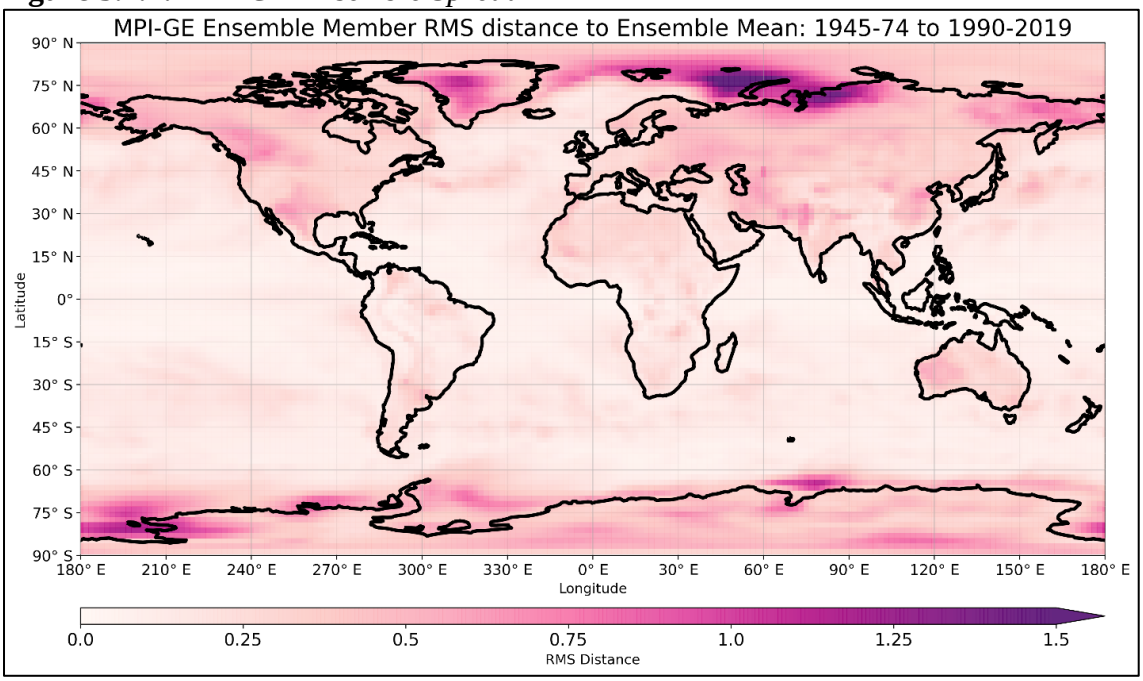
Amplitude of the Fourier fit to the seasonal warming cycle in the MPI-GE model ensemble mean from 1945-1974 to 1990-2019.

Figure 5.2.6: *MPI-GE Phase*



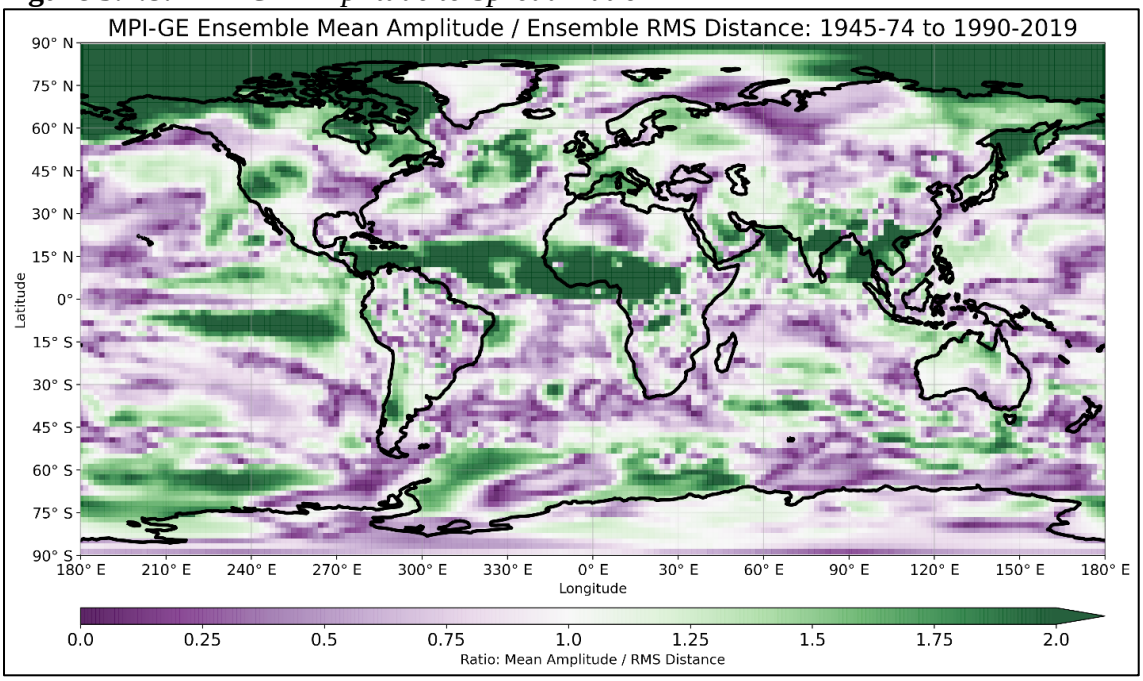
Phase of the Fourier fit to the seasonal warming cycle in the MPI-GE model ensemble mean from 1945-1974 to 1990-2019. The month of the peak of the Fourier fit is indicated by the color.

Figure 5.2.7: *MPI-GE Ensemble Spread*



The root mean squared (RMS) distance between the Fourier fit to the seasonal warming cycle of the 100 ensemble members of MPI-GE and the ensemble mean, using monthly temperature trends from 1945-1974 and 1990-2019.

Figure 5.2.8: *MPI-GE Amplitude to Spread Ratio*



The ratio of MPI-GE ensemble mean amplitude (fig 5.2.5) to the RMS distance (fig 5.2.7). Monthly temperature trends from 1945-1974 and 1990-2019.

6. ZONAL BAND ANALYSIS

6.1. HadCRUT4.6

Figure 6.1.1 displays the zonally averaged phase and amplitude of annual surface temperature trends for using levels of land and ocean masking. Classifications of the data are based on Figure 3.1.1. Grid points with land-based observations, including coasts and islands, are labeled “All-CRUTEM4”, while the subset of land points that do not have any water-based observations are represented by “No-HadSST3”. Similarly, data labeled “All-HadSST3” is the entire spatial coverage HadSST3, which has some overlap with CRUTEM4, whereas “No-CRUTEM4” removes any shared grid points from the CRUTEM4 dataset so that only water-based observations (SST points) remain.

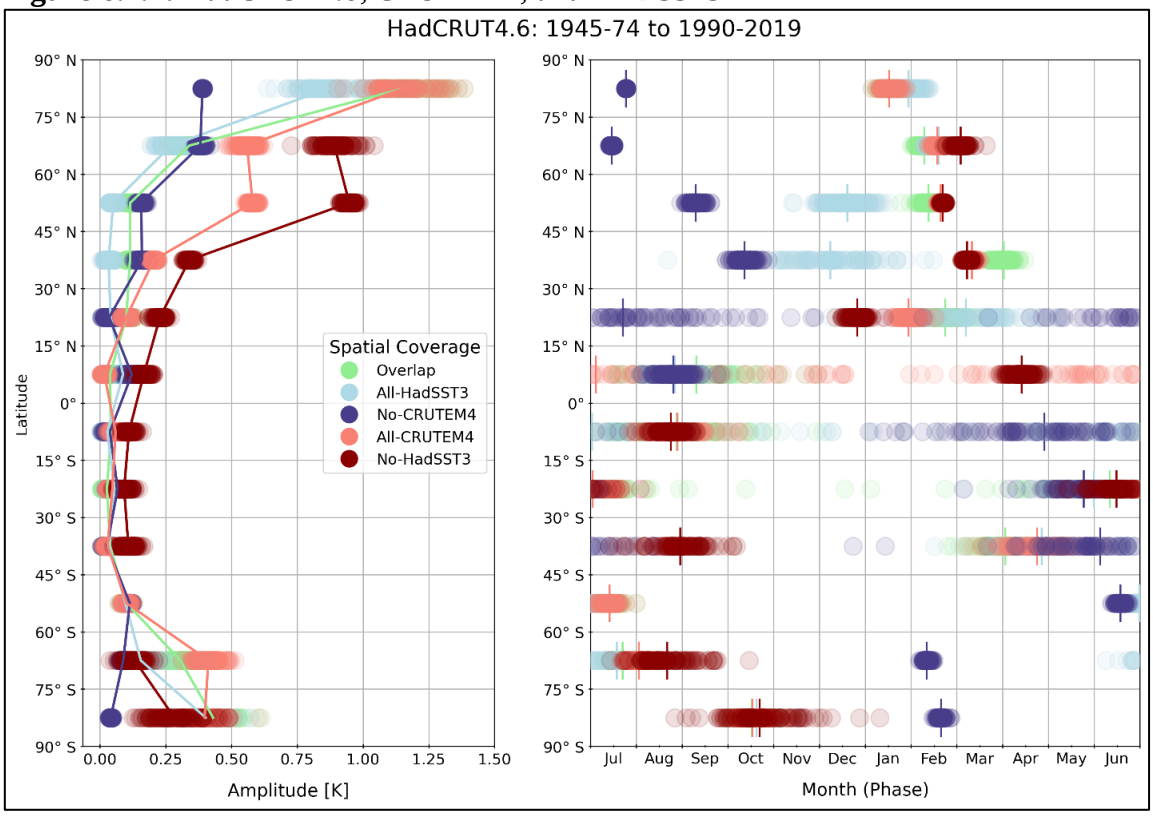
Land generally has a higher amplitude of seasonal warming than sea in most latitude bands, with the highest amplitudes present in Northern Hemisphere middle to high latitudes. Similar to the seasonal warming features identified from figure 5.2.1, the highest amplitudes for land-only points are from 45°N to 75°N (with land-only data not present beyond 75°N) and for ocean-only points are from 60°N to 90°N (Figure 6.1.1). There is a relative maximum amplitude of sea surface temperature seasonal warming in 60°S to 45°S, as noted in the previous section (5.2).

The latitude band 75°N to 90°N contains seasonal cycle variations of the highest amplitude of all the latitude bands and surface types. The averages exhibiting the high amplitudes all include grid points with at least partial land data, whereas “No-CRUTEM4”, which contains no SAT data, has a seasonal warming amplitude only half as large. A similar feature of higher amplitude at the overlap points can be noted on the

high latitude transition from 45°S to 90°S. This indicates that the regions of blended SAT and SST data, typically along coasts, islands, and sea-ice margins at the poles, have a different phase and amplitude of seasonal warming than land or ocean. In general, the inclusion of coastal (overlap) points dampens the amplitude of seasonal warming at mid-latitudes, and exaggerates it at high latitudes compared to SAT and SST alone. When excluding blended SAT and SST points, SAT over land and SST over ocean most clearly exhibit opposite phases of seasonal warming, with reversal across the equator. The certainty in the amplitude of observations is higher in the NH than in the SH, noted by the spread of the ensemble members for either variable.

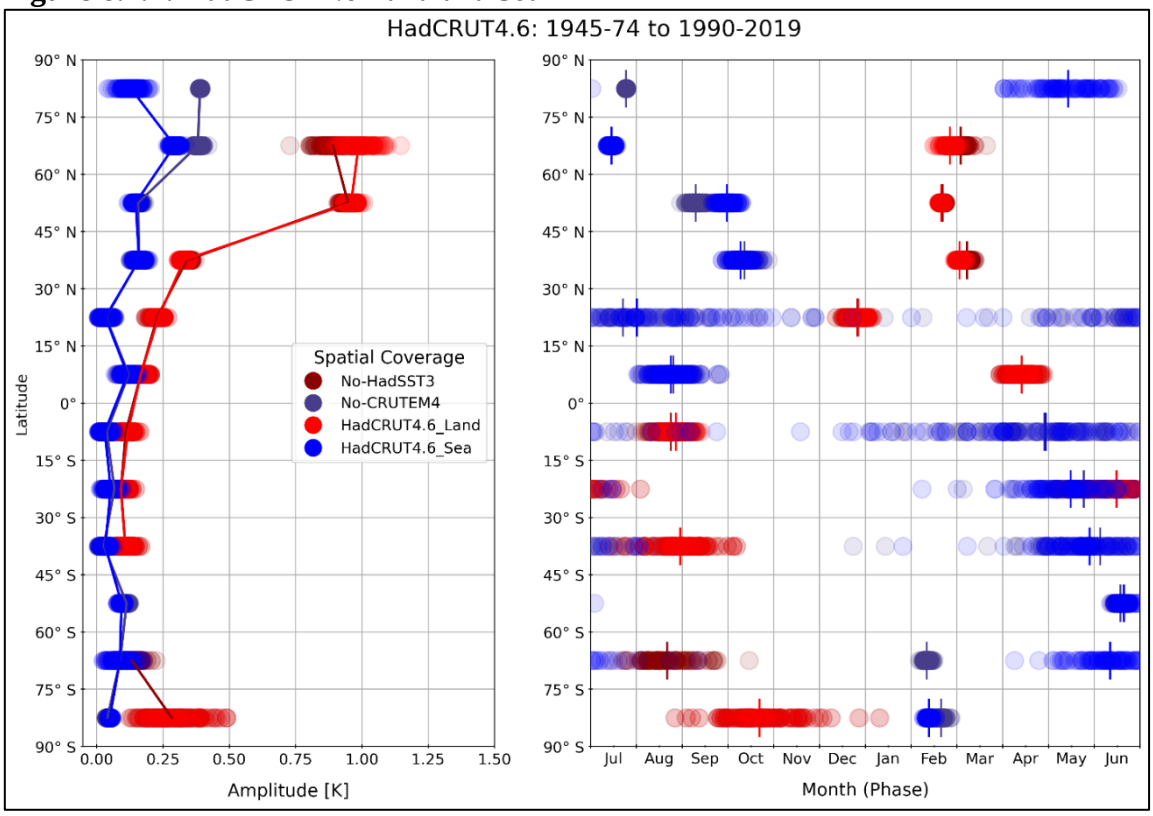
Figure 6.1.2 demonstrates the sensitivity to the choice of land-sea masking methodology. The differences in spatial coverage and the details of their generation are explained in section 3.1. At the latitude band 60°N to 75°N, the use of the GLOBE mask (denoted HadCRUT4.6_Land and HadCRUT4.6_Sea, described in section 3.1) reduces the seasonal warming amplitude for ocean points and increases it for land points. In general, both choices of land-sea masking yield the same phase of seasonal warming to within a month, except for SST at high latitude bands.

Figure 6.1.1: *HadCRUT4.6, CRUTEM4, and HADSST3*



Amplitude (left) and phase (right) of the Fourier fit to the seasonal warming cycle from 1945-1974 to 1990-2019. Vertical lines corresponding to the same-colored dot are the ensemble mean of observation ensemble. Land, including coasts and islands is labeled "All-CRUTEM4", while land points that do not have any overlapping HadSST3 grid points are represented by "No-HadSST3". Similarly, data labeled "All-HadSST3" is the entire spatial coverage HadSST3, which has some overlap with CRUTEM4, whereas "No-CRUTEM4" removes any overlapping points from the CRUTEM4 dataset. Figure 3.1.1 maps the spatial coverage of these HadCRUT4.6 datasets.

Figure 6.1.2: *HadCRUT4.6 Land and Sea*



Amplitude (left) and phase (right) of the Fourier fit to the seasonal warming cycle from 1945-1974 to 1990-2019. Vertical lines corresponding to the same-colored dot are the ensemble mean of observation ensemble. “No-CRUTEM4” denotes regions from the HadCRUT4.6 dataset that only contain SST grid points, and “No-HadSST3” is from regions that only contain SAT grid points. HadCRUT4.6_Land is the HadCRUT4.6 dataset only using grid points of 100% land according to the GLOBE dataset, and HadCRUT4.6_Sea is the HadCRUT4.6 dataset only using grid points of 100% ocean according to the GLOBE dataset. Figure 3.1.2 outlines the differences between these two types of land and sea masks.

6.2. MPI-GE

Figure 6.2.1 compares the use of the standard model variable tas (near surface air temperature) and ts (temperature of the surface) over ocean regions in order to assess if using one variable or the other reveals different results. Other than slight differences at

the highest latitude band, in general, ts and tas behave similarly in each latitude band. Therefore, the variable ts, using 100% sea grid points according the mask outlined in section 3.1, will be used for the remainder of the SST analyses, though the tas model variable could have alternatively been used in place of ts.

Figure 6.2.2 compares the amplitude (left-side) and phase (right-side) of the monthly sea surface temperature (SST) trends for the MPI-GE model ensemble, the MPI-GE model ensemble using the same grid as available observations, and HadCRUT4.6 observation ensemble. Between 60°N and 60°S, the SST amplitude is comparable between models and observations, with the HadCRUT4.6 amplitude tending to be larger. The NH mid-latitude SST seasonal warming peaks in mid boreal autumn according to HadCRUT4.6 ensemble, however the MPI-GE simulates a large spread of potential peak SST warming months, with the ensemble mean in late boreal summer. In the SH between 75°S and 45°S, observed SSTs exhibit peak warming in June, while the month of peak warming according the MPI-GE ensemble mean is in August. In these midlatitude regions for both hemispheres, the HadCRUT4.6 SST ensemble mean seasonal warming signal is statistically different from that of the MPI-GE, even when the MPI-GE is masked to only include grid boxes with available observations.

The MPI-GE overestimates the seasonal warming signal at high latitudes (above 60°N and below 60°S). The high amplitude in the MPI-GE high latitudes is also apparent in figure 5.2.5, as well as figure 6.2.1. In the northern high latitudes, this overestimation of seasonal warming amplitude, as well as the offset of seasonal warming phase compared to the observations, results in the MPI-GE ensemble not capturing the

observed phase and amplitude, made apparent by the p-values for these latitudes. This indicates the MPI-GE could be incorrectly simulating the seasonal warming trend over open ocean at high latitudes; however, it could alternatively be due to the model variable ts (or tas) not accurately capturing the seasonal warming cycle in sea surface temperature due to the presence of sea-ice in some months. Compared to the MPI-GE without masking, the use of the available observation mask reduces the amplitude of the SST seasonal trend, and increases the variability in phase. It doesn't affect the statistical difference between models and observations at northern high latitudes, but does offer some mitigation to the observation-model difference at the southern high latitudes, where observations are sparse and potentially less reliable.

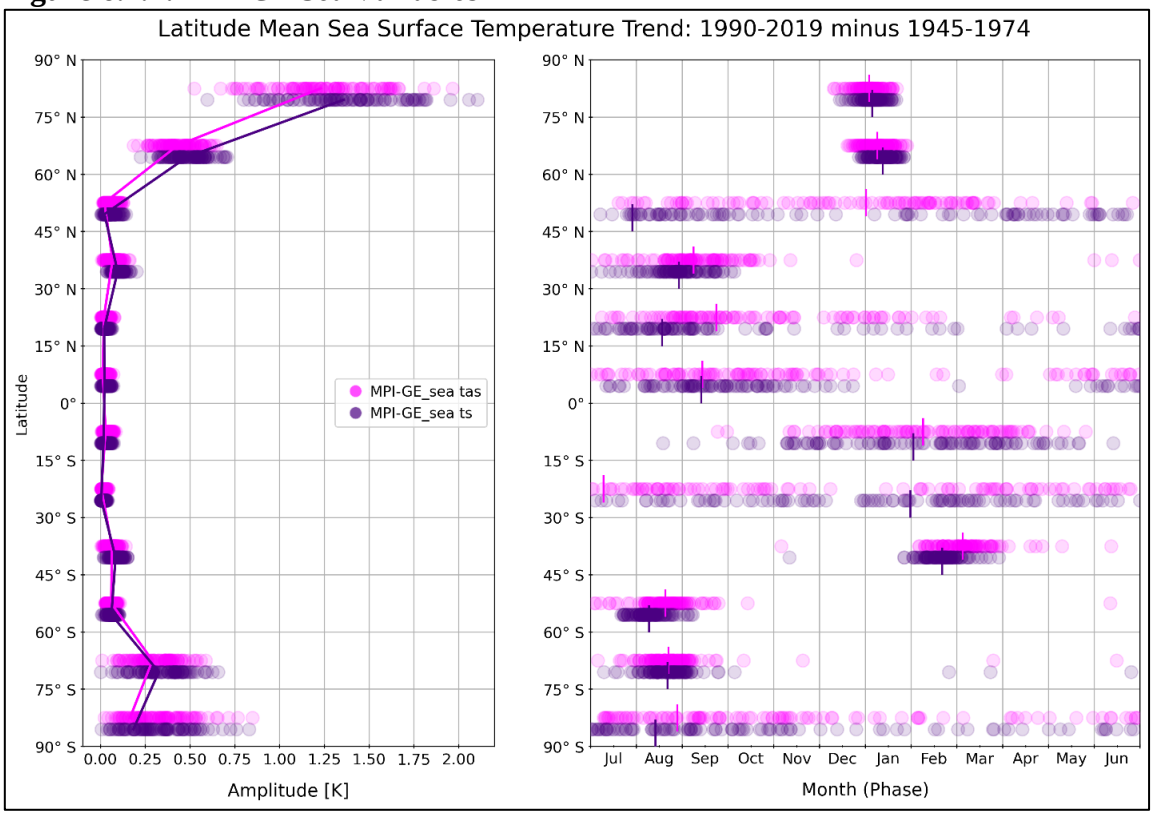
Lastly, tropical latitudes exhibit small seasonal warming amplitudes, which corresponds to incoherent phase values in these regions. This is true for both models and observations. Simulated and observed seasonal warming are more comparable in these latitude bands, and are statistically indistinguishable at 15°S to 0° and 15°N to 30°N. There is a small but distinct observed peak warming that occurs in August from 0° to 15°N but is not present in the MPI-GE ensemble.

The phase and amplitude of seasonal warming trends for the MPI-GE for land-only points of tas (SAT) is compared to the ocean-only points of ts (SST) in figure 6.2.3. Between 60°S and 60°N, surface temperature trends over land have higher seasonal amplitude than over ocean, which is consistent with observations. Poleward of 60°N, MPI-GE_Land tas amplitude is equal to or less than MPI-GE_Sea ts, whereas the HadCRUT4.6 shows the amplitude over land is higher than sea in all latitude bands

where land-only data is present, including the high latitudes above 60°N (compare figure 6.2.3 to figure 6.1.2). In all latitude bands, the month of simulated peak warming over land and over sea occurs within three months of each other, contrary to the tendency for phase over land and sea to be opposite of each other in observations (Fig. 6.1.2).

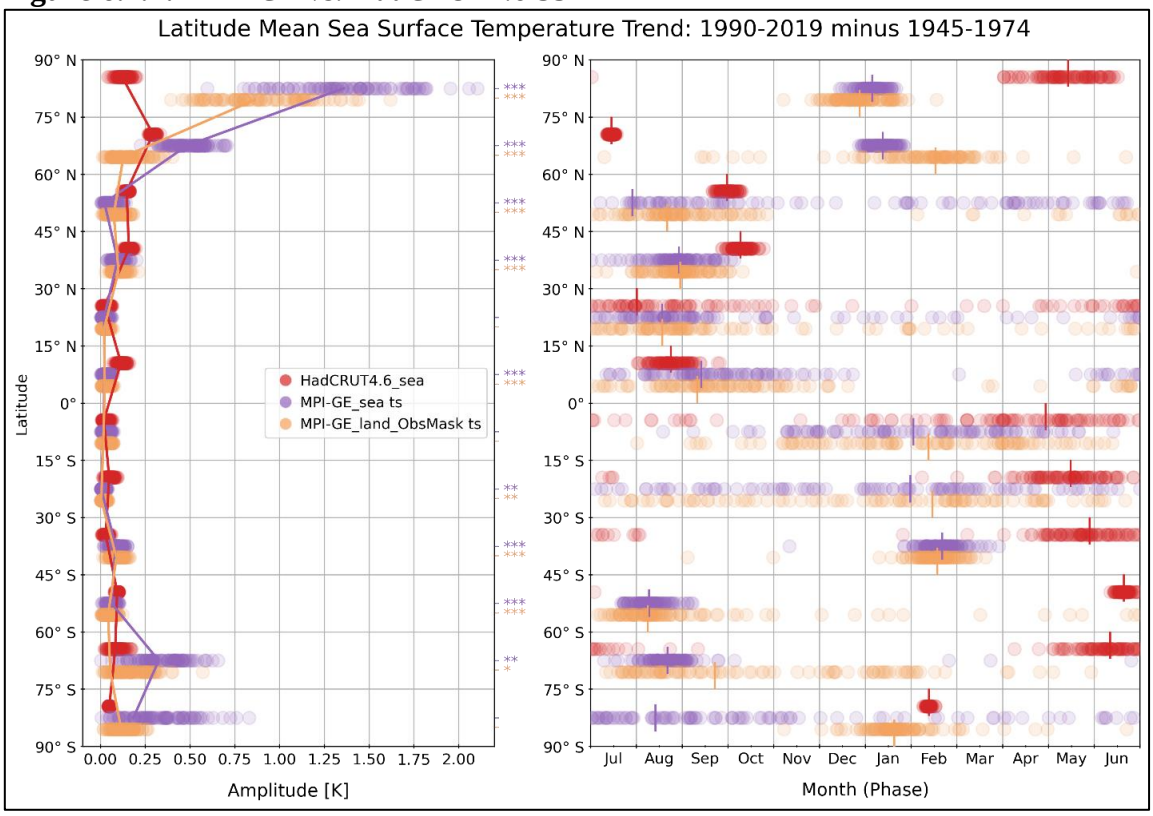
Figure 6.2.4 compares the phase and amplitude of near-surface air temperature (SAT) trends over land for HadCRUT4.6 observation ensemble, the complete MPI-GE ensemble zonally averaged gridded data, and MPI-GE ensemble using grid points only with available observations. In zonal bands from 15°N to 75°N, the MPI-GE underestimates the amplitude of the seasonal warming trend over land compared to observations. In other regions with available observations, modeled and observed amplitudes are comparable. Throughout latitude bands 0° to 45°N, the MPI-GE month of peak warming is 1 to 5 months out of phase with the observations, even when using model grid points that contain available observations. Elsewhere, the MPI-GE has large spread in the phase of peak warming. Though this large spread of phase includes some overlap values with the observed phase of peak warming, the overall observed SAT seasonal warming signal is statistically different from the MPI-GE. The exceptions are in zonal bands that lack observed grid points containing 100% land surface, and in the Antarctic. In the 60°N to 75°N band, the available observation mask brings the simulated seasonal warming amplitude closer to observations, but does not shift the phase of seasonal warming to lessen the statistical difference.

Figure 6.2.1: *MPI-GE Sea Variables*



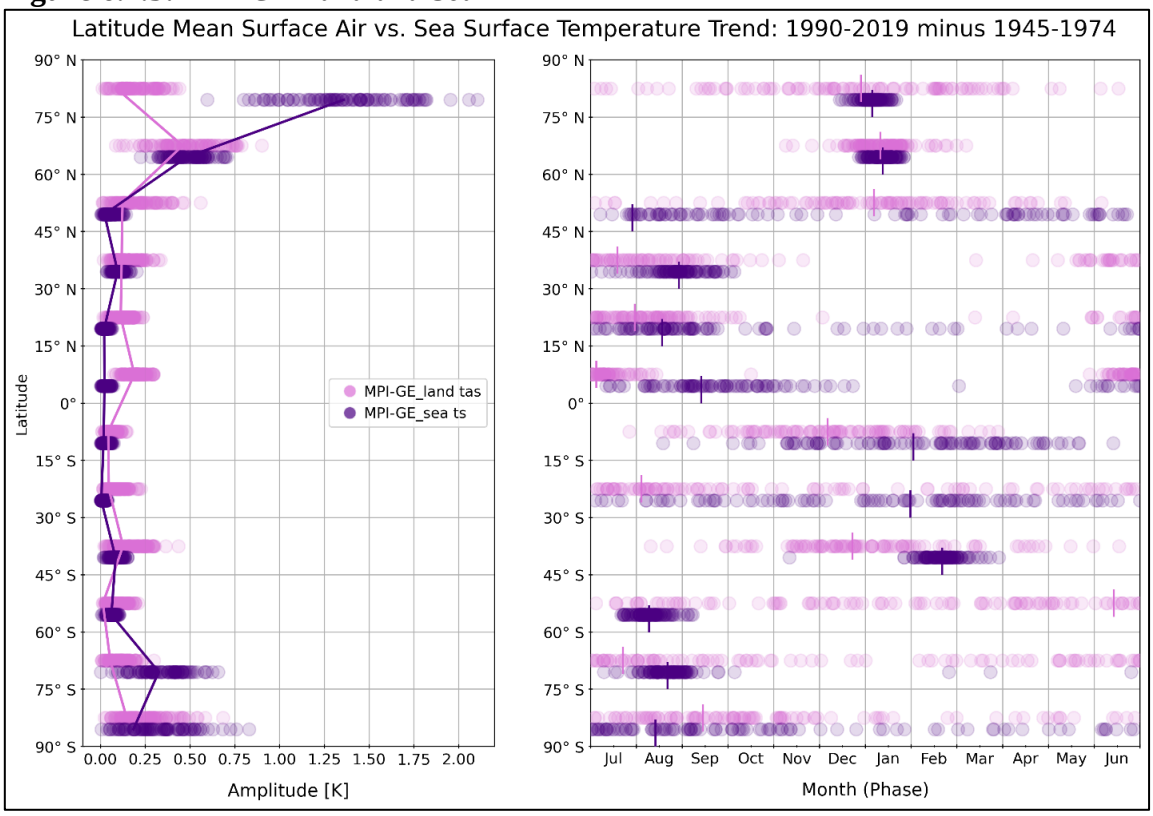
Amplitude (left) and phase (right) of the Fourier fit to the seasonal warming cycle from 1945-1974 to 1990-2019. Vertical lines corresponding to the same-colored dot are the ensemble mean of the observation or model ensemble. MPI-GE tas model variable from sea-only grid points, and MPI-GE ts model variable from sea-only grid points.

Figure 6.2.2: *MPI-GE vs. HadCRUT4.6 SST*



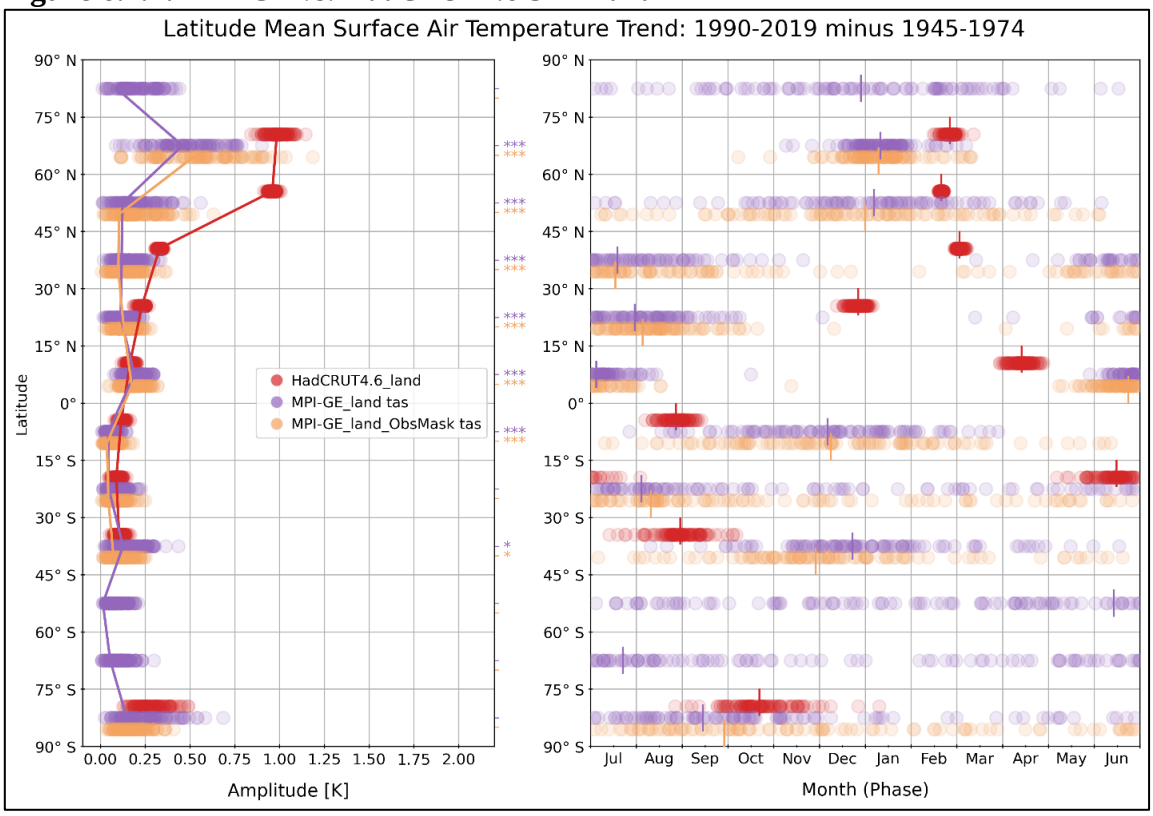
Amplitude (left) and phase (right) of the Fourier fit to the seasonal warming cycle from 1945-1974 to 1990-2019. Vertical lines corresponding to the same-colored dot are the ensemble mean of the observation or model ensemble. MPI-GE ts model variable from sea-only grid points, MPI-GE ts using HadCRUT4.6 sea-only grid points as a mask, and HadCRUT4.6 with sea-only grid points. Asterisks denote whether or not the HadCRUT4.6 ensemble is significantly different from the distribution of model ensemble members at each latitude band.

Figure 6.2.3: *MPI-GE Land and Sea*



Amplitude (left) and phase (right) of the Fourier fit to the seasonal warming cycle from 1945-1974 to 1990-2019. Vertical lines corresponding to the same-colored dot are the ensemble mean of the model ensemble. MPI-GE tas model variable from land-only grid points, and MPI-GE ts model variable with sea-only grid points.

Figure 6.2.4: *MPI-GE vs. HadCRUT4.6 SAT Land*



Amplitude (left) and phase (right) of the Fourier fit to the seasonal warming cycle from 1945-1974 to 1990-2019. Vertical lines corresponding to the same-colored dot are the ensemble mean of the observation or model ensemble. MPI-GE tas model variable from land-only grid points, MPI-GE tas using HadCRUT4.6 land-only grid points as a mask, and HadCRUT4.6 with land-only grid points. Asterisks denote whether or not the HadCRUT4.6 ensemble is significantly different from the distribution of model ensemble members at each latitude band.

6.3. CMIP5 and CMIP6

Figure 6.3.1 is of the amplitude and phase of the near-surface air temperature (SAT) monthly trends over land for all zonal bands, for the observation ensemble and model ensembles. In general, the CMIP5, CMIP6 and MPI-GE amplitude of SAT are consistent with each other across latitude bands. In the Northern Hemisphere,

HadCRUT4.6 has a larger amplitude of seasonal warming between the two time periods than the GCMs. The MPI-GE, CMIP5, and CMIP6 are similarly poor at replicating the amplitude and phase observed in the Northern Hemisphere. This mismatch in seasonal SAT warming signal between models and observations is statistically significant across the Northern Hemisphere, noted by the asterisks in figure 6.3.1. In the Southern hemisphere, the phase and amplitude of SAT seasonal trends in zonal bands have higher observational uncertainty, particularly in the timing of the peak warming month. In most cases throughout the Southern Hemisphere, the observed seasonal SAT trend and associated uncertainty is statistically indistinguishable from the model ensembles.

The amplitude and phase of sea surface temperature (SST) monthly trends over ocean for all zonal bands, for observations and models, is plotted in figure 6.3.2. The amplitudes of seasonal warming modeled by the MPI-GE, CMIP5, and CMIP6 are generally similar to one another across latitude bands, and they all overestimate the observed amplitude at high latitudes in both hemispheres. In the Northern Hemisphere north of 60°N, the simulated phase of SST seasonal warming occurs in mid-winter, however the observed phase is in mid-summer. Southward from 60°N to 30°N, the observed phase of seasonal warming is in mid-fall. For these mid to high latitudes in the Northern Hemisphere, the observed monthly sea surface temperature trend is significantly different than the mean and variability of SST trends simulated by GCMs. The same can be said about significance of observed SST trends versus the simulations in SH midlatitude zonal bands. However, except for the distinct zonal band 0° to 15°N,

most of the tropical temperature trends have little seasonality and the observations are undifferentiated from the GCMs.

Figures 6.3.4 through 6.3.9 are of the same data as figure 6.3.1, but for each 15-degree zonal band, the phase and amplitude of surface air temperature (SAT) seasonal trends are plotted on polar diagrams. The data on figures 6.3.10 through 6.3.15 are the same as are in figure 6.3.2, though similar to the SAT figures, the phase and amplitude of sea surface temperature (SST) seasonal warming trends are plotted on polar diagrams. For all figures 6.3.4 to 6.3.15, the same corresponding 15-degree latitude bands for the Northern Hemisphere and Southern Hemispheres are plotted side by side. These figures and the following discussion demonstrate the relationship between phase and amplitude of the monthly temperature trends and allow for a closer look at covariate distributions of the model ensembles compared to the observation ensemble.

Figure 6.3.5 (left-hand side) shows the phase and amplitude of SAT seasonal temperature trends from 60°N to 75°N over land. HadCRUT4.6 observation ensemble members have a seasonal warming amplitude of 0.8 to 1.2 K, with a peak rate of warming that occurs in late February. The ensemble means from the GCMs simulate peak warming 1.5 to 2.5 months earlier, and the seasonal amplitudes are only half as large as the observations. There are no GCMs that simultaneously simulate the observed phase and amplitude of surface air temperature trends over land in this latitude band. SAT trends over land in the 45° to 60° zonal band in figure 6.3.6 are similar to those in figure 6.3.5 for the Northern Hemisphere; both of which reveal statistical differences between models and observations. As for the Southern Hemisphere, there aren’t enough

land-exclusive grid points in these adjacent latitude bands in order to conclusively compare observations to models (Figure 6.3.5 and 6.3.6 right-hand side). This is also true for the arctic regions plotted in figure 6.3.4, the left-hand side (north) has no data, and the right-hand side (south) has approximately one observation grid point.

The left and right side of figure 6.3.7 (30°N to 45°N (left) and 45°S to 30°S (right)), have a similar phase in their seasonal warming trends, but with weaker amplitude in the Southern Hemisphere. For both hemispheres, the observed maximum rate of SAT warming over land occurs in very early spring, whereas the GCMs have a peak warming mid to late summer. The discrepancy in seasonal temperature trends in the NH latitude band are statistically significant, however they aren't nearly as significant (or at all) in this adjacent SH band.

The tropical zonal bands 15° to 30° (Figure 6.3.8) and 0° to 15° (Figure 6.3.9) in both hemispheres have smaller observed SAT seasonal warming amplitude than the mid- and high- latitude bands. In the Northern Hemisphere tropical zonal bands, the simulated SAT seasonal warming signal does not match the pattern of seasonal warming in the observations, and this mismatch is statistically significant. In the Southern Hemisphere, from the equator to 45°S (Figure 6.3.6 to 6.3.9), the observed seasonal warming signal is generally not statistically distinguishable from simulations. The amplitude of seasonal warming of SAT in these bands are so small there may not be a distinct seasonal warming signal present in the observations or GCMs.

Sea surface temperatures in in the tropical zonal bands 0° to 15° (Figure 6.3.10) and 15° to 30° (Figure 6.3.11) have low amplitude of seasonal warming, which results in

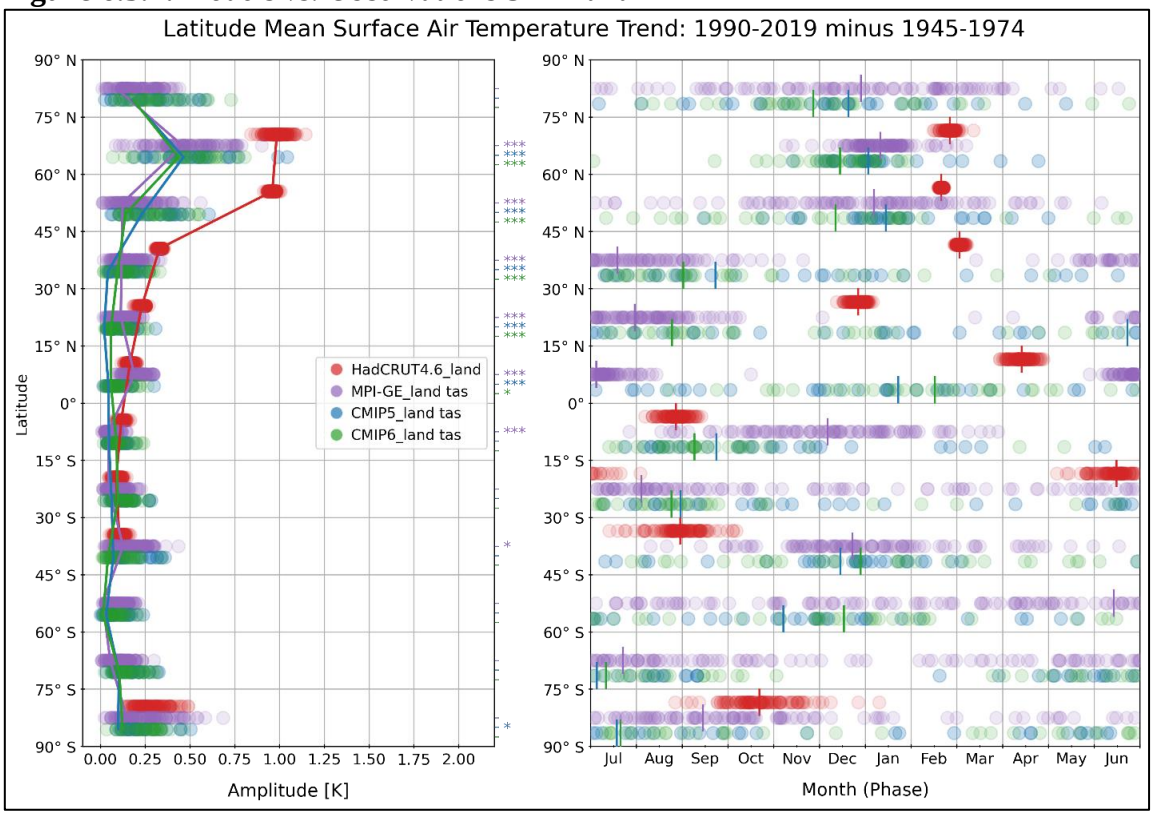
a wide range of phase. Both observations and models exhibit this seasonal warming behavior (or lack thereof) on these zonal bands. The exception is 0° to 15°N, in which the mean observed phase is in mid-August, with an SST amplitude of about 0.1K (left-hand side, figure 6.3.10). None of the GCMs simulate this anomalous warming, and it is statistically significant according to a p-value of less than 0.001 from the Mahalanobis statistic.

The left-hand sides of Figure 6.3.12 (30° to 45°) and 6.3.13 (45° to 60°) exhibit similar results to one another; there is an observed seasonal warming amplitude of around 0.15K in early fall for the Northern mid latitudes. All the models simulate a comparable ensemble mean seasonal amplitude to observations, along with a maximum rate typically one to two months earlier. This mismatch in phase between models and observations is significant in the zonal band 30°N to 45°N for all simulations and in 45°N to 60°N for MPI-GE and CMIP5. On the right-hand side of these figures, the peak warming in SH mid latitude SSTs occurs in June (austral winter). Models, however, distinctly exhibit a maximum rate of warming in February for zonal band 45°S to 30°S and in August for 60°S to 45°S. This suggests the distribution of model natural variability does not represent the observed phase, and this is confirmed by the significant p-value from the Mahalanobis distance at these latitude bands.

Figure 6.3.14 and figure 6.3.15 are of the SST seasonal trends for the latitude band 60° to 75° and 75° to 90°, respectively, for both hemispheres. The GCMs over-estimate the observed amplitude in both hemispheres, and the CMIP ensembles exhibit higher amplitudes than the MPI-GE. The observed Northern Hemisphere peak warming

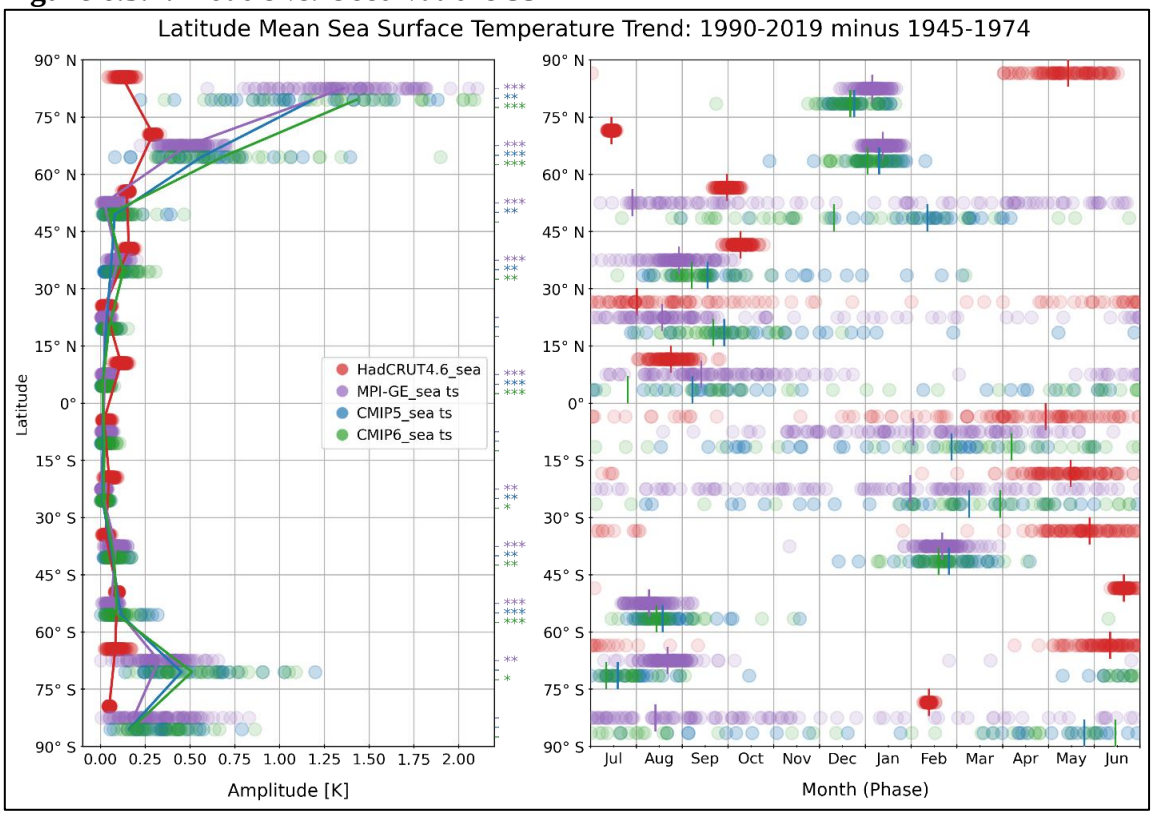
is in the opposite season of that simulated by the GCMs for the high latitudes, and this is statistically significant. The phase simulated by the GCMs in the Southern Hemisphere high latitudes is more variable, therefore the observations are not found to be statistically indistinguishable from simulations across each model ensemble.

Figure 6.3.1: *Models vs. Observations SAT Land*



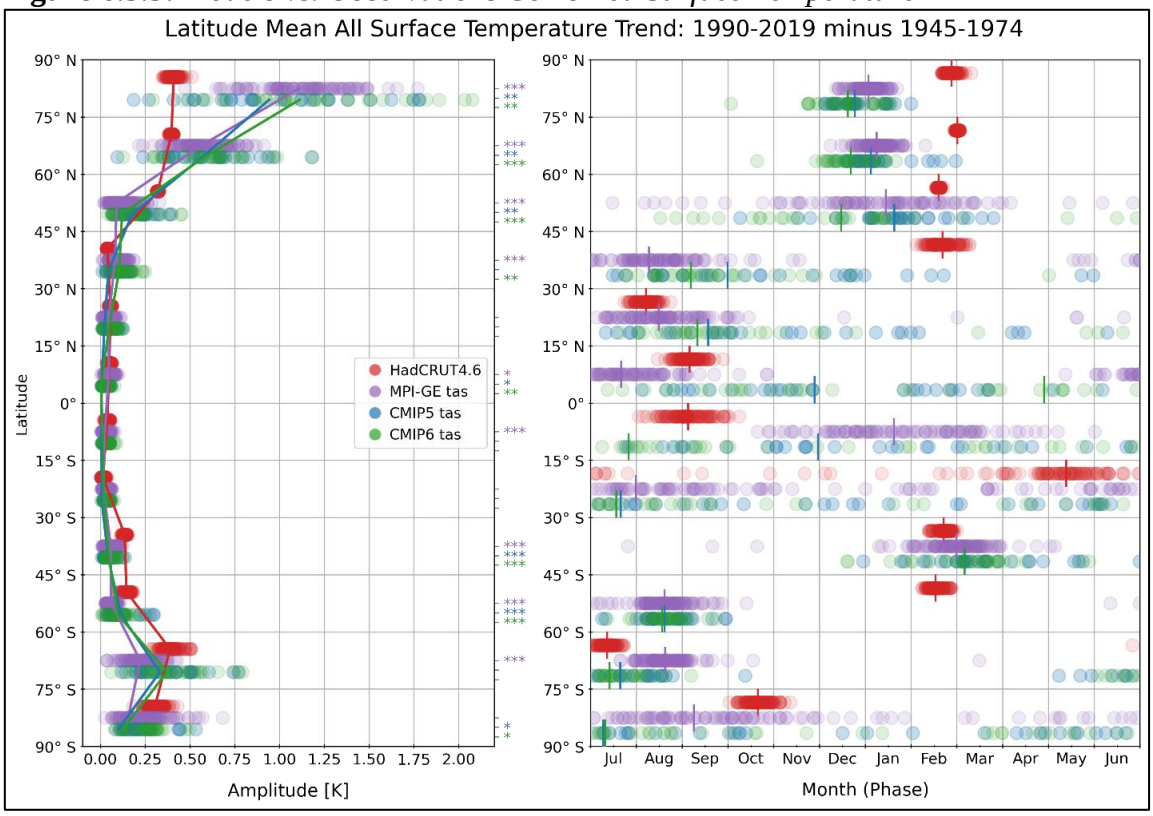
Amplitude (left) and phase (right) of the Fourier fit to the seasonal warming cycle from 1945-1974 to 1990-2019. Vertical lines corresponding to the same-colored dot are the ensemble mean of the observation or model ensemble. Data are of SAT over land-only grid points for each dataset. Asterisks denote whether or not the HadCRUT4.6 ensemble is significantly different from the distribution of model ensemble members for each model ensemble at each latitude band.

Figure 6.3.2: *Models vs. Observations SST*



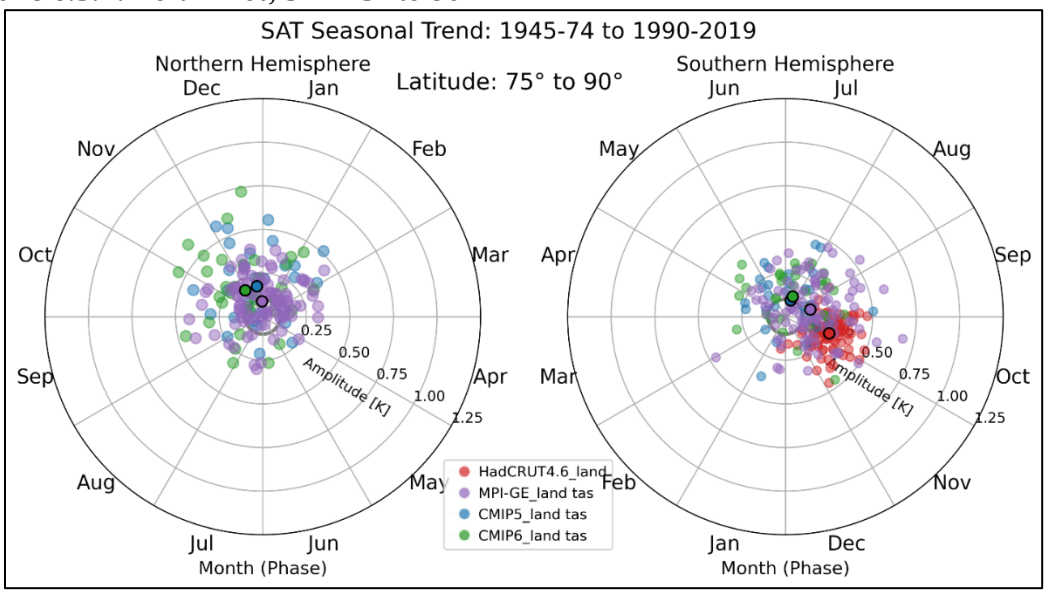
Amplitude (left) and phase (right) of the Fourier fit to the seasonal warming cycle from 1945-1974 to 1990-2019. Vertical lines corresponding to the same-colored dot are the ensemble mean of the observation or model ensemble. Data are of SST over sea-only grid points for each dataset. Asterisks denote whether or not the HadCRUT4.6 ensemble is significantly different from the distribution of model ensemble members for each model ensemble at each latitude band.

Figure 6.3.3: *Models vs. Observations Combined Surface Temperature*



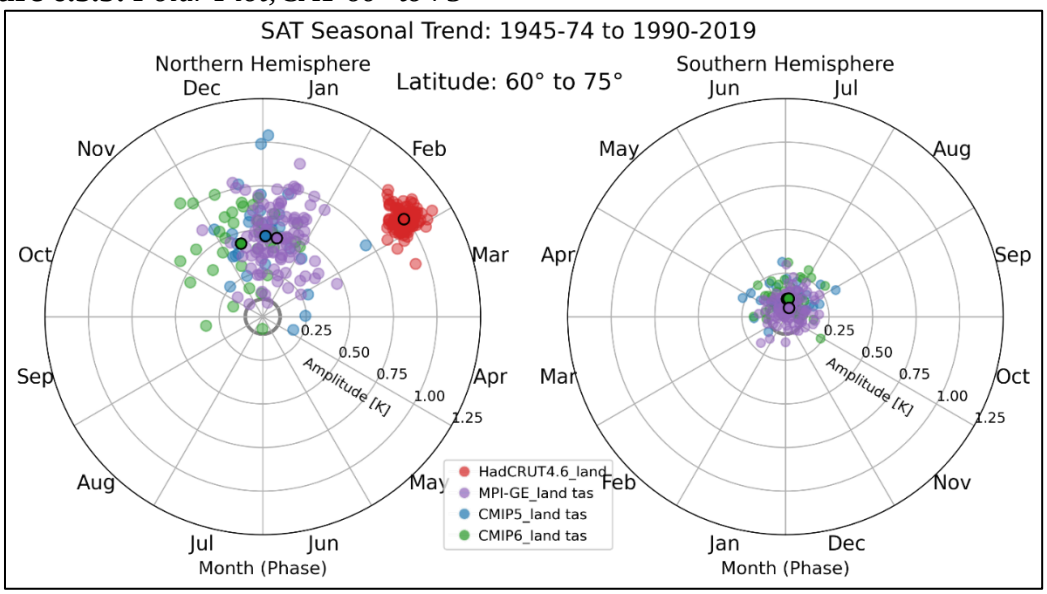
Amplitude (left) and phase (right) of the Fourier fit to the seasonal warming cycle from 1945-1974 to 1990-2019. Vertical lines corresponding to the same-colored dot are the ensemble mean of the observation or model ensemble. Each dataset uses all available grid points without masking. The variable tas is used for models. Asterisks denote whether or not the HadCRUT4.6 ensemble is significantly different from the distribution of model ensemble members for each model ensemble at each latitude band.

Figure 6.3.4: *Polar Plot, SAT 75° to 90°*



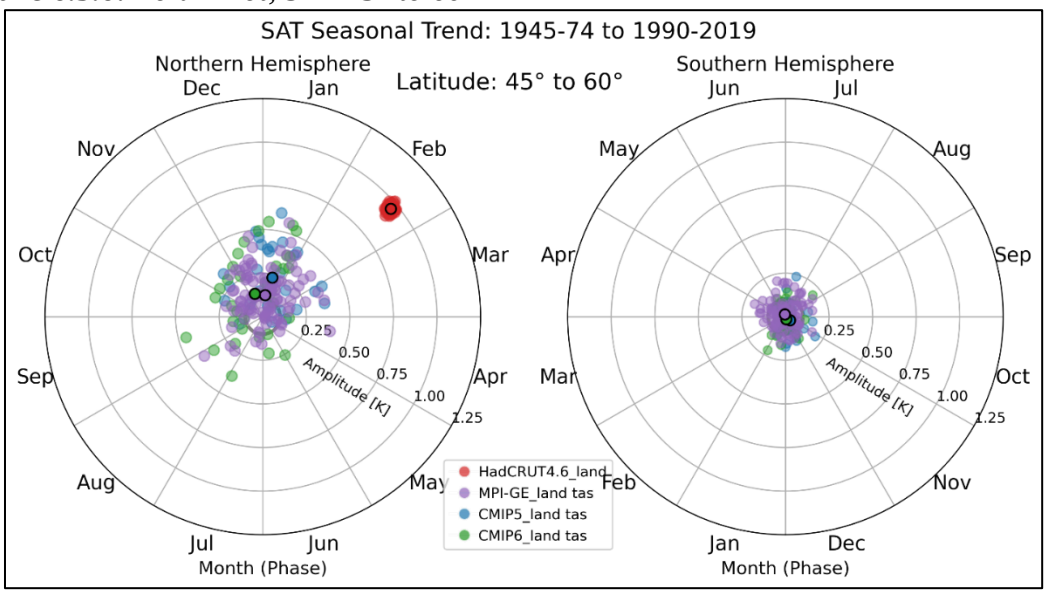
SAT (model tas) over land-only grid points for each dataset. Averaged across grid-points in the zonal band 75° to 90° latitude, displayed for both hemispheres. The amplitude and phase of the Fourier fit to the seasonal warming cycle are represented by r and θ on the polar diagram. Circles with the black outline are of the observation ensemble or model ensemble mean. The Southern Hemisphere polar plot is rotated by 180° to compare boreal and austral seasons. Grey line drawn at 0.1K amplitude for reference of scale.

Figure 6.3.5: *Polar Plot, SAT 60° to 75°*



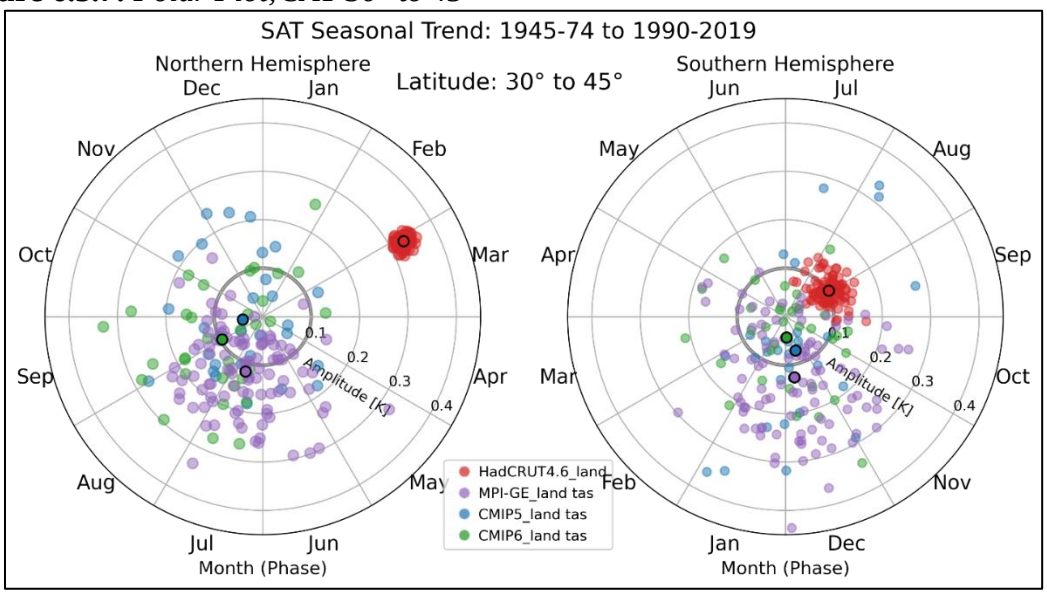
SAT (model tas) over land-only grid points for each dataset. Averaged across grid-points in the zonal band 60° to 75° latitude, displayed for both hemispheres. See Fig 6.3.4's caption for a more detailed figure description.

Figure 6.3.6: *Polar Plot, SAT 45° to 60°*



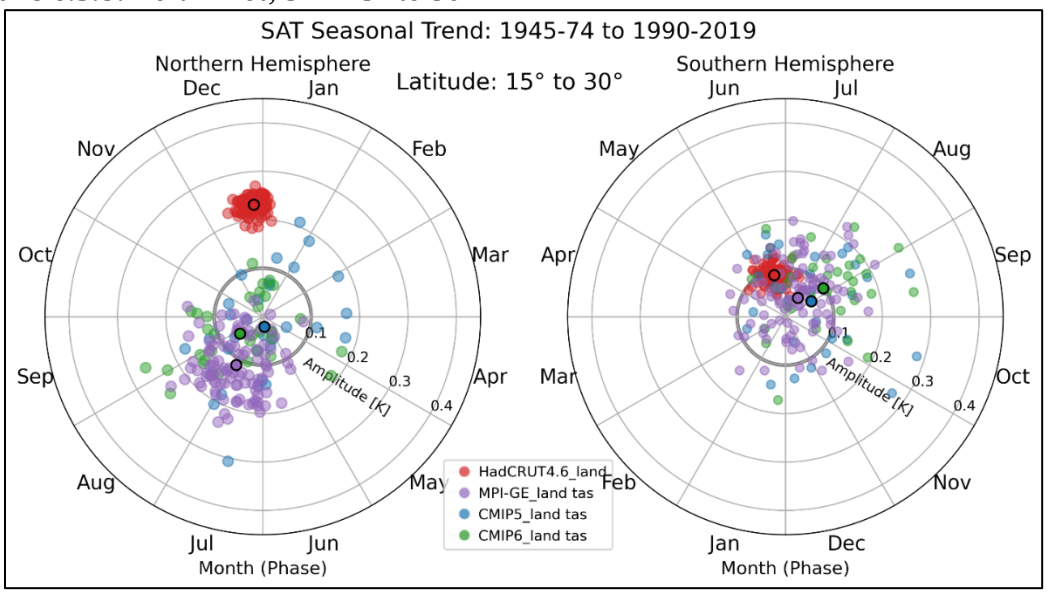
SAT (model tas) over land-only grid points for each dataset. Averaged across grid-points in the zonal band 45° to 60° latitude displayed for both hemispheres. The amplitude and phase of the Fourier fit to the seasonal warming cycle are represented by r and θ on the polar diagram. Circles with the black outline are of the observation ensemble or model ensemble mean. The Southern Hemisphere polar plot is rotated by 180° to compare boreal and austral seasons. Grey line drawn at 0.1K amplitude for reference of scale.

Figure 6.3.7: *Polar Plot, SAT 30° to 45°*



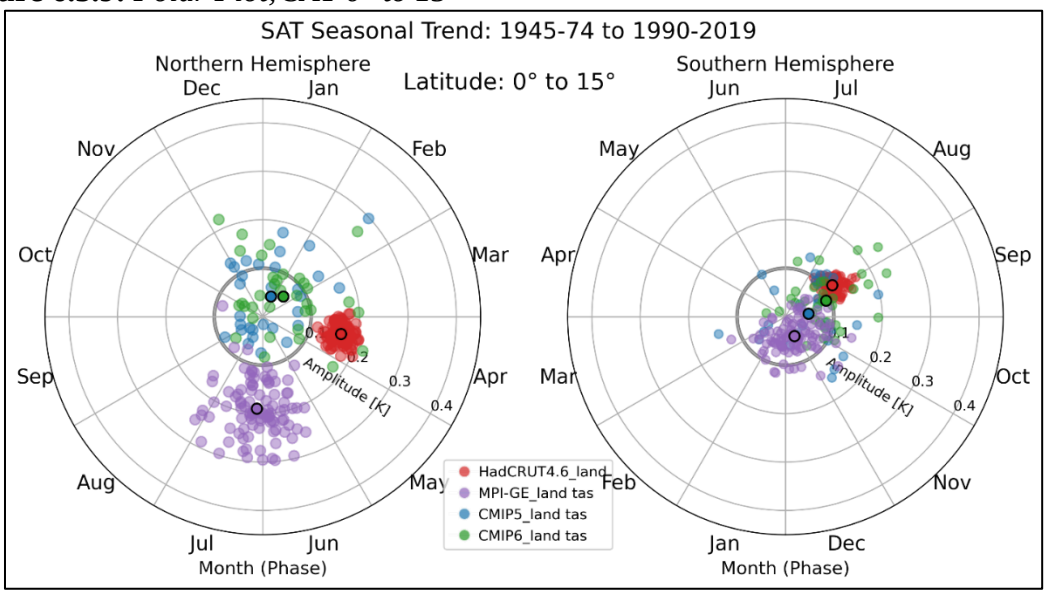
SAT (model tas) over land-only grid points for each dataset. Averaged across grid-points in the zonal band 30° to 45° latitude, displayed for both hemispheres. See Fig 6.3.6's caption for a more detailed figure description.

Figure 6.3.8: *Polar Plot, SAT 15° to 30°*



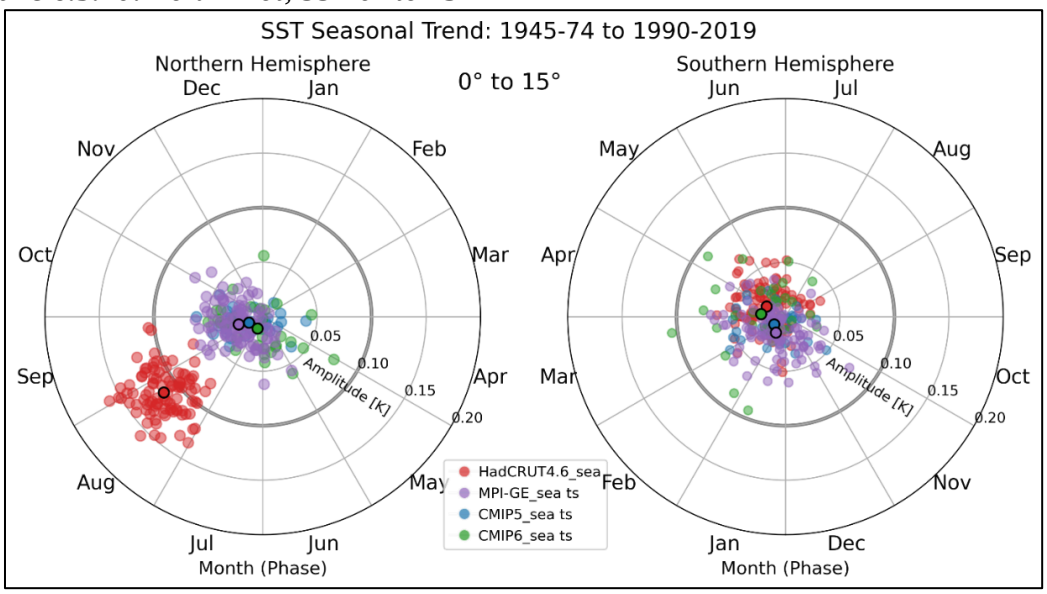
SAT (model tas) over land-only grid points for each dataset. Averaged across grid-points in the zonal band 15° to 30° latitude, displayed for both hemispheres. The amplitude and phase of the Fourier fit to the seasonal warming cycle are represented by r and θ on the polar diagram. Circles with the black outline are of the observation ensemble or model ensemble mean. The Southern Hemisphere polar plot is rotated by 180° to compare boreal and austral seasons. Grey line drawn at 0.1K amplitude for reference of scale.

Figure 6.3.9: *Polar Plot, SAT 0° to 15°*



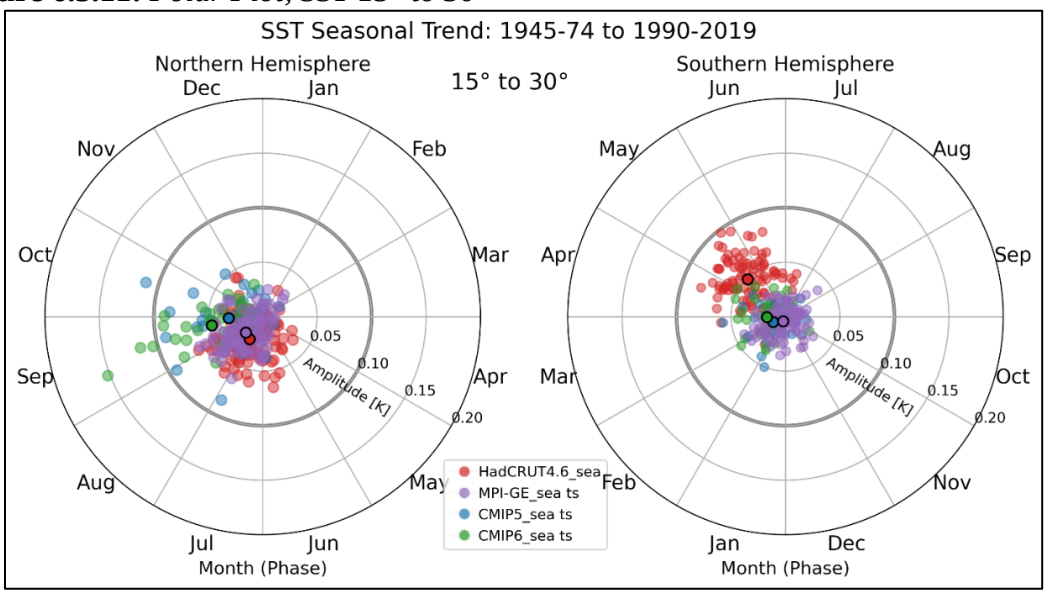
SAT (model tas) over land-only grid points for each dataset. Averaged across grid-points in the zonal band 0° to 15° latitude, displayed for both hemispheres. See Fig 6.3.8's caption for a more detailed figure description.

Figure 6.3.10: *Polar Plot, SST 0°to 15°*



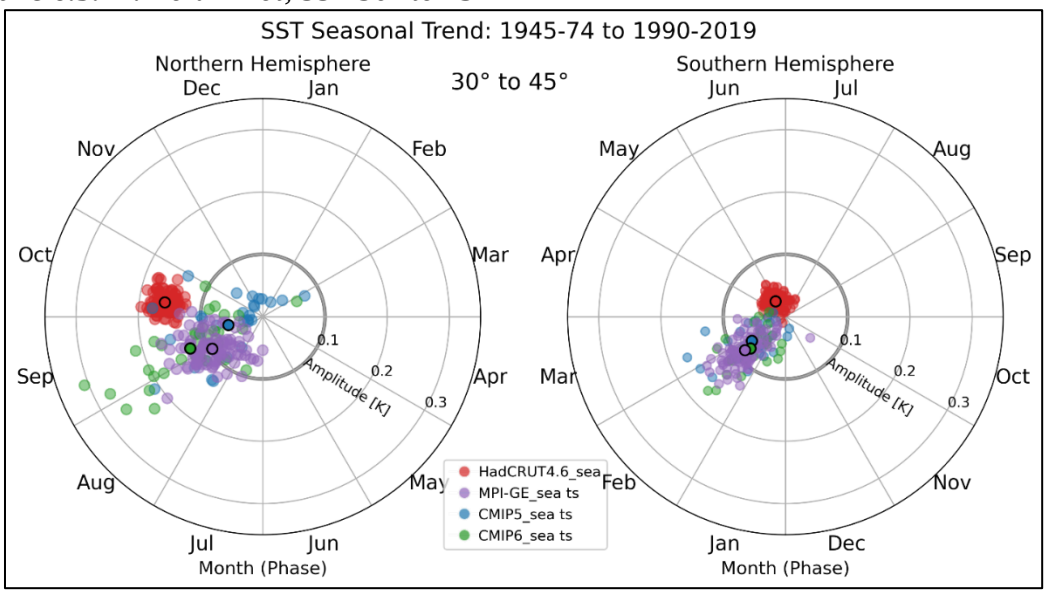
SST (model ts) over sea-only grid points for each dataset. Averaged across grid-points in the zonal band 0° to 15° latitude, displayed for both hemispheres. The amplitude and phase of the Fourier fit to the seasonal warming cycle are represented by r and θ on the polar diagram. Circles with the black outline are of the observation ensemble or model ensemble mean. The Southern Hemisphere polar plot is rotated by 180° to compare boreal and austral seasons. Grey line drawn at 0.1K amplitude for reference of scale.

Figure 6.3.11: *Polar Plot, SST 15°to 30°*



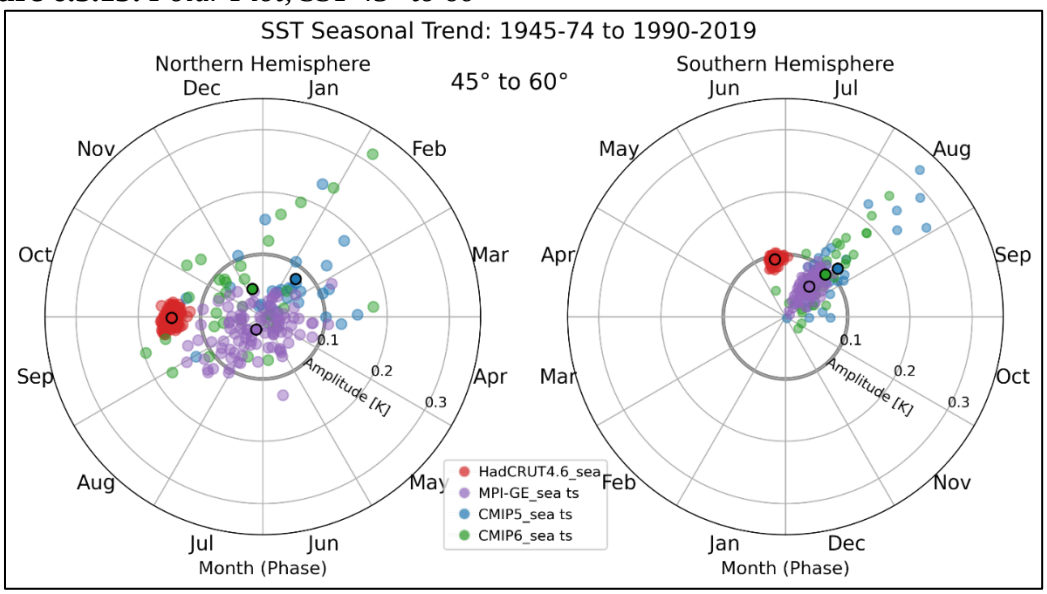
SST (model ts) over sea-only grid points for each dataset. Averaged across grid-points in the zonal band 15° to 30° latitude, displayed for both hemispheres. See Fig 6.3.10's caption for a more detailed figure description.

Figure 6.3.12: *Polar Plot, SST 30°to 45°*



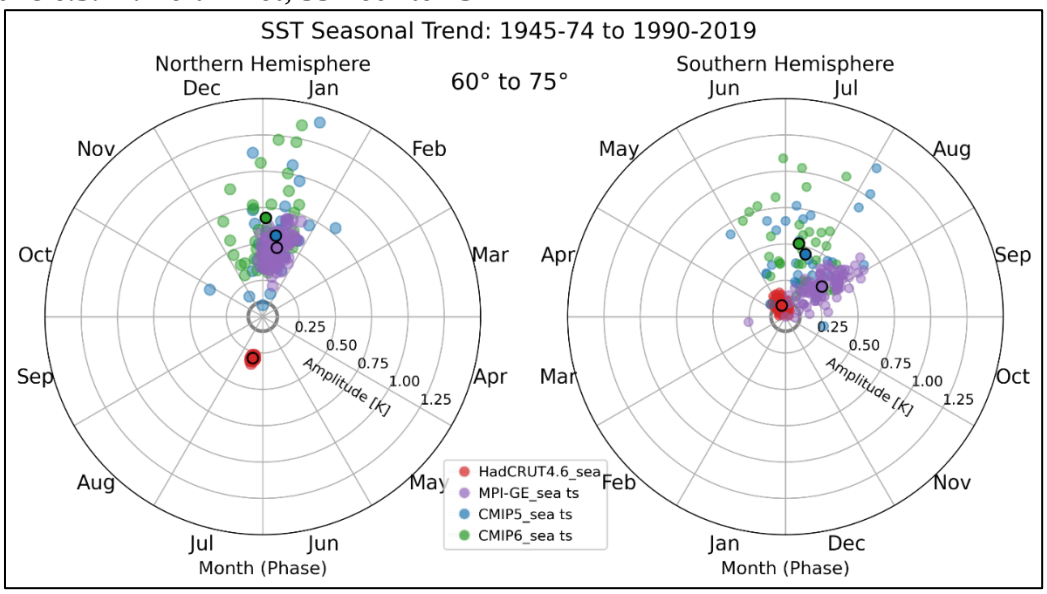
SST (model ts) over sea-only grid points for each dataset. Averaged across grid-points in the zonal band 30° to 45° latitude, displayed for both hemispheres. The amplitude and phase of the Fourier fit to the seasonal warming cycle are represented by r and θ on the polar diagram. Circles with the black outline are of the observation ensemble or model ensemble mean. The Southern Hemisphere polar plot is rotated by 180° to compare boreal and austral seasons. Grey line drawn at 0.1K amplitude for reference of scale.

Figure 6.3.13: *Polar Plot, SST 45°to 60°*



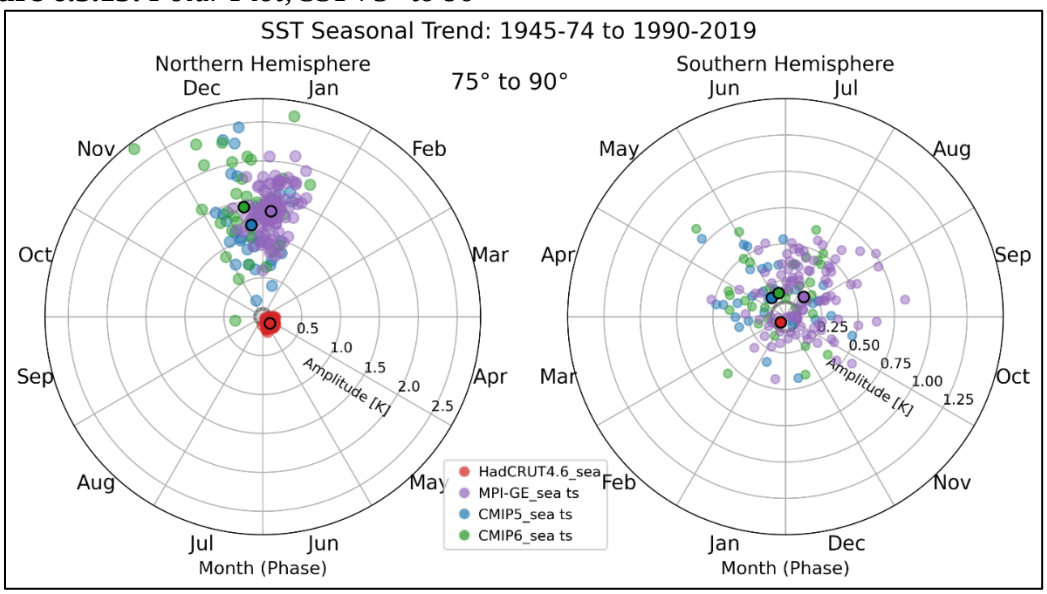
SST (model ts) over sea-only grid points for each dataset. Averaged across grid-points in the zonal band 45° to 60° latitude, displayed for both hemispheres. See Fig 6.3.12's caption for a more detailed figure description.

Figure 6.3.14: *Polar Plot, SST 60°to 75°*



SST (model ts) over sea-only grid points for each dataset. Averaged across grid-points in the zonal band 60° to 75° latitude, displayed for both hemispheres. The amplitude and phase of the Fourier fit to the seasonal warming cycle are represented by r and θ on the polar diagram. Circles with the black outline are of the observation ensemble or model ensemble mean. The Southern Hemisphere polar plot is rotated by 180° to compare boreal and austral seasons. Grey line drawn at 0.1K amplitude for reference of scale.

Figure 6.3.15: *Polar Plot, SST 75°to 90°*



SST (model ts) over sea-only grid points for each dataset. Averaged across grid-points in the zonal band 75° to 90° latitude, displayed for both hemispheres. See Fig 6.3.14's caption for a more detailed figure description.

7. CONCLUSIONS

7.1. Results and Discussion

Observed seasonal warming trends and the notable inconsistencies with GCMs' seasonal warming trends have been identified by previous studies (Mann and Park 1996, Wallace and Osborne 2002, Stine et al. 2009, Dwyer et. al 2012, Qian and Zhang 2015, etc.). The study presented in this paper expands on previous studies by performing a comprehensive analysis of seasonal temperature trends on a global scale. Specifically, the phase and amplitude of the annual cycle of monthly temperature trends are quantified in a global average, on the spatial grids of the datasets, and in zonally averaged latitude bands. The trends are analyzed both collectively and separately for near-surface air temperature (SAT) over land and sea surface temperature (SST) over ocean regions.

If a climate model ensemble is accurately simulating the natural variability and/or the forced response recorded in historical seasonal warming trends, the mean of the observation ensemble seasonal warming (the best estimate of the observed trend given structural uncertainty) should be statistically indistinguishable from members of the climate model ensemble and/or the climate model ensemble mean. Discrepancies may indicate a deficiency in the models' representation of observed natural climate variability and/or forced response due to climate change in the region identified as mismatched. It is important to look to the conceptual model to understand what physical mechanisms are potentially responsible for the observed trends to identify what aspects of the climate models may need improvement.

The observed global average surface temperature trend between 1945-1974 and 1990-2019 results in a seasonal warming maximum relative to the average annual trend of about 0.05K. The mean (forced) response of the global climate model ensembles slightly underestimates this value, but includes a wide range of simulated variability. The observed seasonal maximum warming rate occurs in late February to early March, whereas the simulated maximum occurs between late October to mid-December, depending on the model ensemble. The model to observation discrepancy in the timing and magnitude of maximum seasonal warming is significant to a p-value of <0.01 using multiple observation datasets and <0.001 using the HadCRUT4.6 ensemble.

The maximum observed seasonal warming signal found in the HadCRUT4.6 ensemble mean spans the Northern Hemisphere (NH) mid to high latitude land regions. The amplitude of this feature is around 1 K and occurs between February (mid latitudes) and April (high latitudes). The maximum seasonal warming amplitude simulated by the MPI-GE ensemble in these regions does not extend as far southward into the mid-latitudes, and on average occurs a few months earlier than what is observed. There is a second feature of high temperature trend seasonality observed in the mid-latitude Southern Hemisphere (SH) ocean, collocated with phase of values similar to that of the NH land feature and globally averaged observed trend, ranging from February to May. This region of observed temperature trend seasonality in the SH is not present in the MPI-GE ensemble mean, nor does the model ensemble mean simulate a distinct phase and amplitude of warming in the NH ocean. There may be a small maximum rate of

annual warming over tropical land regions, specifically over Africa, but MPI-GE simulates this maximum warming farther south than observations.

Figure 7.1.a - 7.1.c summarizes the seasonal warming features discussed in the zonal aggregation analysis. Figure 7.1.a is of the SST seasonal trends over ocean, and their spatial contribution to the seasonality in the global mean SST trend, Figure 7.1.b is of the SAT seasonal trends over land and their spatial contribution to the seasonality in the global mean SAT trend, and Figure 7.1.c is the total seasonal surface temperature trend averaged across all data points by each latitude band, and represents how the mean signal across land and ocean present themselves when considering the seasonal warming trend at each latitude band as a whole.

Figure 7.1.a displays the ensemble means' seasonal SST trends by latitude. The SST warming trends that carry the largest weight in the global mean SST signal are in the Southern Hemisphere. From 45°S to 75°S, the observed warming seasonal cycle peaks in July, consistent with a sea ice trend mechanism for altering the seasonal cycle. Model ensemble means have somewhat larger amplitude and tend to peak 1-2 months later, leading to moderate statistical significance differences compared to the observed cycle of surface temperature trends. The overestimation of the amplitude is eliminated by masking the model to only include grid boxes where there is observation data available.

There is a distinct seasonal warming trend in the observations that peaks in May in regions from 45°S to 15°S. The month of peak warming simulated by GCMs leads the observations phase by 2-3 months. The observations are statistically different from all

the GCM's with a p-value of <0.1 to 0.01 for the CMIP ensembles, and <0.01 and <0.001 for MPI-GE ensemble distribution. The peak warming in May in observations, and in March for models doesn't directly suggest any mechanisms laid out in the conceptual model for sea surface temperatures in the mid-latitudes and tropics of the Southern Hemisphere. The mechanism could involve the seasonal cycle of mixing that occurs in the ocean temperature profile, or other oceanic processes not explored in this thesis.

The Northern Hemisphere mid to high latitude SST trend has similar seasonality in latitudinal temperature trends as those identified in the Southern Hemisphere, but opposite in sign (6 months apart). The ocean to land ratio is smaller in the Northern Hemisphere, so even high amplitude trends contribute less to the global mean SST seasonal warming trend. From 30° to 60°N, the annual cycle of temperature trends simulated by GCMs peak in August, which is 1-2 months earlier than the observed phase for this latitude region. In the NH tropics, from 0° to 15°N, the ensemble means of the GCM's have a similar phase of seasonal temperature trends to what is observed in this region, which is similar to NH mid-latitudes, but they all underestimate the magnitude of warming.

From 60° to 90°N, GCMs peak in mid-winter (December to January), which is consistent with the sea-ice albedo feedback's effect on the seasonal cycle of surface temperature. However, observations peak in July on average 60° to 75°N, and in May about 75°N. This is 5-6 months out of phase with GCMs, and it doesn't directly correspond to a forcing mechanism laid out in the conceptual model. This could be a

different driver of change to the seasonal cycle, or a combination of multiple forcing mechanisms that is causing this significant disagreement between models and observations in the NH high latitude SST trends.

Figure 7.1.b is of the observation and model ensemble means' seasonal SAT trends and their contribution to the seasonality of the global mean SAT trend by latitude band. From 30°N to 75°N, the seasonal cycle of observed temperature trends peaks in March, or early boreal spring. This Northern Hemisphere mid to high latitude region has the largest seasonal warming amplitude and largest land area coverage, so this observed trend is the main contributor to the observed global mean SAT seasonal warming signal. This month of peak warming in boreal spring is consistent with the snow feedback mechanism as the driver of change to the seasonal cycle of surface temperature, with the maximum amplitude located between 45°N and 60°N.

The GCM's do not simulate the observed seasonal warming trend in the Northern Hemisphere mid to high latitudes. In regions from 45°N to 75°N, GCMs consistently underestimate the amplitude of observed seasonal warming, and peak a few months earlier, in early to middle boreal winter around December to January. Although this pattern doesn't match observations, it follows a seasonal warming pattern expected by the sea-ice albedo effect outlined in the conceptual model. Since the models SST trends are much larger than their SAT trends in the high latitudes, it's possible that the sea-ice albedo affect is overpowering potential underlying forcing mechanisms originating from land. From 15°N to 45°N, GCM's have a relatively similar phase and amplitude of SAT seasonal warming as they do for SST trends in this region, which have a peak warming

in late boreal summer to early fall. As with SST trends, there is not a single, clear forcing mechanism that was described in this study's conceptual model. Given that SAT and SST are exhibiting similar seasonal warming trends to one another in GCMs, this could be due to changes in the large-scale circulation, such as changes in the Hadley cell extent or the jet stream latitude affecting a region's climate similarly over land and ocean simultaneously, rather than distinct land or ocean process. However, this is not the case for observed SAT and SST warming trends in this latitude region. In the region 15°N to 30°N, the observed seasonal peak warming is in January, which could imply a cooler summer due to increased evapotranspiration; however, a seasonal trend in the annual surface temperature cycle in tropical regions isn't enough to stand alone as evidence of this. North of 30°N, the snow feedback seems to be the primary driver of seasonal warming, which the GCMs all lack or incorrectly simulate.

All of the observed Northern Hemisphere SAT seasonal warming is statistically distinguishable from the GCM ensembles with a p-value of < 0.001. This includes a small region of the tropics, 0° to 15°N. A small but regionally higher area of SAT seasonal warming trend was identified near the Sahel in Africa in the MPI spatial Fourier analysis, and the phase of warming in this zonally averaged latitude band (early boreal summer) is consistent with the grid point analysis in section 5.2 and inconsistent with HadCRUT4 and the CMIP ensembles.

The ratio of land to ocean area is small in the Southern Hemisphere, and is even smaller with respect to observation data availability going back to 1945. The observed seasonal warming trend in the zonal band of 30°S to 45°S peaks in August, or early

austral spring, which is similar to the phase in this zonal band in the Northern Hemisphere. The forcing mechanism driving this early spring seasonal warming trend is unlikely to be the snow feedback, though not impossible, given the limited seasonal snow cover in New Zealand, Chile, and Argentina. The CMIP seasonal warming trends are not significantly different, and the MPI-GE is only marginally so.

Figure 7.1.c is of the ensemble means' seasonal trends by latitude (without any land-sea masking), and weighted by their spatial contribution to the seasonality of the global mean surface temperature trend. In the Northern Hemisphere, from 30°N to 75°N, the seasonality across all data points resembles the seasonality in the SAT land temperature trends (Figure 7.1.b). In the Southern Hemisphere, from 75°S to 30°S, the surface temperature trend's phase and amplitude prominently exhibit SST seasonality (Figure 7.1.a). In other regions, the difference between modeled and observed latitudinal seasonal temperature trends are less significant, contribute less weight to the total global mean, and are a blend of the phases and amplitudes over land and sea.

7.2. Summary and Conclusion

Over the last half century, the observed global seasonal cycle of surface temperature trends peaks in March, with an amplitude of 0.05 K. The MPI-GE, CMIP5, and CMIP6 simulate a global seasonal warming amplitude comparable to observations, but out of phase from observations by a 2 to 5 month lead. The observation-to-model mismatch is significant to a p-value of < 0.01 for the CMIP5 ensemble to all observation datasets, and < 0.001 for the all the GCM ensembles vs. the HadCRUT4.6 observation

ensemble. A similar amplitude, phase, and observation-to-model discrepancy occurs over the course of the entire last century as well, implying a robust trend throughout the observed temperature record and systematic errors in the seasonality of this trend simulated by GCMs.

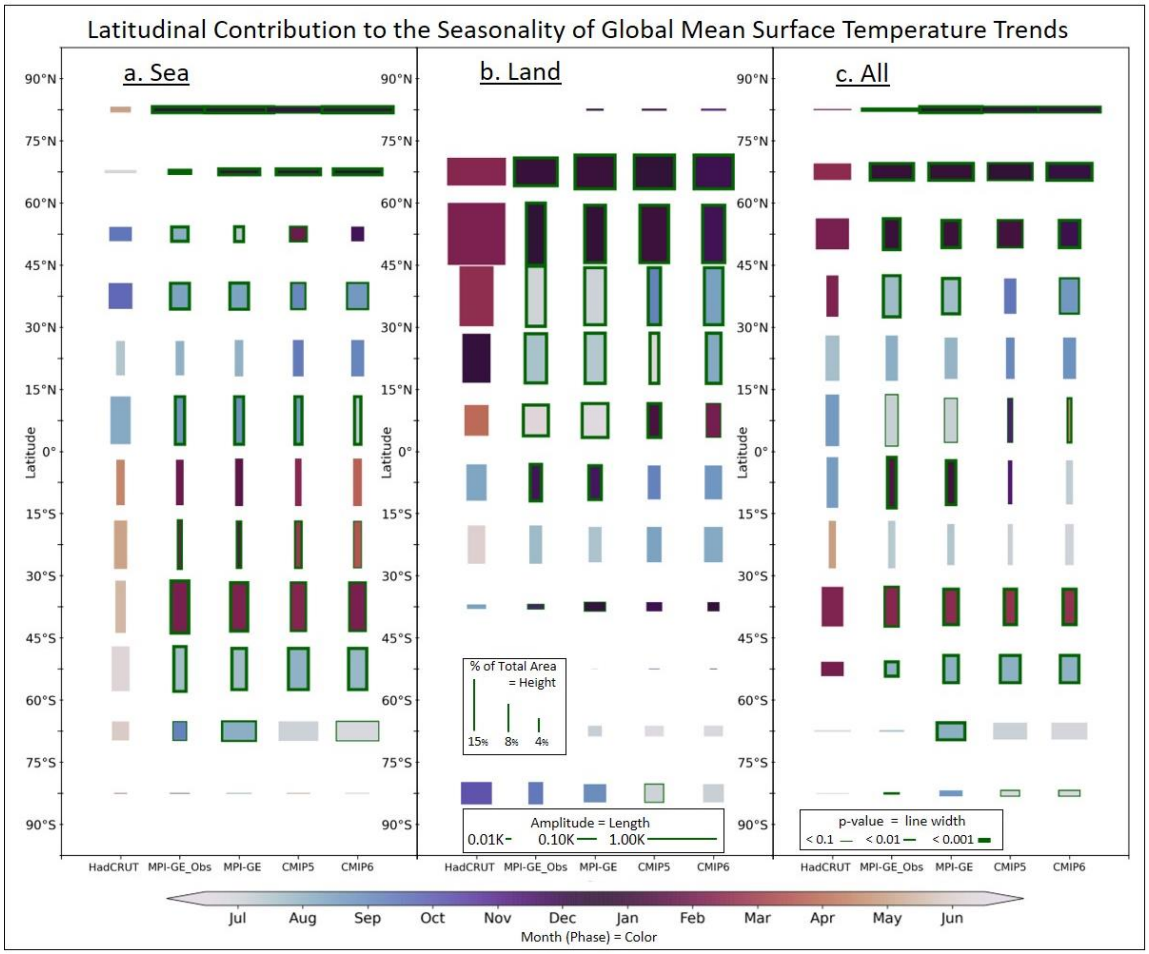
The largest seasonality in surface temperature trends is located from 30°N to 90°N. For SAT trends over land, and SST trends over ocean, all of the GCM ensembles are significantly different from observations to a p-value of < 0.001. Within this broad region, the observed seasonal SAT trends over land regions peak in March, with a maximum seasonal amplitude of about 1.0 K. The models are consistent with one another but do not have a similarly large amplitude extending as far southward from high to mid latitudes as the observations. North of 45°N, the modeled SAT trends lead observations by a few months, peaking in January, and between 45°N and 30°N their phase lags behind observations. The contribution to the global mean seasonal trend is more dispersed across latitudes in GCMs, whereas the largest latitudinal contribution to the observed global mean seasonal trend is from non-tropical NH SATs over land.

The second largest seasonal temperature trends arise from SSTs in 30°S to 60°S. The observed SST trend peaks in late austral fall (Fig 7.1.a), and the combined land-sea mean phase is in March (early austral fall) (Fig 7.1.c). This is the same phase (March) as the peak warming month associated with large amplitudes over NH land, further enforcing the global mean seasonality that peaks around March. The GCMs simulate a statistically different SST trend in this region from the observations with a p-value of < 0.01 in CMIP5 and < 0.001 in the MPI-GE and CMIP6. These mismatched seasonal SST

trends from the GCMs tend to peak a few months later than observations from 60°S to 45°S, and northward to 30°S they peak earlier than observation seasonal SST trends.

A small but distinct annual phase and amplitude of seasonal surface temperature trends have been observed in NH tropical regions from the equator to 30°N. This annual warming trend is mostly absent from models, so the observations are statistically distinguishable from the GCMs' SAT trends to a p-value of < 0.01. Modeled SST trends are also significantly different from observations from the equator to 15°N. Because there aren't many forcing mechanisms that would cause a first-harmonic warming trend in the tropics, it's a good place to observe whether or not precession has affected the seasonal cycle of surface temperature. The concept model describes that precession would induce a maximum seasonal warming between the aphelion (July) and January (perihelion) of Earth's orbit, around October and the same in both Hemispheres. In the most tropical latitude bands, 15°N to 15°S, the combined land-sea temperature trends' time of peak warming in the HadCRUT4.6 is around September. More detailed analyses would be needed to verify this speculation, as well as regional and semi-annual cycle analyses of temperature and precipitation in these latitude regions.

Figure 7.1.1(a-c): *Summary of Latitudinal SAT and SST Trends*



Rectangles are latitudinal ensemble mean phase (color) and amplitude (horizontal length) of the seasonal warming trends between 1945-1974 and 1990-2019 for HadCRUT4.6, MPI-GE using available observations, the full MPI-GE, CMIP5, and CMIP6. The significant p-values from the Mahalanobis distances calculated in sections 6.2 and 6.3 are noted by the thickness of the green border. For land and sea, (a) and (b) the vertical height is scaled by ratio of the area considered 100% land and 100% sea of each latitude band to the area of 100% land and 100% sea across the entire globe. For all points (c), the vertical height is scaled by the total area in each latitude band with respect to the entire area of the globe. The observations account for the missing data points.

REFERENCES

Cowtan, K., and R. Way, 2014: Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Quarterly Journal of the Royal Meteorological Society*, **140,** 1935-1944. https://doi.org/10.1002/qj.2297.

Donohoe, A., and D. Battisti, 2013: The Seasonal Cycle of Atmospheric Heating and Temperature. *J. Climate,* **26,** 4962-4980.

Dwyer, J., M. Biasutti, and A. Sobel, 2012: Projected changes in the seasonal cycle of surface temperature. *J. Climate*, **25,** 6359–6374.

Eyring, V., S. Bony, G. Meehl, C. Senior, B. Stevens, et al., 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development, European Geosciences Union,* **9,** 1937-1958. doi: 10.5194/gmd-9-1937-2016.

Freeman, E., S. Woodruff, S. Worley, S. Lubker, E. Kent, W. Angel, D. Berry, P. Brohan, R.Eastman, L. Gates, W. Gloeden, Z. Ji, J. Lawrimore, N. Rayner, G. Rosenhagen, and S. Smith, 2017: ICOADS Release 3.0: A major update to the historical marine climate record. *Int. J. Climatol. CLIMAR-IV Special Issue,* **37,** 2211-2237, doi:10.1002/joc.4775.

GISTEMP Team, 2020: GISS Surface Temperature Analysis (GISTEMP), version 4. NASA Goddard Institute for Space Studies. Dataset accessed 2020-09 at https://data.giss.nasa.gov/gistemp/.

GLOBE Task Team and others (Hastings, David A., Paula K. Dunbar, Gerald M. Elphingstone, Mark Bootz, Hiroshi Murakami, Hiroshi Maruyama, Hiroshi Masaharu, Peter Holland, John Payne, Nevin A. Bryant, Thomas L. Logan, J.-P. Muller, Gunter Schreier, and John S. MacDonald), eds., 1999. The Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, Version 1.0. National Oceanic and Atmospheric Administration, National Geophysical Data Center, 325 Broadway, Boulder, Colorado 80303, U.S.A. Digital data base on the World Wide Web (URL: http://www.ngdc.noaa.gov/mgg/topo/globe.html)

Huang, B., P. Thorne, V. Banzon, T. Boyer, G. Chepurin, J. Lawrimore, M. Menne, T. Smith, R. Vose, and H. Zhang, 2017: NOAA Extended Reconstructed Sea Surface Temperature (ERSST), Version 5. NOAA National Centers for Environmental Information. 2020-09. doi:10.7289/V5T72FNM.

Jones, P. D., D. H. Lister, T. J. Osborn, C. Harpham, M. Salmon, and C. P. Morice, 2012: Hemispheric and large-scale land surface air temperature variations: An extensive revision and an update to 2010. *J. Geophys. Res.*, **117**, D05127, doi:10.1029/2011JD017139

Karin, Todd. Global Land Mask. October 5, 2020. https://doi.org/10.5281/zenodo.4066722

Kennedy, J., N. Rayner, R. Smith, M. Saunby, and D. Parker, 2011b: Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 1: measurement and sampling errors. *J. Geophys. Res.*, **116,** D14103, doi:10.1029/2010JD015218.

Kennedy, J., N. Rayner, R. Smith, M. Saunby, and D. Parker, 2011c: Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 2: biases and homogenisation. *J. Geophys. Res.*, **116,** D14104, doi:10.1029/2010JD015220.

Lehner, F., C. Deser, N. Maher, J. Marotzke, E. M. Fischer, L. Brunner, R. Knutti, E. Hawkins, 2020: Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. *Earth Syst. Dynam.*, **11**, 491–508. https://doi.org/10.5194/esd-11-491-2020.

Lenssen, N., G. Schmidt, J. Hansen, M. Menne, A. Persin, R. Ruedy, and D. Zyss, 2019: Improvements in the GISTEMP uncertainty model. *J. Geophys. Res. Atmos*., **124,** 6307-6326, doi:10.1029/2018JD029522.

Maher, N., S. Milinski, L. Suarez-Gutierrez, M. Botzet, M. Dobrynin, L. Kornblueh, et al., 2019: The Max Planck Institute Grand Ensemble: Enabling the exploration of climate system variability. *Journal of Advances in Modeling Earth Systems*, **11,** 2050–2069. https://doi.org/10.1029/2019MS001639.

Mann, M., and J. Park, 1996: Greenhouse warming and changes in the seasonal cycle of temperature: Model versus observations. *Geophys. Res. Lett.*, **23,** 1111–1114. https://doi.org/10.1029/96GL01066.

Menne, M., C. Williams, B. Gleason, J. Rennie, and J. Lawrimore, 2018: The Global Historical Climatology Network Monthly Temperature Dataset, Version 4. *J. Climate*, in press. https://doi.org/10.1175/JCLI-D-18-0094.1.

Morice, C., J. Kennedy, N. Rayner, and P. Jones, 2012: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 dataset. *J. Geophys. Res.*, **117,** D08101, doi:10.1029/2011JD017187.

Nigam, S., N. Thomas, A. Ruiz-Barradas, S. Weaver, 2017: Striking Seasonality in the Secular Warming of the Northern Continents: Structure and Mechanisms. *J. Climate,* **30,** 6521–6541. https://doi.org/10.1175/JCLI-D-16-0757.1.

Qian, C., and X. Zhang, 2015: Human Influences on Changes in the Temperature Seasonality in Mid- to High-Latitude Land Areas. *J. Climate*, **28,** 5908- 5921.

Rayner, N., D. Parker, B. Horton, C. Folland, L. Alexander, D. Rowell, E. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.,* **108,** No. D14, 4407 10.1029/2002JD002670.

Sarma, K.V.S, and R.V. Vardharn, 2019: Comparison of Multivariate Means, *Multivariate statistics made simple: a practical approach*. Boca Raton, FL : CRC Press, Taylor & Francis Group, 33-43.

Stine, A., P. Huybers, and I. Fung, 2009: Changes in the phase of the annual cycle of surface temperature. *Nature*, **457,** 435–440, doi:10.1038/nature07675.

Stine, A., and P. Huybers, 2012: Changes in the seasonal cycle of temperature and atmospheric circulation. *J. Climate,* **25,** 7362–7380. https://doi.org/10.1175/JCLI-D-11-00470.1.

Taylor, K., R. Stouffer, and G. Meehl, 2012: An Overview of CMIP5 and the Experiment Design. *Bull. Amer. Meteor. Soc.* **93,** 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Thomson, D., 1995: The seasons, global temperature, and precession. *Science*, **268,** 59–68, doi:10.1126/science.268. 5207.59.

Trouet, V. and G. Oldenborgh, 2013: KNMI Climate Explorer: A Web-Based Research Tool for High-Resolution Paleoclimatology. *Tree-Ring Research*, **69,** 3-13. https://doi.org/10.3959/1536-1098-69.1.3

Wallace, C., and T. Osborn, 2002: Recent and future modulation of the annual cycle. *Climate Res.*, **22,** 1–11. doi:10.3354/cr022001.

Wang, H., S. Schubert, M. Suarez, J. Chen, M. Hoerling, A. Kumar, and P. Pegion, 2009: Attribution of the Seasonality and Regionality in Climate Trends over the United States during 1950–2000. *J. Climate*, **22,** 2571–2590. https://doi.org/10.1175/2008JCLI2359.1.

Zhang, H., B. Huang, J. Lawrimore, M. Menne, and T. Smith, NOAA Global Surface Temperature Dataset (NOAAGlobalTemp), Version 5. NOAA National Centers for Environmental Information. doi:10.25921/9qth-2p70 2020-09.