On the structure of atmospheric warming in models and observations: Implications for the lapse rate feedback

Stephen D. Po-Chedley

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2016

Reading Committee:

Qiang Fu, Chair

Kyle Armour

Cecilia Bitz

Program Authorized to Offer Degree:
Department of Atmospheric Sciences

University of Washington

Abstract

On the structure of atmospheric warming in models and observations: Implications for the lapse rate feedback

Stephen D. Po-Chedley

Chair of the Supervisory Committee:
Professor Qiang Fu
Department of Atmospheric Sciences

This dissertation investigates the structure of atmospheric warming in observations and general circulation models (GCMs). Theory and GCMs suggest that warming is amplified in the tropical upper troposphere relative to the lower troposphere and the surface – a phenomenon known as vertical amplification. We assess model and observational agreement using several amplification metrics derived from the satellite-borne microwave sounding unit (MSU) atmospheric temperature trends.

An important correction to the satellite microwave record is the removal of temperature drifts caused by changes in diurnal sampling. This correction the principal source of uncertainty in microwave temperature datasets. Furthermore, in three existing datasets, the ratio of tropical warming between the upper troposphere (T24 channel) and the surface ($dT24/dT_s$ $\sim$0.6 – 1.3) is lower than that of GCMs ($\sim$1.4 – 1.6). To better understand these issues, we produced an alternate MSU dataset with an improved diurnal correction. We show that existing MSU datasets likely underestimate tropical mid-tropospheric temperature trends. Subsequent improvements to MSU datasets using similar diurnal correction techniques leads to amplification ratios (between T24 and the surface) that are in accord with models.

Another measure of tropical tropospheric amplification is the relative warming between the upper troposphere (T24) and the lower-middle troposphere (TLT). We show that most

GCMs have excessive T24/TLT amplification compared to satellite microwave observations, even when models are forced with prescribed sea-surface temperatures (SSTs). A number of possible reasons for this discrepancy are assessed. Observational uncertainty in the satellite microwave record is substantial and, when taken into account, many models agree with observations within the observational uncertainty range, though about half of the model ensemble members considered still have significant discrepancies compared to observations. Our findings indicate that the prescribed ozone and stratospheric aerosol forcings do not effect T24/TLT amplification in models. On the other hand, model parameterizations for convection and microphysics and, to a lesser degree, uncertainty in the prescribed SST dataset can influence model amplification behavior and bring models into closer accord with observations. In all, significant T24/TLT discrepancies between models and observations remain, but may be reduced with improved model parameterizations.

An underlying motivation for understanding the structure of atmospheric warming is that it is responsible for a large negative lapse rate feedback in future climate simulations. To understand factors that control the global lapse rate feedback across models, we use principal component analysis to find the modes of variability that best explain variance in the local lapse rate feedback. We find that models exhibit marked variability in the lapse rate feedback in the southern hemisphere extratropics. This mode is strongly correlated with the global average lapse rate feedback and is largely a function of the competing influence of tropical and Antarctic surface warming. We show that muted southern ocean sea surface warming and the non-local influence of tropical surface warming contributes to a highly variable lapse rate feedback in the sub-Antarctic across models. This behavior is dissimilar to northern hemisphere high latitudes, which are characterized by strong Arctic amplification and a relatively uniform local lapse rate feedback across GCMs. Climatological Antarctic sea ice extent influences Antarctic warming and, as a result, influences both the meridional profile of warming in the southern hemisphere and the global lapse rate feedback.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

This dissertation was made possible with the support of a number of colleagues, friends, and family:

The Department of Atmospheric Sciences represents a vibrant community of curious scientists. My interactions with department members motivated me to explore challenging and interesting problems and gave me the skills and education necesary to embark on increasingly independent research. The technical expertise and support from individuals in the department greatly accelerated my research endeavors, helping me focus on scientific issues. My committee was generous with their time and provided a number of ideas which improved this dissertation and ongoing research efforts. Much of my thesis work builds upon the ideas of my academic advisor who has patiently helped me grow as an atmospheric scientist, supported my endeavors within and outside of the study of atmospheric sciences, and imparted endless encouragement that has given me the confidence to pursue research in this field.

Friends and family members have supplied ongoing support throughout my time in graduate school. These individuals have shaped my life in unique ways, helping to make me a better scientist and a better person. They have helped me through challenging obstacles and difficult choices. Thank you to all of those who have helped me along the way – your contributions, large and small, have been important and have not gone unnoticed.

Chapter 1
INTRODUCTION

1.1 Background

Atmospheric temperature is a critical metric for understanding climate variability and change. Atmospheric temperature responds to a range of phenomenon such as the El Niño-Southern Oscillation [Yulaeva and Wallace, 1994], poleward movement of the subtropical jet stream [Fu et al., 2006], ozone concentration changes [Young et al., 2014], Brewer-Dobson circulation variability [Lin et al., 2009], and anthropogenic climate change [Fu et al., 2004]. The vertical structure of atmospheric warming can also influence important phenomenon such as the behavior of tropical cyclones [Vecchi et al., 2013], the Brewer-Dobson circulation [Lin and Fu, 2013], and the lapse rate feedback [Soden and Held, 2006]. This basic measure of climate is thus relevant to a number of atmospheric phenomena on a range of timescales.

A critical dataset for measuring atmospheric temperature is derived from microwave sounding unit (MSU) measurements since 1978 and advanced MSU (AMSU) measurements since 1998. These observations measure thermal microwave emissions from atmospheric oxygen, which can be related to the temperature in deep atmospheric layers. MSU/AMSU instruments have flown on over 15 satellites and there are continuous measurements since November 1978, which makes it possible to intercalibrate the various records to create a long-term homogenized time series. These observations were originally intended to improve numerical weather forecasts [Thorne et al., 2011] and AMSU measurements currently contribute the most toward improving numerical weather prediction systems [e.g., Lorenc and Marriott, 2014]. Since these instruments were developed for weather and not climate purposes, they have instrument noise on the order of 0.3 K [Spencer and Christy, 1992]. Despite relatively large instrument noise, Spencer and Christy [1990] showed that satellite measure-

ments were remarkably stable over time. This was an important contribution because it indicated that it was possible to accurately estimate multi-decadal atmospheric temperature changes. Satellite microwave measurements are largely insensitive to clouds and water vapor, have near-global coverage on weekly timescales, have continuous measurements since the late 1970s, and measure several broad atmospheric layers [Grody et al., 2004, Chen and Zou, 2014]. For these reasons, the MSU/AMSU record has been an important dataset for estimating long-term changes in atmospheric temperatures.

Although other records of atmospheric temperature exist, they each have important weaknesses. Radiosonde measurements (i.e., measurements from weather balloons) have sparse spatial coverage, especially over the ocean, and have cooling biases that are difficult to quantify [Randel and Wu, 2006, Sherwood et al., 2005, Titchner et al., 2009]. Infrared observations, for example from the NASA Atmospheric InfraRed Sounder (AIRS) [Le Marshall et al., 2006], have better vertical resolution relative to microwave observations, but cannot penetrate clouds and have a much shorter continuous record with measurements since 2002. Radio occultation from Global Positioning System satellites (GPS RO) take very accurate measurements of temperature in the upper troposphere and lower stratosphere, although continuous data exists only since 2001. GPS RO measurements in the lower troposphere are sensitive to both temperature and water vapor, making it a less reliable temperature record near Earth's surface [Steiner et al., 2011]. So although there are a number of complementary temperature datasets, none match the time and space scales of the microwave record, and emphasis has been placed on the MSU/AMSU measurements for assessing long-term atmospheric temperature trends.

The microwave record is not without issue, though. Every few years, novel biases are discovered and corrected, including: the effects of satellite orbit altitude decay [e.g., Wentz and Schabel, 1998], warm target instrument calibration issues [e.g., Christy et al., 2000], corrections for changes in diurnal sampling [e.g., Mears and Wentz, 2005], and channel frequency drift [e.g., Zou and Wang, 2011]. The effect of some of these bias corrections can be quite large; the diurnal correction and the warm target correction for the tropical mid-troposphere

is 0.040 K decade^{-1} and -0.023 K decade^{-1}, respectively, while the reported trend is 0.115 $\pm$ 0.024 K decade^{-1} [Po-Chedley et al., 2015]. Over land, the tropical diurnal correction (0.160 K decade^{-1}) is larger than the trend itself (0.104 K decade^{-1}) [Po-Chedley et al., 2015]. Furthermore, tropical trend differences between teams that homogenize MSU/AMSU records differ by more than a factor of three and several co-orbiting satellites have large relative trends indicating that not all biases have been removed [Mears et al., 2011, Po-Chedley et al., 2015]. Recent analyses have suggested that the radiometer measurement frequency may not be stable for all instruments [Dee and Uppala, 2009, Lu and Bell, 2014] and that residual non-linear calibration biases and diurnal drift corrections [Christy et al., 2010, Po-Chedley et al., 2015] could affect MSU/AMSU time series homogenization efforts. It is therefore important to continue to improve the microwave temperature record.

This thesis emphasizes the tropics for several reasons. First, temperature trends and variability are amplified aloft relative to the surface, and this tropospheric vertical amplification is largest in the tropics [Solomon et al., 2007]. Therefore, temperature trends in the tropical troposphere should be relatively large compared to surface or lower tropospheric trends. Second, important feedbacks, such as the lapse rate feedback are directly linked to the warming in the tropical troposphere [Soden and Held, 2006] and cloud feedbacks have been linked indirectly via the effect of tropical moist static stability on the overturning circulation [Tomassini et al., 2015]. Last, homogenized microwave temperature trends are produced by the University of Alabama at Huntsville (UAH) [Christy et al., 1998, 2000, 2003], Remote Sensing Systems (RSS) [Mears et al., 2003, Mears and Wentz, 2009a,b, Mears et al., 2011], and the National Oceanic and Atmospheric Administration (NOAA) Center for Satellite Applications and Research (STAR) [Zou et al., 2006, 2009, Zou and Wang, 2010, 2011, Wang and Zou, 2014] and the trend discrepancies between these datasets are largest in the tropics [Po-Chedley, 2012].

A fundamental expectation in the tropics is that the troposphere will warm faster than the surface – a phenomenon known as tropical tropospheric vertical amplification. Vertical amplification results because the atmosphere is linked to the surface via moist convection.

As the surface warms, convective updrafts of saturated air will condense more water due to the Clausius-Clapeyron relationship, which leads to enhanced latent heat release. This moist adiabatic warming due to latent heat release has a greater effect with increasing height [Sobel et al., 2002]. Warming in convective regions is then quickly communicated throughout the entire tropics via internal gravity waves [Bretherton and Sobel, 2003]. A number of studies have shown that when accounting for natural variability and uncertainty, tropical tropospheric trends between models and general circulation models (GCMs) are not significantly different from one another [e.g., Flannaghan et al., 2014, Santer et al., 2008, Thorne et al., 2007, Pralungo and Haimberger, 2015, Sherwood and Nishant, 2015, Po-Chedley et al., 2015, Kamae et al., 2015]. While this is generally true for the bulk of the tropical troposphere, analysis from radiosondes and satellite temperature observations have suggested that GCMs warm the tropical upper troposphere too much even when natural variability is taken into account [Po-Chedley and Fu, 2012a, Mitchell et al., 2013a, Seidel et al., 2012, Sherwood and Nishant, 2015]. Recent analyses have shown that the pattern of sea surface temperature (SST) warming [Po-Chedley and Fu, 2012a, Garfinkel et al., 2013, Flannaghan et al., 2014] and the atmospheric forcing [Polvani and Solomon, 2012, Santer et al., 2014] can play an important role in atmospheric temperature trends, which could help explain discrepancies between models and observations for upper tropospheric warming.

While much of this dissertation focuses on the tropics, we also consider the global lapse rate feedback. In the high latitudes, lower tropospheric warming can exceed warming in the upper troposphere due to the near surface temperature inversion [e.g., Payne et al., 2015]. This difference in behavior between the tropics, where warming is larger aloft, and polar regions, where warming is largest near the surface, is important for the lapse rate feedback. The lapse rate feedback is the radiative response to vertically non-uniform atmospheric warming [Soden and Held, 2006]. Warming the upper troposphere allows longwave radiation to escape to space more efficiently than near-surface warming. When warming is largest near the upper troposphere (surface) a negative (positive) feedback results. Thus the lapse rate feedback is robustly negative in the tropics and positive at the poles, which reinforces

Arctic amplification [Pithan and Mauritsen, 2014]. Given the competing behavior between the tropics, where surface warming is coupled to the upper troposphere, and the poles, where surface warming can be trapped in the lower troposphere, the meridional pattern of warming is important in determining the global average lapse rate feedback [Soden and Held, 2006].

This dissertation focuses on understanding both the differences in independent observational estimates of tropical tropospheric warming and differences between GCMs and satellite observations of the structure of atmospheric temperature change. Understanding such discrepancies across observational records is important, because satellite-borne microwave observations represent the only long-term (38 year), global, and vertically resolved temperature dataset. These records have led to the discovery of important atmospheric phenomenon [Fu et al., 2006] and are critical for climate monitoring [Spencer and Christy, 1990], attribution of climate change [Santer et al., 2005], and assessing the fidelity in which GCMs simulate the recent climate [Fu et al., 2011]. Given the wide array of climate applications, such data are used in national [Karl et al., 2006] and international climate assessments [Solomon et al., 2007] and are used to inform public policy [Christy et al., 2015]. It is also critical to understand differences between model simulations and observations of atmospheric temperature change. Discrepancies between observations and GCMs could be a result of observational error, biases in data used to force models, or problems with GCM parameterizations. Thus quantifying model-observational discrepancies and factors that contribute to such differences can help improve both observations and GCMs that are used to project future changes in climate. This dissertation also seeks to understand factors that influence the model spread in the tropical and global lapse rate feedback. This is important because the global lapse rate feedback is highly variable across models and covaries with other important feedbacks, such as the water vapor feedback [Caldwell et al., 2016]. Thus, understanding intermodel differences in this feedback may shed light on some of the underlying physics that lead to different model responses to increased atmospheric greenhouse gas concentrations.

1.2 Overview of Research

This thesis investigates a number of issues related to the rate of atmospheric warming and its representation in general circulation models. In Chapter 2, we begin by investigating the large differences between various MSU/AMSU datasets in the rate of tropical mid-tropospheric warming. In response to suggestions that these differences result from different strategies for correcting changes in diurnal sampling [e.g., Christy et al., 2011], we propose a novel diurnal correction for satellite measurements of tropical mid-tropospheric temperature. A key result of this chapter is that models and observations largely agree on the amplification of warming between the surface and the troposphere. In Chapter 3, we explore tropical vertical tropospheric amplification in more detail, assessing model-observational differences between the upper and lower-middle troposphere. This work follows research by Fu et al. [2011], which compared model and satellite estimates of warming in the upper and lower-middle tropical troposphere. Their results suggested that coupled atmosphere-ocean models from the Coupled Model Intercomparison Project Phase 3 (CMIP3) tended to have too much warming in the tropical upper troposphere compared to the lower-middle troposphere. We expand on this research using coupled atmosphere-ocean GCMs from CMIP5. We also utilize atmospheric GCMs with prescribed SSTs to reduce the role of natural variability in any model-observational differences in the structure of atmospheric warming. We demonstrate that the latest generation of models (CMIP5) still tend to exaggerate warming in the tropical upper troposphere, even when constrained with prescribed SSTs. In Chapter 4, we investigate the degree to which model forcing, model physics, observational uncertainty in the prescribed pattern of SST warming, and observational uncertainty in the satellite temperature record affects this result. This analysis demonstrates that model-observational differences in amplification are influenced by uncertainty in the pattern of SST changes and model parameterizations for moist physics. We also test various satellite metrics for the lapse rate feedback and the degree to which warming over the satellite era corresponds to projected changes in the vertical structure of warming. We find that recent multi-decadal amplification

is only weakly predictive of future warming behavior. Furthermore, the strength of the tropical lapse rate feedback explains less than half of the intermodel spread in the global lapse rate feedback. In Chapter 5, we look beyond the tropics and attempt to understand controls on the intermodel spread of the global lapse rate feedback. While it has been long been known that the meridional pattern of warming is predictive of the global lapse rate feedback [Soden and Held, 2006], we show that this largely results because intermodel differences in meridional warming modulates the local lapse rate feedback strength in the southern extratropics. Across models, the southern extratropics exhibits widely varying feedback strengths that stem from suppressed southern ocean warming and the remote influence of the tropics on extratropical upper tropospheric warming. Antarctic surface warming (60°S – 90°S), largely controlled by models' initial sea ice extent, also plays a significant role in determining the global lapse rate feedback.

Chapter 2

REMOVING DIURNAL CYCLE CONTAMINATION IN SATELLITE-DERIVED TROPOSPHERIC TEMPERATURES

Stephen Po-Chedley, Tyler J. Thorsen, and Qiang Fu, 2015, "Removing diurnal cycle contamination in satellite-derived tropospheric temperatures: Understanding tropical tropospheric trend discrepancies," Journal of Climate, doi: http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-13-00767.1.

2.1 Abstract

Independent research teams have constructed long-term tropical mid-tropospheric temperature (TMT) time series using satellite microwave sounding unit (MSU) and advanced MSU (AMSU) measurements. Despite careful efforts to homogenize the MSU/AMSU measure-

ments, tropical TMT trends beginning in 1979 disagree by a factor of three. Previous studies suggest that the discrepancy in tropical TMT temperature trends is caused by differences in both the NOAA-9 warm target factor and diurnal drift corrections. This work introduces a new observationally-based method for removing biases related to satellite diurnal drift. Over land, we find that the derived diurnal correction is similar to a general circulation model (GCM) diurnal cycle. Over ocean, our diurnal corrections have a negligible effect on TMT trends, indicating that oceanic biases are small. It is demonstrated that this method is effective at removing biases between co-orbiting satellites and biases between nodes of individual satellites. Using our homogenized TMT dataset, the ratio of tropical tropospheric temperature trends relative to surface temperature trends is in accord with the ratio from GCMs. It is shown that bias corrections for diurnal drift based on a GCM produce tropical trends very similar to those from the observationally-based correction, with a trend differences smaller than 0.02 K decade^{-1}. Differences among various TMT datasets are explored further. Large differences in tropical TMT trends between this work and the University of Alabama in Huntsville (UAH) are attributed to differences in the treatment of the NOAA-9 target factor and the diurnal cycle correction.

2.2 Introduction

Atmospheric deep layer temperature trends from satellite microwave sounding unit (MSU) and advanced MSU (AMSU) measurements are frequently used both as a measure of climate change and as a reference to evaluate global circulation models (GCMs) [e.g., Wallace and Coauthors, 2000, Santer et al., 2005, Karl et al., 2006, Solomon et al., 2007, Fu et al., 2011, Po-Chedley, 2012, Mitchell et al., 2013a]. A common metric is to compare the ratio of warming between the troposphere and the surface in the tropics since GCMs and basic atmospheric physics suggest that tropospheric temperature change should be amplified relative to surface temperature change.

Over the past decade many studies based on MSU/AMSU observations have shown that the tropical tropospheric temperature increases more than the surface temperature on multi-

decadal time scales [Fu et al., 2011, Fu and Johanson, 2005, Mears and Wentz, 2005, Thorne et al., 2007, Santer et al., 2008] while others suggest reduced tropospheric warming relative to the surface [e.g., Christy et al., 2007, 2010, Douglass et al., 2008, Mckitrick et al., 2010]. Our understanding of tropical tropospheric temperature trends are complicated by the fact that tropical radiosonde stations are limited in number, their trends are cold biased, and the bias magnitude cannot be accurately determined [e.g., Sherwood et al., 2005, Randel and Wu, 2006, Titchner et al., 2009]. Complications with radiosonde measurements reinforce the need to ensure that satellite records are unbiased in order to advance our understanding of climate and climate change.

Three research teams, including the University of Alabama in Huntsville (UAH), Remote Sensing Systems (RSS), and the National Oceanic and Atmospheric Administration (NOAA) Center for Satellite Applications and Research (STAR), have developed up-to-date, homogenized datasets for the mid-tropospheric temperature (TMT) [Christy et al., 2003, Mears et al., 2003, Zou and Wang, 2011]. These groups employ over thirty years of temperature measurements from MSU and, starting in 1998, AMSU. These instruments have flown onboard a number of satellites beginning with MSU on Tiros-N in late 1978 through AMSU on NOAA-19. Despite using the same basic radiometer measurements, tropical TMT trend differences between these groups differ by a factor of three. The UAH dataset has a tropical TMT trend that is close to zero (0.029 K decade^{-1}) whereas RSS and NOAA have trend values of 0.089 and 0.105 K decade^{-1}, respectively. Zou et al. [2009] analyzed TMT trends over the tropical ocean, where the diurnal drift effects are negligible for trend comparisons. They found large trend differences between MSU/AMSU datasets and enhanced tropical amplification relative to both RSS and UAH. Reasons for TMT trend differences stem from different processing choices and bias corrections made by the three research teams, including the choice of satellites included [Mears et al., 2003], corrections for the influence of the warm target temperature on the measured brightness temperature [e.g., Christy et al., 2000, Mears et al., 2003, Po-Chedley and Fu, 2012b], and corrections for the drift of satellites through the diurnal cycle [e.g., Christy et al., 2000, Fu and Johanson, 2005, Mears and Wentz, 2005].

Recent studies by Po-Chedley and Fu [2012b, 2013] suggest that UAH uses a biased NOAA-9 warm target factor that results in artificial cooling in the UAH TMT time series, though UAH has not changed its treatment of NOAA-9 Christy and Spencer [2013]. Accounting for this bias, however, can only explain part of the discrepancy in TMT trends between UAH and other teams over the tropics. Much of the remaining discrepancies in TMT trends are likely caused by differences in diurnal drift corrections employed by various MSU/AMSU research teams.

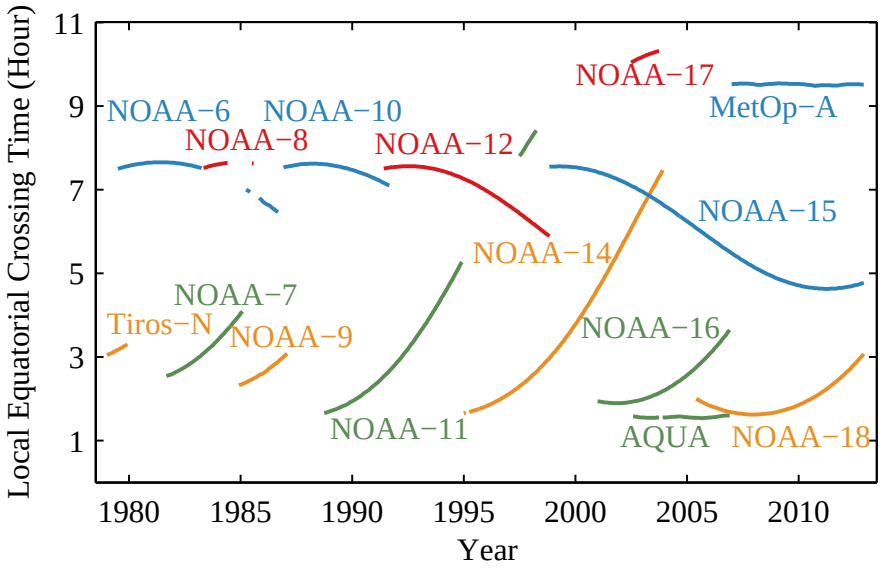

Figure 2.1: LECT for the satellites used in this study. We include LECT values for all the months in which satellites are incorporated into our analysis. LECTs in this figure are for the descending node. LECTs for the ascending node are 12 h later.

As a polar orbiting satellite drifts east or west relative to the sun, the local sampling time on Earth, generally indicated by the local equatorial crossing time (LECT) changes accordingly (Fig. 2.1). As the satellite drifts over time, changes in the satellite-measured brightness temperature due to changes in the local sampling time can be much larger than the temperature change associated with the long-term climate change. This is especially true over land where the surface skin temperature has a large diurnal range. It is therefore

very important to remove the effect of diurnal drift on the measured brightness temperature when estimating the magnitude of decadal temperature change. The RSS and NOAA research teams apply a drift correction based on the diurnal cycle from a GCM [Mears et al., 2003, Zou and Wang, 2011] whereas UAH produces an MSU TMT diurnal correction based on temperature comparisons between three co-orbiting satellites carrying AMSU with different local sampling times. UAH does not yet correct the diurnal drift for satellites carrying AMSU, because they attempt to use these satellites during periods when diurnal drift is small (Dr. John Christy, personal communication, 2013). By comparing MSU/AMSU datasets with radiosondes, several studies suggest that diurnal adjustments derived from a GCM and/or other bias corrections lead to a warm bias in the trends for lower tropospheric temperature data (TLT) for RSS [Christy and Norris, 2006, Christy et al., 2007, Randall and Herman, 2008, Christy et al., 2010]. Mears et al. [2012] found that many of the conclusions from these studies were dependent on the datasets and time considered and found that the methodological choices significantly impacted the results. Mears et al. [2012] also point out that the time period considered is complicated by effects from ENSO and the eruption of Mount Pinatubo. In addition, radiosonde networks undergo artificial cooling, which may not be completely removed [Sherwood et al., 2005, 2008]. In general, it is impossible to conclude which dataset is most accurate when comparing two potentially biased measurements [Mears et al., 2012].

This paper focuses on developing a novel approach to the diurnal drift correction based on MSU/AMSU observations. We use MSU/AMSU level 1 data calibrated with an instrument calibration technique that minimizes intersatellite differences, reducing intersatellite residuals by an order of magnitude compared to the prelaunch MSU/AMSU calibration [Zou et al., 2009]. This state-of-the-art instrument calibration allows us to derive a diurnal drift correction by attributing remaining intersatellite differences to diurnal drift biases. In Section 2.3 we describe the data that we use in this study and list comparison datasets. We describe our basic processing decisions for homogenizing MSU/AMSU level 1 data in Section 2.4. In Section 2.5 we present an observationally-based approach for removing biases related to the

diurnal drift of satellites and compare this correction to the GCM derived correction. We then show the efficacy of our approach with an emphasis on the tropics. Our homogenized TMT time series is compared to the UAH, RSS, and NOAA TMT products in Section 2.6 and in Section 2.7 we attempt to understand major underlying causes for differences. In Section 2.8 we conclude with a summary and discussion of our results.

2.3 Data

2.3.1 Level 1 MSU/AMSU dataset

In this study, we use the most recent version of the level 1C (L1C) MSU/AMSU data produced by NOAA STAR, which was newly released in 2013. The dataset has been described by Zou and Wang [2011, 2013]. The L1C data has been intercalibrated using simultaneous nadir overpasses (SNO) [Zou et al., 2006, Zou and Wang, 2011]. This on-orbit calibration technique utilizes simultaneous nadir observations in polar regions from co-orbiting satellites to solve for the MSU/AMSU calibration coefficients such that intersatellite biases related to the warm target temperature are minimized. The L1C swath data minimizes or removes a number of biases using the NOAA STAR Integrated Microwave Inter-Calibration Approach (IMICA).

The NOAA IMICA calibration is both effective and important to our analysis. The limb view correction used in the IMICA removes temperature differences due to vertical sampling differences across view angles with accuracy below the noise of the instrument and the corrected temperatures have similar distributions for the limb and near-nadir measurements [Goldberg et al., 2001]. Zou et al. [2009] found that the trends are independent of the number of footprints used implying that the limb correction does not have a large influence on our results and our product effectively represents the nadir view of the MSU/AMSU instrument. NOAA-16 exhibits steady bias drifts in the radiometer signal counts, which leads to a spurious warm drift of 0.3 K decade^{-1} in the brightness temperatures [Zou and Wang, 2011]. For this reason, RSS does not include NOAA-16 in its analysis. This bias drift is removed in

the IMICA analysis, which allows us to consider NOAA-16 data in our analysis, though the inclusion of this data does not significantly affect the diurnal cycle correction derived in this work. The IMICA calibration also minimizes biases related to sun heating instrument variability and the scene temperature, which may introduce regional and seasonal biases and spurious trends [Zou and Wang, 2011]. Using IMICA, the global ocean mean intersatellite bias and standard deviation is on the order of 0.05 K and 0.03 K, respectively, compared to the prelaunch calibration which has biases on the order of 0.5 K and 0.13 K [Zou and Wang, 2013]. IMICA brings intersatellite trends close to zero over the ocean [Zou et al., 2009]. Since the diurnal drift effect is small over the ocean, this implies that the instrument calibration is excellent.

In summary, the IMICA procedure removes biases including i) global mean inter-satellite biases, ii) scene temperature dependent biases due to inaccurate calibration non-linearity, iii) sun-heating induced instrument variability biases, iv) drift bias for NOAA-16 and v) limb adjustment to nadir [Zou and Wang, 2011, Dr. Cheng-Zhi Zou, personal communication 2013]. Because the IMICA calibration eliminates or minimizes these biases, we are able to largely attribute remaining intersatellite differences to diurnal drift effects, though we cannot completely rule out unknown biases. Note that any unknown residual biases that are correlated with the diurnal cycle but independent of individual satellite ascending and descending nodes will effectively be removed through our diurnal cycle correction.

2.3.2 Comparison datasets

In order to compare our MSU/AMSU TMT homogenization and merging efforts to independent datasets, we utilize the most recent versions of existing MSU/AMSU climate records, including UAH TMT v5.6 [Christy et al., 1998, 2000, 2003], RSS TMT v3.3 [Mears et al., 2003, Mears and Wentz, 2009a, Mears et al., 2011], and NOAA TMT v3.0 [Zou et al., 2006, 2009, Zou and Wang, 2010, 2011]. We also employ lower stratospheric channel data, TLS, from these groups in order to remove the stratospheric signal from the TMT dataset [e.g., Fu et al., 2004]. These datasets represent independently homogenized monthly MSU/AMSU

datasets that can be used for comparison to our work. We also utilize HadCRUT4 data as a measure of surface temperature trends over land and ocean [Morice et al., 2012].

The largest differences between MSU/AMSU datasets relate to how each group removes instrument calibration drift and biases due to diurnal drift. UAH applies a diurnal drift correction for MSU based on the diurnal cycle derived from co-orbiting AMSU measurements and avoids using data when large diurnal drifts are present. RSS and NOAA utilize corrections based on the diurnal cycle simulated by a climate model. RSS and UAH generally utilize the prelaunch calibration whereas NOAA (and this analysis) utilizes data that is corrected using IMICA. All of the datasets remove residuals related to the warm target temperature onboard the satellite, though differences in methodology lead to different bias corrections [e.g., Po-Chedley and Fu, 2013]. Another difference relates to how data is included in the merging procedure. RSS and NOAA use a consensus approach that averages all overlapping satellites together. UAH chooses stable backbone reference satellites and ignores data when diurnal drifts noticeably affect the time series [e.g., Christy and Spencer, 2013]. Other differences include corrections to limb data, removing differences between MSU and AMSU measurements, and how scene-dependent temperature biases are removed.

2.4 Procedures and processing decisions for L1C MSU/AMSU data

We begin our analysis with NOAA STAR L1C data, utilizing MSU channel 2 and AMSU channel 5 data over 1979 through 2012. These two channels have historically been used to represent the mid-tropospheric temperature, TMT. We use the five central view angles from MSU and 12 central view angles for AMSU [e.g., Mears and Wentz, 2009a]. We correct the L1C data for diurnal drift by adjusting the data to local noon utilizing a GCM derived diurnal cycle, while we also process L1C data with no diurnal correction, which we will use to derive and apply an observationally-based diurnal correction developed in this study (see section 4). For quality control purposes, we remove measurements that are greater than 5σ above or below the daily zonal (2.5°) average brightness temperature for each satellite. Following Wang and Zou [2014], we also remove observations that contain precipitation. Daily data is

then placed into 2.5° x 2.5° grids for the ascending and descending nodes. Measurements that are more than 4σ above or below the long-term grid point mean are eliminated. Both of these quality control steps remove less than 0.07% of the data for each satellite. We also manually removed data in cases when the data is suspect, including Tiros-N after 1979, NOAA-6 after day 274 of 1986, NOAA-9 after day 62 of 1987, NOAA-14 after 2003, NOAA-16 and AQUA beginning in 2007, and NOAA-17 after day 300 of 2003. Several individual days were also removed due to obvious data quality issues.

After quality control and gridding, we follow Mears et al. [2003] to remove the warm target temperature effect. The warm target temperature effect was described by Christy et al. [2000], in which it was determined that the evolution of the warm target calibration temperature onboard various satellites could explain intersatellite differences between co-orbiting satellites. This issue is caused by sun heating induced instrument variability and is largely removed via the IMICA calibration [Zou and Wang, 2011], but we apply the warm target temperature calibration to remove any remaining residuals. Following Mears et al. [2003], we used global mean oceanic pentad (five day averages) data to linearly regress out any intersatellite difference residuals related to the warm target temperature onboard each satellite. The regression uses intersatellite brightness temperature differences as predictands and warm target temperature anomalies onboard each satellite as predictors. The coefficients of this regression are referred to as warm target factors. In constructing global mean pentads, we required that each pentad had data from 85% of the ocean grid cells and that there were at least three days of data. By using oceanic data to determine the warm target factors, the influence of the diurnal cycle is minimized since the diurnal cycle over the ocean is approximately an order of magnitude smaller than over land. We solved for our target factors using oceanic data that was diurnally corrected with the RSS climate model data, but the differences in our target factors and the resulting temperature trends are negligible if we use oceanic data that has no diurnal correction.

After solving for the target factors, we create monthly averages for each satellite's ascending and descending nodes and utilize our target factors to remove the warm target

temperature effect from each grid cell. In order to remove MSU/AMSU differences related to differences in the weighting function we take the monthly average value for NOAA-15 minus NOAA-14 over 1999 through 2003 to create an average monthly offset climatology for each satellite node following Mears and Wentz [2009a]. In computing our monthly gridded data, this offset climatology is removed from AMSU data to match AMSU to MSU as in Mears and Wentz [2009a].

2.5 Diurnal drift corrections

2.5.1 GCM-based diurnal drift correction

In order to remove diurnal drift from TMT measurements, the RSS team uses the diurnal cycle simulated from five years of model simulation in the National Center for Atmospheric Research Community Climate Model Version 3 (CCSM3) [Kiehl et al., 1996, Mears et al., 2003]. Mears et al. [2003] derived the TMT diurnal cycle with hourly data for each month of the year and each view angle using microwave radiative transfer and surface emissivity models [Wentz and Meissner, 1999]. We use the RSS diurnal correction and scale the diurnal amplitude by 0.875 for MSU and 0.917 for AMSU following Zou et al. [2009].

The diurnal drift correction is sensitive to the applied diurnal cycle in terms of both phase and amplitude [Po-Chedley, 2012]. Dai and Trenberth [2004] showed that the GCM simulated diurnal cycle in CCSM2 likely has biases. Therefore, model biases may impact diurnal bias corrections applied from GCM data. It has been suggested that RSS and NOAA overcorrect for diurnal drift, which leads to spurious warming relative to UAH in the tropics [Christy et al., 2010, 2011]. On the other hand, NOAA [Zou et al., 2009] and RSS [Mears et al., 2003] have compared the magnitude of the GCM derived diurnal drift correction with global intersatellite residuals and have shown that the adjustments applied minimize intersatellite residuals. Other studies have also found consistency between GCM-derived and observed diurnal cycle climatology in tropospheric channels of the High-Resolution Infrared Radiation Sounder, which provides evidence that diurnal cycle climatology from a GCM should be

reasonably valid for microwave measurements [Jackson and Soden, 2007, MacKenzie et al., 2012].

Because it is possible that GCM-derived diurnal cycle corrections have biases, we develop a regression technique to derive a TMT diurnal cycle correction based on satellite observations for the purpose of homogenizing the MSU/AMSU observations. This diurnal correction will be used as a comparison to understand the corrections implemented by NOAA, RSS, and UAH.

2.5.2 Observationally-based technique for removing diurnal drift biases

In analyzing data for which no diurnal corrections have been applied, the implications of the diurnal drift bias become immediately clear for long-term trends. Over land, brightness temperature differences between co-orbiting satellites or differences between the ascending (PM) and descending (AM) node of a single satellite can drift by more than 1 K as the satellite drifts toward earlier or later local measurement times. Figure 2.1 shows the LECT for each satellite over the satellite's lifetime for the descending node (the ascending node LECT is 12 hours later). This figure only includes the time periods in which we used data from each satellite. Several satellites drift more than an hour during their lifetime, which means that a large diurnal signal will be aliased into their TMT measurements.

In order to remove the effects of diurnal drift, we develop a regression technique for estimating a TMT diurnal cycle correction based on MSU/AMSU measurements that include diurnal drift information. In section 3, we created a gridded monthly TMT time series that included warm target bias corrections, but not diurnal drift corrections. This uncorrected data will allow us to estimate a diurnal cycle correction. In our analysis, most satellites drift between LECT values of 1:30 and 8 am (pm) for the descending (ascending) node (see Figure 2.1). Over the tropical ocean, the GCM diurnal cycle is well approximated with a quadratic function during these hours for both the AM and PM nodes. Over tropical land, however, the diurnal cycle has more inflection points and is better approximated with a cubic fit (the normalized residuals for fits to the GCM diurnal cycle are reduced by a factor of 2-3 using

a cubic rather than a quadratic function). We can then write down the measured land TMT brightness temperature for given month as:

$$T^M_{i,N} = T^E + \alpha_i T^W_i + a_{I,N} t_i^3 + b_{I,N} t_i^2 + c_{I,N} t_i + d_{I,N} + \epsilon_i \tag{2.1}$$

where T^E is the actual unbiased Earth temperature, T^M is the measured temperature, α is the warm target factor, T^W is the warm target temperature anomaly, and ϵ represents unresolved errors for the ith satellite. The constants a, b, and c are the coefficients for the diurnal cycle correction and d represents a constant offset for each node of each satellite. Our diurnal cycle is dependent on the LECT, t, and the coefficients vary by the MSU or AMSU instrument (I) and the node (N). The measured ocean brightness temperature can similarly be written as:

$$T^M_{i,N} = T^E + \alpha_i T^W_i + e_{I,N} t_i^2 + f_{I,N} t_i + g_{I,N} + \epsilon_i \tag{2.2}$$

where, in this case, e and f are the diurnal cycle coefficients and g is a constant offset.

Our numerical regression technique relies on being able to attribute differences between satellites to a common diurnal cycle, which we represent via 12 coefficients for land (three terms in the cubic fit times two instruments and two nodes) and 8 coefficients for the ocean (two terms in the quadratic fit times two instruments and two nodes). By first removing the warm target effect in the monthly gridding process, we can ignore the warm target temperature term. Our strategy is then to fit intersatellite differences using the following equation for land based on Eq. 2.1:

$$\begin{aligned} T^M_{i,N_1} - T^M_{j,N_2} = (a_{I_1,N_1} t_i^3 + b_{I_1,N_1} t_i^2 + c_{I_1,N_1} t_i + d_{I_1,N_1}) \\ - (a_{I_2,N_2} t_i^3 + b_{I_2,N_2} t_i^2 + c_{I_2,N_2} t_i + d_{I_2,N_2}) + \epsilon_i - \epsilon_j \end{aligned} \tag{2.3}$$

and the following equation for ocean from Eq. 2.2:

$$T^M_{i,N_1} - T^M_{j,N_2} = (e_{I_1,N_1} t_i^2 + f_{I_1,N_1} t_i + g_{I_1,N_1}) \\ - (e_{I_2,N_2} t_i^2 + f_{I_2,N_2} t_i + g_{I_2,N_2}) + \epsilon_i - \epsilon_j \quad (2.4)$$

where i and j represent the satellites under consideration. The coefficients for our diurnal cycle correction vary by instrument (I) and node (N), but it is not necessary to use the same node or instrument for satellites i and j in Eqs. 2.3 and 2.4. In other words, we do not need to compare two satellites with the same instrument or node. We have therefore denoted the instrument (I) and node (N) with subscripts 1 and 2 (for the two measurements being compared) to make it clear that N_1 and N_2 and I_1 and I_2 are not necessarily equal to one another.

From Eqs. 2.3 and 2.4, we can use monthly brightness temperature differences from co-orbiting satellites (i.e., $i \neq j$) and differences between the ascending and descending node of the same satellite (i.e., $i = j$ and $N_1 \neq N_2$) to solve for the diurnal cycle coefficients using multiple linear regression. This procedure is similar to that used for determining warm target factors for each satellite, except in this case we are fitting higher order functions (a quadratic and cubic for ocean and land, respectively) and we use LECT as a predictor instead of the warm target temperature anomaly. We use LECT values between 0 and 12 for both the ascending and descending nodes, even though the ascending node is in the afternoon. In our analysis we use monthly means, which is a timescale sufficient to track the gradual changes in LECT, but still leaves more than 4,000 predictands. The NOAA-10 descending node constant offset is set to zero. We focus on the gradual diurnal drift by removing the long-term mean MSU/AMSU seasonal cycle from each satellite although we note that the diurnal drift may also induce small changes in the seasonal cycle.

Some analyses have used Fourier series to solve for the diurnal cycle in infrared sounder [e.g., Lindfors et al., 2011] and microwave [e.g., Mo, 2009, Kottayil et al., 2013] measurements. This approach uses overlapping satellites and fits brightness temperature measurements as a function of local time using a second order Fourier series. It requires that measurements

are absolutely calibrated, which cannot be guaranteed for MSU/AMSU measurements; the absolute uncertainty in the IMICA calibration is 0.5 - 1.0 K with global mean intersatellite biases between 0.1 and 0.2 K [Zou and Wang, 2013]. Our method instead tracks the differences between satellite measurements over time and attributes these differences to diurnal sampling biases. Note that in our method we remove constant offsets for each node of each satellite, which means we do not require an absolute calibration of brightness temperature measurements. We also use the differences between satellites, which means that natural variability common to both satellites is removed, leaving diurnal sampling bias as the biggest component of slowly evolving intersatellite differences. Our method uses polynomial fits since polynomials reproduce the diurnal cycles from climate models and are easily incorporated into our linear regression approach.

To assess the impact of the seasonal cycle on the diurnal drift bias removal, we merged our time series using a GCM correction with and without a seasonal cycle. The tropical land trend difference between the merged time series with and without a seasonal cycle was 0.006 K decade^{-1}. Over tropical oceans the difference is 0.001 K decade^{-1}. These small differences are expected since the seasonal cycle in the tropics is much smaller compared to higher latitudes, though a diurnal drift bias correction that accounts for the seasonal cycle should be developed in future efforts. Another important assumption we make in our formulation is that the diurnal cycle does not change over time. This is an assumption also made by RSS, NOAA, and UAH though it is not necessarily true [e.g., Hansen et al., 1995]. We expect that although this assumption is not strictly true, diurnal cycle changes for TMT over the tropics are small compared to the diurnal drift bias we remove in our technique. Note that by using carefully calibrated L1C data, instrument calibration issues are small and our regression is addressing diurnal drift biases. It is, however, still possible that our technique also removes residual instrument calibration biases that are correlated with the LECT. In this case our observationally-derived diurnal cycle correction may deviate from the real diurnal cycle, especially over the ocean where the diurnal cycle signal is small. But the trends based on this observationally-derived correction are expected to be more accurate

than those using the real diurnal cycle because any residual instrument calibration biases related to satellite drift are further removed.

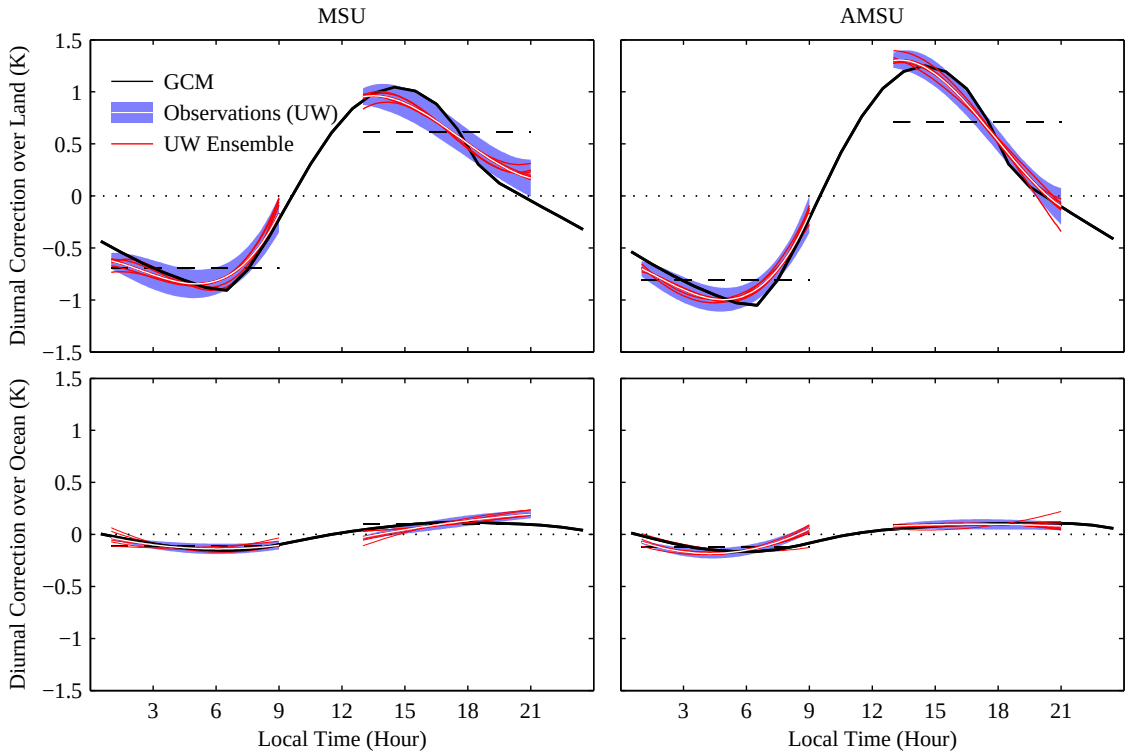

Figure 2.2: Tropical mean (20°N - 20°S) diurnal cycle corrections in TMT for (left) MSU and (right) AMSU over (top) land and (bottom) ocean from a GCM (black line) and from our regression analysis based on observations (white line). Blue shading represents the 95% confidence interval from our regressed diurnal cycle. We also calculated 15 diurnal cycle corrections by excluding data from each of the 15 satellites and recalculating the diurnal cycle corrections, which is represented by red lines. This ensemble shows the sensitivity to individual satellites. Since our regression removes a constant offset from each satellite node (so the absolute temperature difference between the AM and PM node is not preserved), we forced the mean of each of our regressed diurnal cycle segments (0100-0900 and 1300-2100 LT) to match the mean of the GCM over the same time period (black dashed line) to aid in comparing the GCM and regressed diurnal cycle. The GCM diurnal cycle shown here has a scaling factor of 1, though we scaled the MSU correction by 0.875 and the AMSU correction by 0.917.

We solve Eqs. 2.3 and 2.4 for the MSU/AMSU diurnal cycle correction using tropical (20°N - 20°S) mean time series for land and ocean. Table 2.1 presents the diurnal cycle correction coefficients from our multiple linear regression, which correspond to the constants presented in Eqs. 2.1 and 2.2. Figure 2.2 shows the tropical TMT diurnal cycle corrections derived from multiple linear regression using MSU/AMSU satellite observations compared to the diurnal cycles from the GCM [Mears et al., 2003] for MSU and AMSU over land and ocean. We show our diurnal cycle correction over ocean for completeness and transparency. Since the oceanic diurnal cycle should be small [e.g., Zou et al., 2009] this correction gives some indication of the degree that uncorrected biases remain in our time series after warm target calibration. For example, if we found large oceanic diurnal cycle corrections this would indicate that calibration errors remain in the time series. From Figure 2.2, we see that the ocean diurnal cycle corrections are small indicating that large biases do not remain. For this analysis we added white noise to our monthly average tropical mean brightness temperatures and re-solved for the warm target factors and diurnal coefficients 10,000 times in order to estimate the uncertainty in our regression technique. The error estimate from this Monte-Carlo analysis accounts for the uncertainty in our intersatellite offsets, warm target correction, and diurnal correction. Mears et al. [2011] do a detailed uncertainty estimate and find larger uncertainties by accounting for intersatellite trends between NOAA-14 and NOAA-15 and because they include diurnal uncertainty from several GCMs. We also solve for diurnal cycle corrections, which have had a single satellite of the 15 satellites removed from consideration (referred to as "UW Ensemble"). The results from this sensitivity analysis show that the derived diurnal cycle correction is not very sensitive to individual satellites (generally the sensitivity is smaller than our estimated error in the diurnal cycle correction over land and similar in magnitude to our estimated error over ocean). Removing NOAA-15 from the diurnal cycle calculation does reduce the magnitude of the AMSU AM ocean diurnal cycle, but also increases the magnitude of the AMSU PM ocean diurnal cycle correction. This likely occurs because NOAA-15 is an important satellite for solving for the AMSU diurnal cycle correction since it drifts through much of the diurnal cycle. When we remove NOAA-

	Land			Ocean	
	a	b	c	d	e
MSU AM	0.0038	-0.0290	-0.0005	0.0040	-0.0494
MSU PM	0.0029	-0.0485	0.1262	-0.0015	0.0419
AMSU AM	0.0037	-0.0209	-0.0652	0.0105	-0.0895
AMSU PM	0.0038	-0.0690	0.1667	-0.0009	0.0091

Table 2.1: Average value of the diurnal cycle correction co- efficients in Eqs. (1) and (2) over land and ocean in the tropics (20°N - 20°S) for the MSU and AMSU instruments.

15, the diurnal cycle later than 3pm for the ascending node and 3am for the descending node is unconstrained for AMSU measurements. We also solved for the diurnal cycle correction without applying a warm target correction and the resulting diurnal cycle corrections were negligibly different compared to those with a warm target correction (not shown). This result implies that instrument calibration drifts are not systematically aliased into our diurnal cycle.

Our derived diurnal cycle correction is similar to that produced from a climate model, but there are notable differences in the phase and amplitude over land for each node. However, we do not expect our observationally derived diurnal cycle correction would necessarily agree with one derived from a GCM. This is because our correction may include residual calibration biases that are correlated with the diurnal cycle and the GCM diurnal cycle may also have biases. Other observational studies have noted differences in the representation of phase and amplitude in the diurnal cycle over land compared to GCMs [e.g., Dai and Trenberth, 2004, Seidel et al., 2005, MacKenzie et al., 2012]. For example, Dai and Trenberth [2004] show that the peak in the tropical land surface air temperature diurnal cycle occurs later in CCSM2 than in observations. We similarly find that the peak in the TMT diurnal cycle occurs earlier than in CCSM3 (Fig. 2.2). Note that the peaks of the tropical mean land AMSU GCM diurnal cycle used in Mears et al. [2011] occur at approximately 12:30, 1:30, and 2:30 PM for MERRA, HadGEM, and CCSM3 (Dr. Carl Mears, personal communication,

2014). Our tropical mean AMSU land peak occurs at 1:19 PM, which is well within the range of model estimates. We also note that differences in the amplitude between the MSU and AMSU diurnal cycle are also enhanced in the UW-derived diurnal cycle correction compared to the GCM. We expect the AMSU diurnal amplitude to be larger over land, because the AMSU weighting function peaks closer to the surface than the MSU weighting function.

The diurnal cycle correction over the ocean is small (see Fig. 2.2) and the difference between tropical ocean trends with and without a diurnal cycle correction is negligible (0.002 K decade^{-1} from Table 2.3). Although much of our derived ocean diurnal cycle correction is significantly different from zero, the corrections, when applied to the drifting satellite data, are not statistically significant. This suggests that a significant bias is detected, but this bias has a negligible impact on our time series. The AMSU AM ocean diurnal cycle correction and the MSU PM diurnal cycle correction are larger in our technique compared to the GCM. These differences may be caused by residual calibration biases that are correlated with the diurnal cycle and/or GCM biases. Other studies [e.g., Dai and Trenberth, 2004, Seidel et al., 2005, MacKenzie et al., 2012], noted that the amplitude of GCM diurnal cycles over ocean is underestimated relative to observations. Over the ocean, the observationally-derived diurnal correction for MSU has increasing temperatures during the afternoon, but that for AMSU has relatively constant temperatures. It is not obvious which afternoon oceanic diurnal cycle correction is closer to the real diurnal cycle. Note that the peak in the AMSU tropical ocean diurnal cycle is approximately 3:30, 7:30, or 8:30 PM for HadGEM, MERRA, and CCSM3, respectively (Dr. Carl Mears, personal communication, 2014). The inconsistency in observationally-derived MSU and AMSU oceanic diurnal cycle corrections suggests that residual calibration biases are present. Regardless, these biases are effectively removed through our diurnal cycle correction, which has a negligible effect on our derived trends. Our results indicate that our oceanic diurnal cycle correction represents a combination of a small diurnal cycle and small residual calibration corrections.

The possibility that unresolved errors introduce a significant impact on our observationally-derived diurnal cycle correction is likely small since removing individual satellites from con-

sideration typically produces diurnal cycles within the uncertainty of our regression (Fig. 2.2). Some satellites, such as NOAA-14 or NOAA-15, do have greater influence over our derived diurnal cycle correction and also have unexplained trend differences in their times series [Mears and Wentz, 2009a, Mears et al., 2011], which could be aliased into our regression technique. Since instrument calibration drift is similar for the ascending and descending node, incorporating ascending minus descending TMT comparisons into our diurnal cycle correction estimate helps minimize the effect of instrument calibration drift biases. The insensitivity of the diurnal cycle estimate to our treatment of the warm target bias also indicates that there is not a large influence of instrument calibration drift on our analysis.

We further solve for the diurnal cycle correction as a function of latitude and surface type (land or ocean) within the tropics. This approach allows us to construct the spatial evolution of TMT and to determine the combined (land and ocean) tropical TMT time series, though we note that the land and ocean trends using this gridded method are the same as those calculated using tropical mean data. In this calculation, grid cells containing greater (less) than 50% land fraction are considered as land (ocean). We solve Eqs. 3 and 4 for the diurnal cycle correction coefficients for land and ocean in each latitude band, respectively. We then apply the diurnal correction and constant offsets equally to all of the respective land or ocean grid points in that zonal band. We tested to see if the zonal spatial scale or the annual mean diurnal correction had a large influence on our results by applying a tropical, annual mean GCM land and ocean correction to each land and ocean grid cell, respectively. This test forces all regions (e.g., convective versus non-convective) to share a common tropical mean diurnal cycle with no seasonal dependence, even though it has been demonstrated that the diurnal cycle varies between convective and non-convective regions and over the seasonal cycle [e.g., Tian et al., 2004, Yang and Slingo, 2001]. The tropical land trend difference versus the full GCM correction was 0.010 K decade^{-1} and less than 0.005 K decade^{-1} over ocean. These sensitivity experiments demonstrate that although the spatial pattern of the diurnal cycle evolution (e.g., across varied surface characteristics within a latitude band or over regions varied convection) may be important locally they do not have large effects on

the tropical mean TMT warming. The remainder of the analysis will be based on our gridded TMT dataset that utilizes zonal diurnal corrections. We will refer to data corrected with our observationally-derived diurnal correction as UW (University of Washington) and data processed with a GCM diurnal correction as UW_{GCM}.

2.5.3 The efficacy of observationally-derived TMT diurnal cycle corrections

Figure 2.3 presents the time series of differences between the ascending and descending nodes for satellites over tropical land without a diurnal correction, with a correction using a GCM diurnal cycle (UW_{GCM}), and with our observationally-derived diurnal cycle correction (UW). Comparing differences in the AM and PM nodes of satellites is an important check for removing diurnal drifts, because instrument calibration drift should be similar in both the ascending and descending node (reducing the effect of ϵ in Eqs. 3 and 4). Removing the diurnal drift with a GCM diurnal cycle is generally effective, but leaves small residual trends in the ascending minus descending difference time series. Our regression approach further reduces residuals and trends between the ascending and descending nodes.

Effectively removing diurnal drifts should also lead to near-zero differences between satellites for monthly mean brightness temperatures. Figure 2.4 presents the differences in tropical land anomalies between satellites for various pairs of ascending nodes for TMT data with no diurnal correction, data corrected with a GCM diurnal cycle (UW_{GCM}), and data corrected using our observationally-derived diurnal cycle correction (UW). We chose to display the ascending (PM) node because of larger drifts in this node, providing a powerful test of our technique. We attempted to choose pairs of satellites in which one satellite is relatively stable (i.e., not rapidly drifting east or west relative to the sun) while the comparison satellite has large drifts (see Fig. 2.1). The colors of the residuals represent the LECT of the rapidly drifting satellite, which generally cools (warms) relative to our comparison satellite as it drifts later (earlier) in the day if no diurnal corrections are applied. Figure 2.4 shows that in some cases the GCM seems to overcorrect the diurnal drift (e.g., NOAA-11 - NOAA-12), while in other cases some of the diurnal signature remains (e.g., NOAA-16 - AQUA). The GCM

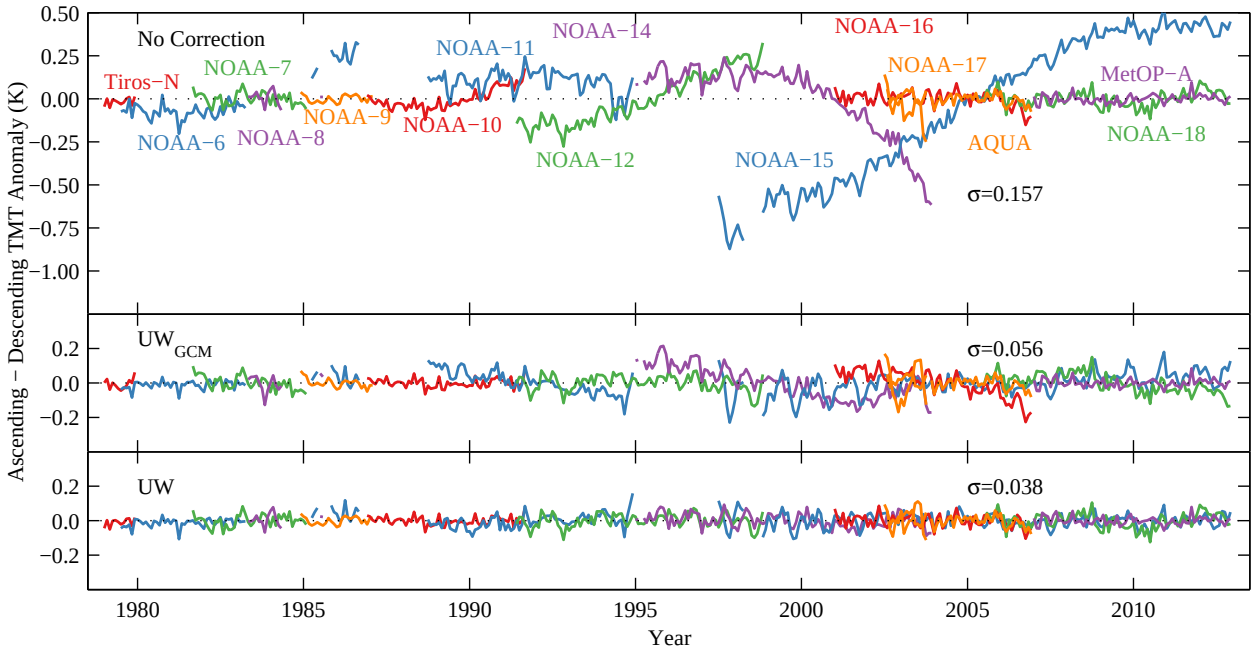

Figure 2.3: Time series of the ascending minus descending node for each satellite over tropical (20°N - 20°S) land using TMT (top) without diurnal corrections, (middle) with diurnal drift corrections following a diurnal cycle from a GCM, and (bottom) with corrections from our regression technique.

diurnal bias correction is largely effective at removing intersatellite residuals and trends, but our technique further reduces these biases. Figure 2.1 shows that AQUA has almost no LECT drift because of its onboard propulsion system. In Figure 2.4, the differences between NOAA-15 and NOAA-16 relative to AQUA can thus largely be attributed to the diurnal cycle. Each difference measurement demonstrated in Figs. 2.3 and 2.4 helps to constrain our fit for the diurnal cycle correction even though there is no stable, non-drifting reference (e.g., AQUA or MetOp-A) for the bulk of the time series. Using all overlaps helps to ensure that unresolved or unknown instrument problems do not give us spurious results.

In order to characterize the overall effectiveness of our method, we show the intersatellite residual standard deviation in Table 2.2. Intersatellite biases over land are effectively reduced using a GCM diurnal correction (UW_{GCM}) and our diurnal correction (UW), as compared to the case in which no diurnal corrections are applied. Our correction represents a small

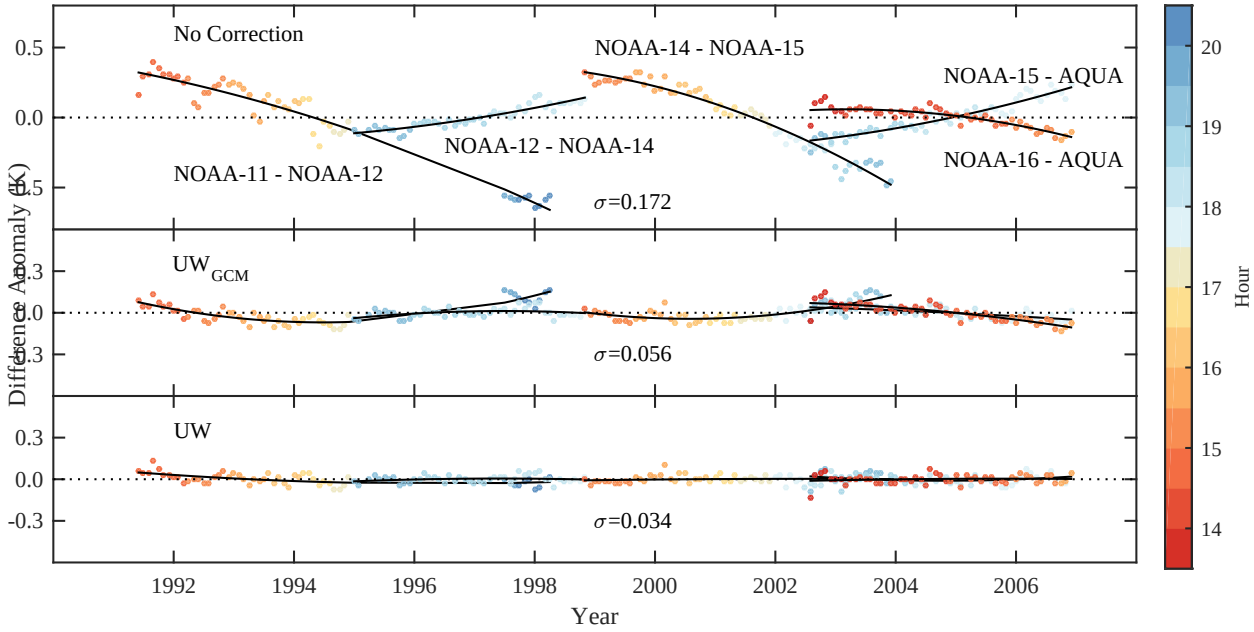

Figure 2.4: Time series of differences between pairs of satellites for the ascending node over tropical (20°N - 20°S) land using TMT (top) without diurnal drift corrections, (middle) with diurnal corrections following a diurnal cycle from a GCM, and (bottom) with corrections from our regression technique. We chose these pairs of satellites and the ascending node over land to maximize diurnal drift biases. For each pair (i.e., satellite 1 2 satellite 2), satellite 1 is rapidly drifting, while satellite 2 has relatively little diurnal drift. The color of each point represents the LECT of the drifting satellite 1 for reference. Quadratic fits to the data are shown to aid in visualization.

but robust improvement compared to the GCM correction. Figures 2.3 and 2.4 and Table 2.2 illustrate that our diurnal correction appears to be successful in removing intersatellite biases due to diurnal drift.

2.6 TMT time series and trends and comparisons with UAH, RSS, and NOAA

We remove the diurnal drift biases, warm target effect, and constant offsets from each satellite time series and average over all of the satellites and both the ascending and descending nodes to form a homogenized MSU/AMSU TMT time series. Figure 2.5 presents the UW TMT time series for the tropical mean over land, ocean, and both the land and ocean. As with

	Intersatellite residual σ (K)		
	Both	Ocean	Land
No correction	0.020	0.019	0.052
UW_{GCM}	0.021	0.020	0.035
UW	0.017	0.018	0.027

Table 2.2: Average monthly intersatellite residual standard deviation (s) for TMT without diurnal corrections (no correction), UW, and UW_{GCM}. We show results for the tropical (20°N - 20°S) land, ocean, and land-ocean averages. The value displayed is a weighted average of the standard deviation of all co-orbiting satellites with at least 6 months of overlap.

other MSU/AMSU datasets there is strong coherence between the land and ocean time series in the tropics because the tropics do not maintain strong temperature gradients. The El Niño Southern Oscillation is an evident feature, especially the large El Niño event during 1997 - 1998. Our TMT time series is quite similar to that from other research groups (not shown).

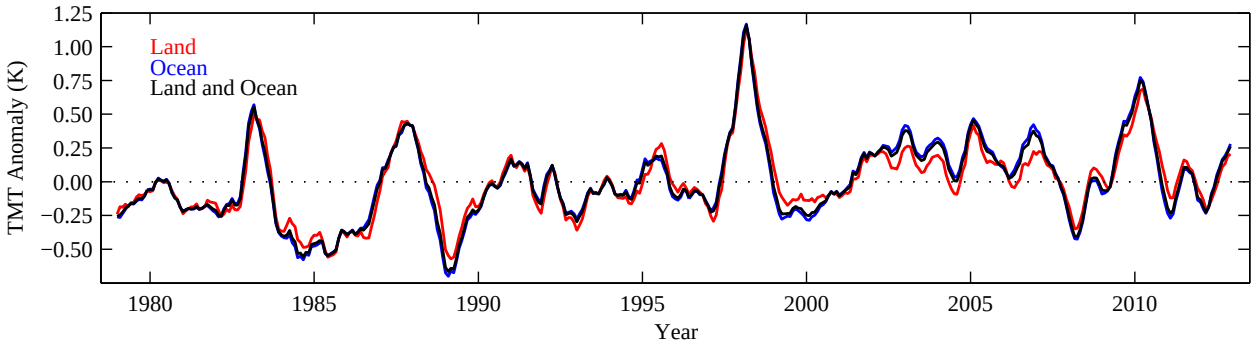

Figure 2.5: Time series of UW TMT for the tropics (20°N - 20°S) for land (red), ocean (blue), and combined land-ocean (black) over 1979-2012. The time series is smoothed using a 5-month moving average.

Figure 2.6 presents the probability distribution functions of our tropical TMT trends over land and ocean along with trend values from other groups. This calculation utilizes

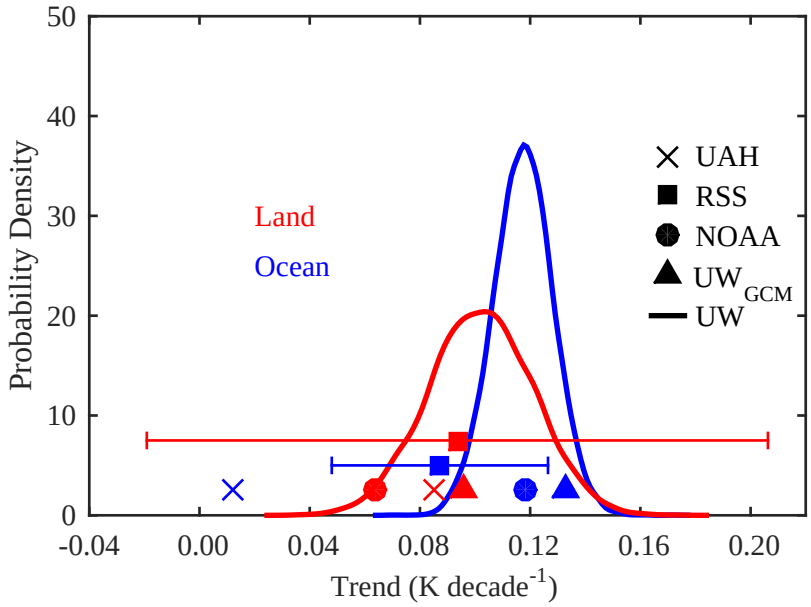

Figure 2.6: Probability distribution function of tropical (20°N - 20°S) UW TMT land (red) and ocean (blue) trends over 1979-2012 from this study. We also show the trends for UW_{GCM}, NOAA, RSS, and UAH for comparison. We include the uncertainty values for RSS trend calculation using data from Mears et al. [2011]. The UW distribution represents the results when we add noise in our regression for the warm target factors and diurnal cycle correction and then re-merge all of the satellites together 10,000 times.

our diurnal correction performed over the tropical land and ocean with the ensemble of trend values coming from our Monte-Carlo simulation described in Section 2.5.2. We find a mean tropical land trend of 0.103 ± 0.037 K decade^{-1} and a mean tropical ocean trend of 0.118 ± 0.021 K decade^{-1} (95% confidence interval). Note that this uncertainty is related to the uncertainty in our diurnal cycle and warm target corrections and does not include uncertainties related to sampling or structural uncertainties. Mears et al. [2011] undertook a comprehensive uncertainty analysis using a Monte-Carlo estimation technique; the resulting RSS trend uncertainty is shown in Figure 2.6. Even though the differences between groups can be large, these differences are not necessarily significant when the full internal uncertainty is considered.

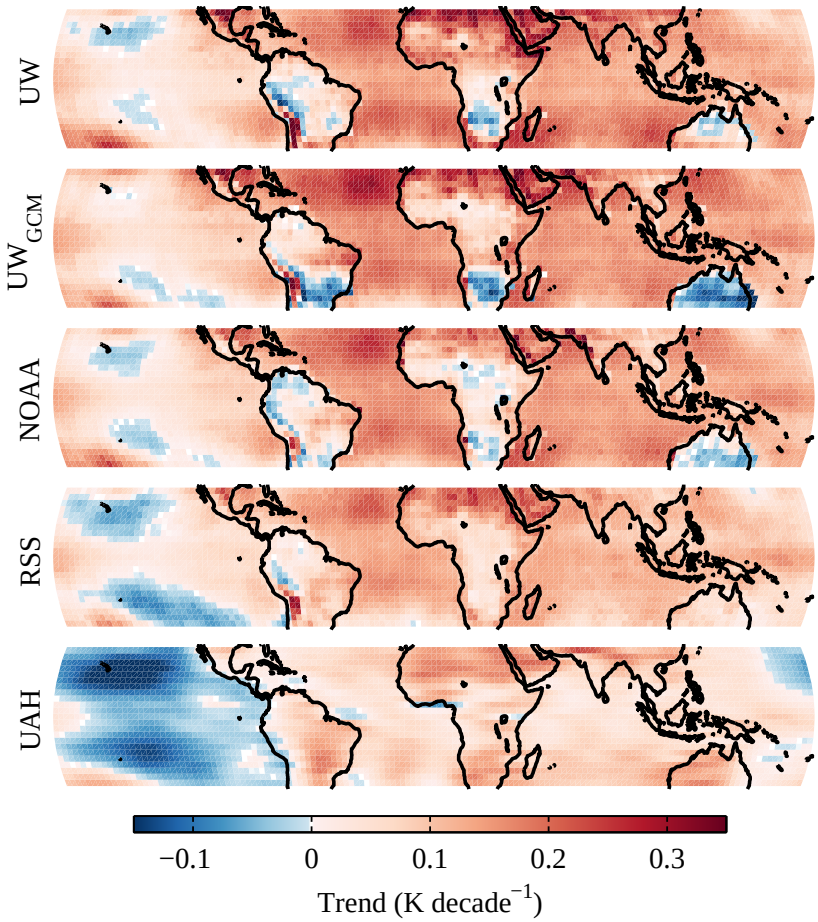

Figure 2.7: Tropical (30°N - 30°S) spatial pattern of TMT trends (K decade^{-1}) from various MSU/AMSU datasets over 1979-2012. We merged the MSU/AMSU observations together using both our regression technique based on observations (UW) and a GCM diurnal drift correction (UW$_{GCM}$) with other processing choices the same.

Figure 2.7 presents maps of tropical TMT trends over 1979 - 2012 based on our regression technique (UW), our data corrected with a GCM diurnal cycle (UW$_{GCM}$), and trends from the UAH, RSS, and NOAA datasets. UW, RSS, and NOAA have a similar warming pattern, though RSS also has a noticeably larger region of cooling in the subtropical Pacific Ocean. The most striking difference is that UAH has large negative cooling throughout most of the Pacific Ocean and has less warming in the other ocean basins. UAH and RSS also lack the

cooling areas over the African and Australian continents that are present in both the UW and NOAA datasets.

Table 2.3 presents the tropical mean TMT trends over 1979 through 2012. In the table, we include uncertainty estimates for RSS [Mears et al., 2011] and for UW, which are based on the uncertainty in merging and bias removal procedures. We can also calculate the statistical uncertainty in the trend for each group. The 95% confidence interval in the least squares linear trend estimates, accounting for autocorrelation, is roughly the same for each dataset: approximately $\pm$ 0.10 K decade^{-1} over land, $\pm$ 0.12 K decade^{-1} over ocean, and $\pm$ 0.12 K decade^{-1} for land and ocean. For the UW time series, the warm target correction reduces trends across land and ocean by 0.023 K decade^{-1}. The diurnal correction is a small positive correction over ocean (0.002 K decade^{-1}), but a very large and positive correction over land (0.159 K decade^{-1}). In general, our trends corrected with a GCM and trends corrected with our observationally derived diurnal cycle correction are similar to trends from NOAA and RSS, though RSS has less warming over ocean and NOAA has less warming over land. As demonstrated in Table 2.3 and Figs. 2.6 and 2.7, the UAH ocean trend is notably lower than trends from all other datasets. Differences among various TMT datasets over the tropics will be interpreted in detail in Section 6.

One drawback of TMT trends is that they include some influence from the stratosphere, which has a large negative temperature trend. It has been demonstrated that this influence can be removed using TLS stratospheric data from MSU channel 4 and AMSU channel 9 to provide a measure of the full tropospheric temperature (T24) trend [e.g., Fu et al., 2004, Fu and Johanson, 2004, 2005]. In Table 2.4, we provide tropical T24 trends, surface temperature trends from HadCRUT4, and the ratio of the T24 trend to the surface trend. Our amplification factor over the tropics is consistent with tropical amplification implied by models, which is approximately 1.5 - 1.6 [Santer et al., 2005, Fu et al., 2011]. Our amplification factor over land is reduced due to enhanced land surface warming relative to sea surface warming [Sutton et al., 2007]. All of the MSU/AMSU datasets demonstrate tropical amplification, except UAH.

	TMT trend (K decade^{-1})		
	Both	Ocean	Land
No diurnal correction	0.075	0.116	-0.056
UW	0.115 ± 0.024	0.118 ± 0.020	0.104 ± 0.035
UW_{GCM}	0.124	0.132	0.096
NOAA	0.105	0.118	0.064
RSS	0.089 ± 0.051	0.087 ± 0.039	0.094 ± 0.113
UAH	0.029	0.012	0.085

Table 2.3: Trends (K decade^{-1}) over 19792012 for TMT over the tropics (20°N - 20°S) for TMT without diurnal corrections (no diurnal correction), UW, and UW_{GCM}. The trends from NOAA, RSS, and UAH are listed for reference. The 95% confidence intervals on UW trends are from our Monte Carlo simulation. The 95% confidence interval for RSS are based on data updated since Mears et al. [2011].

2.7 Interpreting TMT time series differences between MSU/AMSU datasets

Figure 2.8 shows the TMT brightness temperatures from each group minus the UW time series over the tropics. It contains several notable features. Relative to the UW dataset over land, RSS, NOAA, and UW_{GCM} slowly cool from 1979 to 1995. RSS, NOAA, and UW_{GCM} also show warming after 2000 over land, followed by cooling after 2003 relative to UW TMT. Over ocean, RSS cools relative to UW near 1985 - 1987 and near 2004. NOAA warms relative to UW near 1998 and cools after 2003. We note that 1985 - 1987 differences are likely a result of the NOAA-9 target factor [Po-Chedley and Fu, 2012b]; RSS uses a slightly larger target factor than NOAA and UW. Overall, NOAA behaves similarly to UW_{GCM} over ocean and has similar features over land, especially after 1997. UAH shows a large discontinuity during 1985 to 1987 in both land and ocean, a cooling near 1992 (1995) over land (ocean), a gradual cooling from about 1997 to 2003 over both ocean and land, and a large warming

Group	Both	Ocean	Land
UW	0.160 (1.41)	0.163 (1.62)	0.150 (0.86)
UW_{GCM}	0.170 (1.50)	0.179 (1.78)	0.141 (0.81)
NOAA	0.149 (1.32)	0.163 (1.62)	0.106 (0.61)
RSS	0.125 (1.10)	0.123 (1.22)	0.133 (0.76)
UAH	0.064 (0.56)	0.044 (0.44)	0.129 (0.73)
HadCRUT4	0.114	0.101	0.175

Table 2.4: T24 trends (K decade^{-1}) over 19792012 in the tropics (20°N - 20°S) over land, ocean, and the entire tropical region, as derived from various MSU/AMSU datasets. The values in parentheses are the amplification ratio, which is defined here as the T24 trend divided by the HadCRUT4 surface trend. We compute T24 using $T24 = 1.1TMT - 0.1TLS$. The UW and UW_{GCM} T24 trends are calculated using TMT from the present study and TLS from NOAA STAR v3.0 data.

trend over land near 2005 relative to UW.

Despite the differences in both phases and magnitudes of the diurnal cycles between GCM and observations (Fig. 2.2), the differences in derived TMT time series between UW_{GCM} and UW (Fig. 2.8) are small. The trend differences are less than 0.02 K decade^{-1} (Tables 2.3 and 2.4). It is indicated that the diurnal cycle from a GCM can be used to effectively remove the satellite drift biases in TMT.

In Fig. 2.9, we present results that aid in further understanding some of the underlying causes for difference between the tropical TMT time series. Figure 2.9(A) shows UAH minus UW, RSS, and NOAA for the tropical ocean. We utilize the tropical ocean in this case, because this reduces the effect of diurnal drift. In each of these cases, UAH has a large discontinuity during 1985 - 1987. In Fig. 2.9(A) we also use a time series in which we force the NOAA-9 target factor to be the same as the UAH value, $\alpha 9 = 0.0986$ [Po-Chedley and Fu, 2012b], and then re-solve for the rest of our target factors ($UW_{\alpha(9)=UAH}$) to see if we reproduce the UAH discontinuity. Our original target factor for NOAA-9 using data in which

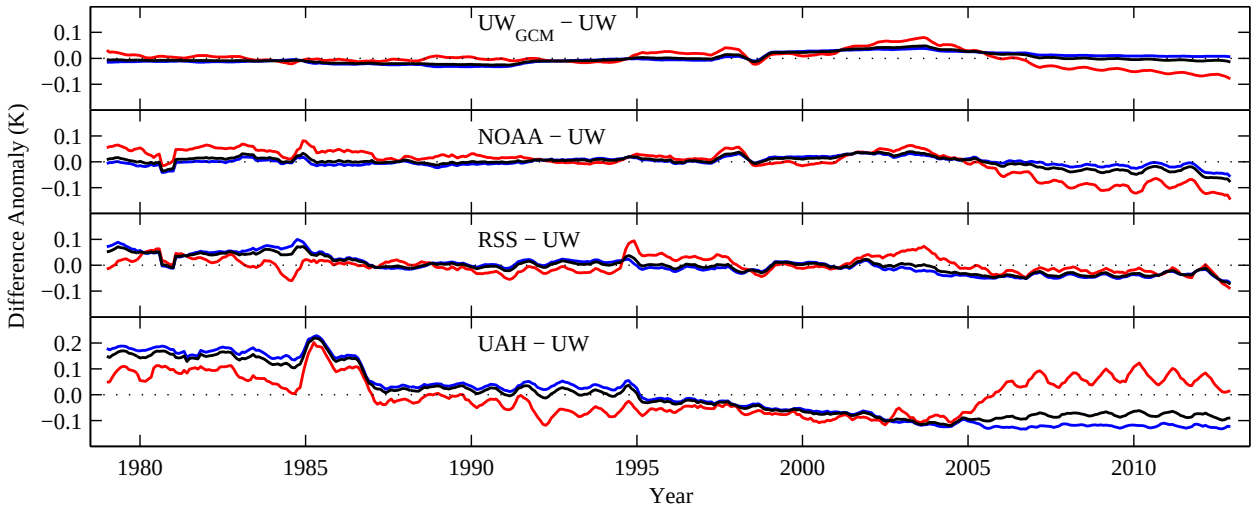

Figure 2.8: Tropical (20°N - 20°S) land (red), ocean (blue), and combined land-ocean (black) difference time series for NOAA, RSS, and UAH minus UW. A time series of UW$_{GCM}$ is also shown to understand the differences caused by diurnal cycle corrections (all other processing decisions are held constant between UW and UW$_{GCM}$). The time series are smoothed using a 5-month moving average.

instrument non-linearity biases are removed via IMICA is -0.0046. The small magnitude of our target factor demonstrates that the NOAA IMICA calibration is effective at removing instrument biases. By solving for the diurnal cycle correction with these perturbed target factors, we reproduce the UAH discontinuity during 1985 - 1987 (see the red line in Fig. 2.9(A)). Relative to UW, UWa(9)=UAH reduces the tropical ocean TMT trend by 0.039 K decade^{-1}, which explains about one-third of the trend difference between UAH and UW over the tropical ocean (Table 2.3). Importantly, UAH has a similar discontinuity relative to RSS, NOAA, and UW, and the NOAA and UW effectively rely on the independent IMICA instrument calibration. Figure 2.9(A) indicates that UAH has a problem with its bias removal during the 1985 - 1987 period. Our recent work showed that the UAH NOAA-9 target factor has a positive bias of 0.05 - 0.08 based on radiosondes and in comparisons between UAH NOAA-9 and NOAA-6 data [Po-Chedley and Fu, 2012b, 2013, Christy and

Spencer, 2013]. This bias leads to a contamination of NOAA-9 warm target temperatures in the UAH TMT dataset, causing a global artificial cooling trend of about 0.035 K decade^{-1} over 1979 - 2009 [Po-Chedley and Fu, 2013]. UAH uses land and ocean data to solve for its target factors. One potential approach to understand this bias is to use UAH global ocean data with and without a diurnal cycle bias correction to solve for the warm target factors. Since the target factors are meant to minimize errors in instrument calibration, the target factors should be invariant for land and ocean. If the target factors change using UAH oceanic data, it would suggest that a UAH diurnal cycle bias correction is influencing the warm target calibration. The ocean diurnal correction should also have a small effect on the target factors since the diurnal cycle over the ocean is small. UAH also utilizes seasonal smoothing and a subset of satellites to determine its NOAA-9 target factor, which may also contribute to this difference [Po-Chedley and Fu, 2013]. The UAH data, as is, suggests that the warming of the tropical troposphere is decoupled from the tropical oceanic warming. On the other hand, three independent MSU/AMSU analyses find a much lower target factor value for NOAA-9, which is also supported by analysis based on independent observations [Po-Chedley and Fu, 2012b, 2013].

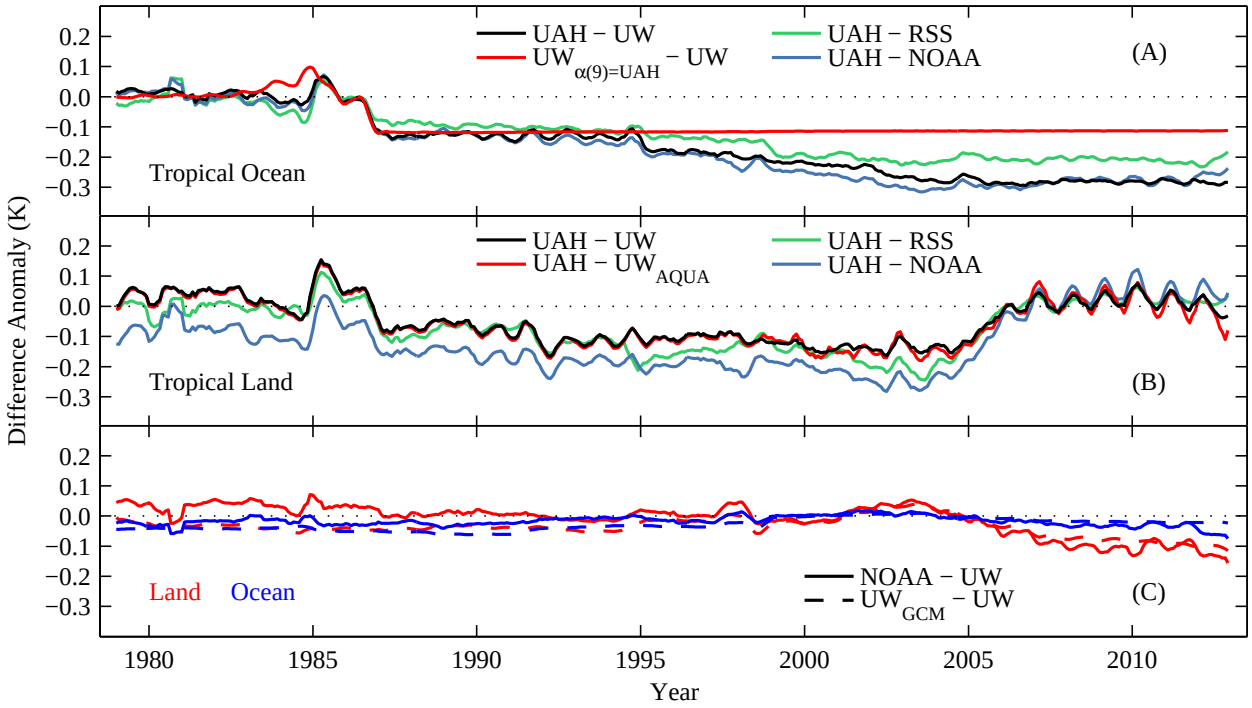

Figure 2.9: Tropical (20°N - 20°S) ocean TMT time series for UAH minus UW, RSS, and NOAA. Ocean TMT data are used to minimize the influence of the diurnal cycle. UW data are also compared to data in which we force the NOAA-9 target factor to match the UAH value of 0.0986 [UWa(9)5UAH]. Each time series is smoothed using a 5-month moving average and has anomalies computed relative to the 1985-87 time period for comparison purposes. (b) Tropical (20°N - 20°S) land time series for UAH TMT minus UW, RSS, and NOAA TMT. The land time series is used here to understand the diurnal cycle bias corrections after 2005. Smoothed time series anomalies are displayed relative to the 2005-11 time period for comparison purposes. We also use UW TMT time series in which we remove NOAA-15, NOAA-16, NOAA-17, and NOAA-18 from our time series (UW_{AQUA}), which forces our data to rely on the Aqua satellite in the mid-2000s (since Aqua has no diurnal drift, UW_{AQUA} cannot have a diurnal drift bias). (c) Comparison of NOAA TMT 2 UW TMT for tropical (20°N - 20°S) land (red solid line) and ocean (blue solid line). We also display UW_{GCM} 2 UW for land (red dashed line) and ocean (blue dashed line) to understand the differences causes by diurnal drift bias corrections. Smoothed time series anomalies are displayed relative to the 2000-07 time period.

The UAH tropical TMT land trend is closer to that of RSS, NOAA, and UW largely because of strong warming after 2005. Christy et al. [2010] suggested that after 2001 model based diurnal corrections for TLT results in artificial warming for NOAA-14 and even greater artificial cooling for NOAA-15, which lead to greater consistency between UAH and RSS by 2008. Figure 2.9(B) shows the UAH tropical land time series relative to UW, RSS, and NOAA. In this case we utilize the tropical land time series to emphasize differences related to the diurnal drift correction. UAH has strong warming near 2005 relative to all of the comparison datasets. Note that UW applied an observationally-based diurnal drift correction while RSS and NOAA used a GCM-derived correction. To further test whether this difference is caused by diurnal drift corrections used in UW, RSS, and NOAA, we remove NOAA-15, NOAA-16, NOAA-17, and NOAA-18 to form a new time series (UW_{AQUA}), which means that TMT brightness temperatures largely rely on AQUA from 2002 - 2006 and MetOP-A afterwards. Since AQUA and MetOP-A do not have diurnal drift, there should be no signature of diurnal drift in UW_{AQUA}. Figure 2.9(B) shows that UAH - UW_{AQUA} is very similar to UAH - UW and UAH still has similar strong land warming after 2005 relative to UW_{AQUA}. This suggests that the UAH TMT dataset has spurious warming related to its diurnal drift correction. In TMT v5.6, UAH uses satellites that have relatively little diurnal drift and does not apply corrections for satellites carrying AMSU, which may leave spurious warming during NOAA-15's lifetime (Dr. John Christy, personal communication, 2013). Comparisons between NOAA-15 and NOAA-16 with AQUA may help resolve this issue.

Figure 2.8 and 2.9(A) suggests that much of the UAH tropical trend difference relative to other datasets is due to the NOAA-9 target factor, with more gradual changes during 1995 - 2005, which also contribute to the discrepancy. A spurious UAH tropical land trend after 2005 relative to other TMT datasets and our analysis with drifting satellites removed suggest that the UAH treatment of diurnal drift has biases. Unlike the RSS and NOAA diurnal corrections, the UAH correction is not publicly available for comparison. UAH constructs its TMT diurnal drift correction using 13 months of data from 3 co-orbiting AMSU satellites:

NOAA-15, NOAA-16, and NOAA-17 (Dr. John Christy, personal communication, 2013). Differences between our approach and UAH likely result because UAH does not apply diurnal corrections to AMSU measurements (AMSU was carried on satellites starting with NOAA-15). It is also possible that diurnal corrections calculated with AMSU may have biases when applied to MSU. A further complication is that the three co-orbiting satellites need to be absolutely calibrated, which is not possible with AMSU since MSU/AMSU radiometers are not SI calibrated. Furthermore, the correction derived from six unique points may be sensitive to assumptions regarding the shape of the diurnal cycle.

Figure 2.9(C) shows NOAA - UW for both tropical land and ocean. We also plot UW_{GCM} - UW, which isolates the differences due to the GCM diurnal correction versus our observationally-based approach since other processing steps are the same. This UW-UW_{GCM} comparison shows that slightly enhanced warming over the tropical ocean may be a result of the GCM derived diurnal drift correction, though the trend difference is very small (0.014 K decade^{-1}). The upward and then downward trends in NOAA and RSS tropical land between 2000 and 2007 (Fig. 2.8 and 2.9(C)), similar to those in UW_{GCM}, are caused by the difference between the GCM and observationally-based diurnal corrections. This behavior may be due to the residual trends in NOAA-14, NOAA-15, and NOAA-16 (Figs. 2.3 and 2.4), which was also noted in Christy et al. [2010] for TLT. Since our diurnal drift correction is observationally-based and also improves error characteristics in the tropical TMT time series relative to a GCM, these features in the NOAA (and RSS) time series may be artificial owing to the use of the GCM-simulated diurnal cycle. This statement is further enforced by examining the difference between NOAA and UW_{AQUA} (not shown). It's also noted that the small jump in 1997/1998 in NOAA - UW appears to be related to the re-emergence of NOAA-11 in the time series (see Figs. 2.4 and 2.8). UW warms relative to NOAA over land prior to 1998. We utilize different time periods for the satellites compared to NOAA and find that there is some sensitivity of the MSU/AMSU merged trend to the exact satellite data used. Other causes for this difference may include processing differences such as small differences in the application of our GCM diurnal correction or warm target correction.

2.8 Discussion and conclusions

This study demonstrates the success of an observationally-based approach to removing diurnal drift biases from the MSU/AMSU TMT record. The approach utilizes information from intersatellite differences and differences in ascending and descending nodes of individual satellites. As satellites drift through the diurnal cycle we can compare observations at various LECTs and construct a common diurnal cycle correction for MSU and AMSU that explains differences in the TMT measurement for each satellite and node. This approach, compared to a diurnal drift bias correction derived from a climate model, has improved error characteristics, though tropical trend values utilizing the observationally-based diurnal correction are very similar to trend values utilizing a GCM correction, with differences smaller than 0.02 K decade^{-1}. By deriving zonally averaged diurnal corrections over land and ocean, our approach may introduce some biases in the spatial pattern of trends over land because of different land surface characteristics in a given latitudinal band. The spatial pattern of our TMT trends, however, compares well with independent TMT datasets.

In this work we combined TMT mid-tropospheric data with TLS lower-stratospheric data to derive the T24 full tropospheric temperature, which effectively removes the influence of the stratosphere from TMT. The ratio of tropical full tropospheric T24 trends to the HadCRUT4 surface temperature trends over ocean is 0.4, 1.2, 1.6, and 1.6 from UAH, RSS, NOAA, and UW, respectively. The ratios from RSS, NOAA, and UW demonstrate tropical amplification and are in general agreement with tropical amplification from climate models, indicating that there is no significant discrepancy between observations and models for lapse rate changes between the surface and the full troposphere.

This work represents an independent analysis of the MSU/AMSU TMT evolution. We start the analysis using the newly released NOAA STAR L1C data that includes the state-of-the-art corrections for instrument calibration biases. Our subsequent analysis using an observationally-based diurnal drift correction shows that our results are generally consistent with RSS and NOAA in the tropics, even though small discrepancies exist. The focus of this

work is on the tropics, given the historic debates regarding tropical tropospheric warming. Large differences between UAH and comparison datasets in the tropics are largely a result of differences in the NOAA-9 target factor and differences in diurnal drift corrections. Although this is generally referred to as a structural uncertainty, careful comparisons between co-orbiting satellites and between the ascending and descending node on individual satellites may help resolve these discrepancies. Work by Po-Chedley and Fu [2012b, 2013], through multiple lines of evidence, suggests that the UAH NOAA-9 target factor is too large. This bias explains roughly one-third of the trend difference in the tropical ocean between UAH and UW. Evidence of diurnal drift bias in the UAH TMT dataset after 2005 was presented in this work. Christy et al. [2011] has suggested the large differences between UAH and RSS (NOAA) in the tropical TMT time series are due to RSS diurnal drift biases that are prominent around 1992, though Mears et al. [2012] demonstrate that evidence presented for these claims is dependent on the datasets, the methodology, and ignores the large uncertainty in observations. Our independent analysis based on an observationally-derived diurnal cycle correction also shows differences between UAH and UW in the 1990s, especially a gradual cooling relative to UW during 1995 - 2005. These differences could be due to a number of effects including data treatment, warm target calibration, and diurnal drift. In the future, it would be useful if each dataset provider made individual satellite node time series and bias corrections available for inter-comparison since this work shows that these comparisons can be helpful in advancing understanding of biases in the MSU/AMSU TMT dataset.

2.9 Acknowledgments

We would like to thank Dr. Cheng-Zhi Zou for the useful discussions on this topic. We would also like to thank two anonymous reviewers for their helpful comments. This work was supported by the National Science Foundation Graduate Research Fellowship (DGE-0718124) and by NASA grant NNX13AN49G.

Chapter 3

DISCREPANCIES IN TROPICAL UPPER TROPOSPHERIC WARMING BETWEEN ATMOSPHERIC CIRCULATION MODELS AND SATELLITES

Po-Chedley, Stephen and Qiang Fu, 2012, "Discrepancies in tropical upper tropospheric warming between atmospheric circulation models and satellites," Environmental Research Letters, doi: http://iopscience.iop.org/article/10.1088/1748-9326/7/4/044018.

3.1 Abstract

Recent studies have compared CMIP3 coupled atmosphere-ocean general circulation model (GCM) simulations with satellite and radiosonde observations of warming in the tropical upper troposphere relative to the lower-middle troposphere. These studies showed that models tended to overestimate increases in static stability in the tropical upper troposphere. We revisit this issue using atmospheric GCMs with prescribed historical SSTs and historical coupled GCMs that participated in CMIP5. It is demonstrated that even with historical SSTs as a boundary condition, most atmospheric models exhibit excessive tropical upper tropospheric warming relative to the lower-middle troposphere as compared with satellite-borne microwave sounding unit measurements. It is also shown that the results from CMIP5

coupled atmosphere-ocean GCMs are similar to findings from CMIP3 coupled GCMs. The apparent model-observational difference for changes in static stability in the tropical upper troposphere represents an important problem, but it is not clear whether the difference is a result of common biases in GCMs, biases in observational datasets, or both.

Keywords: climate, global warming, tropical amplification, tropical troposphere, atmospheric temperature trends, microwave sounding unit, global circulation models, remote sensing

3.2 Introduction

A robust feature of GCM simulations of the 21st century climate is enhanced tropical upper tropospheric warming with temperature trends increasing as a function of height until about 200 hPa [Solomon et al., 2007]. This characteristic accounts for much of the lapse rate and associated water vapor feedbacks in GCMs, which are important factors for climate sensitivity [e.g., Soden and Held, 2006]. Amplification of tropical tropospheric temperature trends relative to the surface has been examined in a number of studies and it has been demonstrated that no serious inconsistency exists between GCMs and observations [Fu et al., 2004, Fu and Johanson, 2005, Douglass et al., 2008, Thorne et al., 2007, Santer et al., 2008]. Recent studies of temperature amplification in the tropical upper troposphere relative to the lower-middle troposphere found that coupled atmosphere-ocean models from Phase 3 of the Community Model Intercomparison Project (CMIP3) tended to exaggerate this amplification compared to satellite microwave sounding unit (MSU) [Fu et al., 2011] and radiosonde [Seidel et al., 2012] observations.

A confounding factor in comparing coupled atmosphere-ocean GCMs with observations is that coupled GCM's internally generated natural variability is not the same as the observational record. Furthermore, historical simulations from coupled GCMs have larger tropical tropospheric temperature trends compared to observations since 1979 [Fu et al., 2011]. Fu et al. [2011] demonstrated that some of the differences in the change of static stability in the tropical upper troposphere between models and observations are simply a result of an

overestimation of tropical warming in CMIP3 coupled GCM simulations. This study utilizes CMIP5 model simulations [Taylor et al., 2012] from the Atmospheric Model Intercomparison Project (AMIP). AMIP model simulations use observational records of sea ice and sea surface temperature (SST) evolution [e.g., Hurrell et al., 2008] as boundary conditions for atmospheric GCMs. Thus the AMIP model's SSTs have the same variability and trends as observations. Previous studies have demonstrated that AMIP style runs can closely reproduce the observed tropical tropospheric temperature variations [Hurrell and Trenberth, 1997]. The use of AMIP models allows us to closely examine the simulated changes in tropical tropospheric static stability in GCMs using the observed SST evolution.

Another advantage of the AMIP simulations is that the pattern of SST changes in the models is constrained to the observations, which may affect tropical tropospheric amplification. Deep convection typically occurs over warm tropical SST regions [e.g., Wallace, 1992, Johnson and Xie, 2010], which acts as a heating source for the entire tropical free troposphere. This large-scale heating occurs because any anomalous heating is distributed in accord with the weak temperature gradient approximation [e.g., Bretherton and Sobel, 2003]. The tropical tropospheric temperature profile should follow a moist adiabat that is determined by warm SST regions [e.g., Sobel et al., 2002]. As a result, the specific pattern of SST warming could affect the structure of the temperature change in the tropical troposphere. Using AMIP models in this study will ensure that the effect of the pattern of SST changes on disagreements between models and observations is minimized.

In this letter observed changes in the temperature structure of the tropical troposphere from MSU observations are used to examine the changes in CMIP5 models that produced simulations for the AMIP and coupled atmosphere-ocean historical experiments. The main focus of this research is to determine if modeled and observational discrepancies seen in Fu et al. [2011] and Seidel et al. [2012] can be reconciled in atmospheric models using prescribed historical SSTs. We show that even by constraining atmospheric models with prescribed SSTs, model upper tropospheric warming relative to the lower-middle troposphere is still significantly larger than observations for most models. We also show that the results for

CMIP5 coupled simulations are similar to CMIP3 results [Fu et al., 2011].

3.3 MSU Datasets and Model Simulations

The MSU and its successor, the advanced MSU, have provided global temperature measurements of deep atmospheric layers since late 1978 over a number of satellites that have been combined into a homogenized climate record. In this work, we will utilize two synthetic channels that have been created to remove the influence of the stratosphere. The lower-middle tropospheric channel, TLT, utilizes a linear combination of view angles from the mid-tropospheric channel, T2, with a weighting function that peaks near 600 hPa [Spencer and Christy, 1992]. Another synthetic satellite channel uses observations from the lower stratospheric channel, T4, to remove the influence of the stratosphere on T2 [Fu et al., 2004, Johanson and Fu, 2006]. This channel, referred to as T24, has a weighting function that peaks near 300 hPa. We will combine MSU T2 and T4 observations to produce a T24 time series using a tropical weighting of T24 = 1.1T2 - 0.1T4 [Johanson and Fu, 2006]. These two channels will then be used as metrics to determine the relative change between the upper troposphere (represented by T24) and the lower-middle troposphere (represented by TLT) following Fu et al. [2011]. Specifically we will use the T24 - TLT trend as a measure of the static stability change in tropical upper troposphere and the ratio of T24 to TLT trends as a measure of tropical upper tropospheric temperature amplification.

Our T24 and TLT MSU observations come from the University of Alabama at Huntsville (UAH) and Remote Sensing Systems (RSS). The National Oceanographic and Atmospheric Administration (NOAA) also provides observations for T2 and T4, which allow us to create a T24 time series, but NOAA does not currently produce a TLT product. We will use RSS v3.3 [Mears and Wentz, 2009a,b], UAH v5.4 [Christy et al., 2003], and NOAA STAR v2.0 [Zou and Wang, 2011]. All of the teams that produce MSU products must account for a number of non-climatic biases such as satellite orbital decay, contamination related to the satellite warm target calibration, and drift in the local sampling time of the satellite [e.g Fu and Johanson, 2005, Mears and Wentz, 2005, Karl et al., 2006, Wentz and Schabel, 1998,

Christy et al., 2000, Po-Chedley and Fu, 2012b]. As a result of the different procedures for bias removal, the various teams derive different long-term temperature trends for TLT and T24, but the relative warming in TLT compared to T24 is consistent for both UAH and RSS likely because the structural uncertainties are similar for TLT and T24 [e.g., Po-Chedley and Fu, 2012b].

We will compare GCMs in the CMIP5 archive to MSU observations. Specifically, we will utilize models that contributed historical simulations (44 models) and models that contributed AMIP simulations (19 models). Taylor et al. [2012] provides more details on the CMIP5 experimental design. In both experiments, the atmospheric composition evolves in the model to reflect natural and anthropogenic forcings, but the atmospheric changes are not the same across all models. For example, not all models include emissions from volcanic eruptions. Hurrell and Trenberth [1997] noted that even when volcanic forcings are not included in atmospheric models, most of their influence is captured by the SST forcing. The historical simulations utilize coupled atmosphere-ocean GCMs and the AMIP simulations utilize atmospheric GCMs with prescribed SSTs (land surface temperature is not constrained). In order to compare the model simulations to satellite observations, we apply a static weighting function (from RSS) to the model's tropical mean temperature profile to produce an MSU equivalent brightness temperature.

3.4 Temperature Trend Difference Between the Tropical Upper and Lower-middle Troposphere

Figure 3.1 shows the time evolution of TLT from both coupled and AMIP model simulations and MSU satellite observations. The model spread is dramatically reduced in the AMIP models owing to the prescribed observational SSTs and there is good agreement between the multi-model mean and observed TLT time series (r=0.89 for 1979 - 2008). There is noticeable disagreement between AMIP models and MSU observations in 1979 and 1980. This disagreement has been discussed in the past and it is unclear if the disagreement is an artifact of the SST datasets used by the AMIP models, the MSU datasets, a problem

with the models, or some combination of these factors Hurrell and Trenberth [1997, 1998], Christy et al. [1997, 1998]. We will ignore this time period in this analysis, but we note that including these years increases the disagreement between modeled and observed measurements of tropical upper tropospheric amplification. The coupled historical models are listed in Table 3.1 and the AMIP models are listed in Table 3.2. In Tables 3.1 and 3.2, we also include the model ensemble mean trend values for TLT, T24, and T24 - TLT and the ratio of T24 to TLT trends. The AMIP models have trend values that are in much better agreement with the observations, but the amplification ratio is larger for both coupled and AMIP GCMs compared to the MSU observations. There is only one instance in which a model ensemble-mean has a lower amplification ratio than the MSU observations. As a whole, the trend amplification in GCMs is 1.2, which is consistent with warming in a moist adiabatic environment; this represents an increase in static stability in the tropical upper troposphere.

Table 3.1: Trend values for TLT and T24 in K decade^{-1} from the MSU teams and the ensemble mean of each coupled climate model (CMIP5 historical simulations). The trends are the simple mean of the TLT, T24, and T24-TLT trends of all of the ensemble members for each model. The ratio is the average T24 trend divided by the average TLT trend. The trends are for the tropics (20°S-20°N) over the period 1981-2005. Trend values of T24-TLT are in units of K decade^{-1} and the ratio of T24 to TLT trends are also given as a measure of the vertical temperature trend amplification in the tropical troposphere.

	TLT	T24	T24 - TLT	Ratio
UAH	0.118	0.129	0.011	1.09
RSS	0.203	0.222	0.020	1.10
NOAA	n/a	0.265	n/a	n/a
ACCESS1.0	0.316	0.375	0.059	1.19
ACCESS1.3	0.227	0.260	0.033	1.14
BCC-CSM1.1	0.274	0.317	0.043	1.16
BCC-CSM1.1(m)	0.318	0.376	0.059	1.19
BNU-ESM	0.261	0.314	0.053	1.20
CanCM4	0.283	0.337	0.054	1.19
CanESM2	0.336	0.398	0.062	1.18
CCSM4	0.287	0.345	0.058	1.20
CESM1(BGC)	0.219	0.266	0.047	1.21
CESM1(CAM5)	0.247	0.297	0.050	1.20
CESM1(CAM5.1,FV2)	0.187	0.226	0.039	1.21
CESM1(FASTCHEM)	0.255	0.307	0.051	1.20
CESM1(WACCM)	0.279	0.332	0.053	1.19
CMCC-CM	0.236	0.291	0.054	1.23
CNRM-CM5	0.229	0.250	0.021	1.09
CSIRO-Mk3.6.0	0.241	0.284	0.043	1.18
EC-EARTH	0.240	0.272	0.032	1.13
FGOALS-g2	0.223	0.260	0.036	1.16
FGOALS-s2	0.274	0.319	0.045	1.16
FIO-ESM	0.241	0.292	0.050	1.21
GFDL-CM2.1	0.398	0.464	0.066	1.16
GFDL-CM3	0.364	0.435	0.071	1.20
GFDL-ESM2G	0.452	0.522	0.070	1.15
GFDL-ESM2M	0.319	0.374	0.055	1.17
GISS-E2-H	0.271	0.320	0.049	1.18
GISS-E2-R	0.266	0.316	0.051	1.19
HadCM3	0.290	0.342	0.052	1.18
HadGEM2-AO	0.335	0.390	0.055	1.16
HadGEM2-CC	0.245	0.284	0.039	1.16
HadGEM2-ES	0.252	0.296	0.044	1.18
INM-CM4	0.100	0.120	0.020	1.20
IPSL-CM5A-LR	0.404	0.487	0.083	1.21
IPSL-CM5A-MR	0.341	0.417	0.075	1.22
IPSL-CM5B-LR	0.287	0.356	0.069	1.24
MIROC-ESM	0.153	0.195	0.042	1.28
MIROC-ESM-CHEM	0.083	0.117	0.034	1.42
MIROC4h	0.363	0.437	0.074	1.20
MIROC5	0.190	0.227	0.037	1.20
MPI-ESM-LR	0.362	0.421	0.060	1.17
MPI-ESM-MR	0.325	0.384	0.058	1.18
MPI-ESM-P	0.296	0.348	0.052	1.18
MRI-CGCM3	0.194	0.229	0.036	1.18
NorESM1-M	0.222	0.263	0.040	1.18
NorESM1-ME	0.257	0.307	0.049	1.19
Model Average	0.271	0.322	0.051	1.19

Table 3.2: As in table 3.1 but for the MSU teams and CMIP5 AMIP GCM simulations over the period 1981 - 2008.

TLT	T24	T24-TLT		
UAH	0.082	0.086	0.004	1.05
RSS	0.149	0.164	0.015	1.10
NOAA	n/a	0.205	n/a	n/a
BCC-CSM1.1	0.166	0.192	0.027	1.16
CanAM4	0.174	0.216	0.042	1.24
CCSM4	0.176	0.217	0.042	1.24
CNRM-CM5	0.184	0.209	0.025	1.14
FGOALS-s2	0.171	0.207	0.036	1.21
GFDL-HIRAM-C180	0.140	0.158	0.018	1.13
GFDL-HIRAM-C360	0.143	0.165	0.022	1.15
GISS-E2-R	0.136	0.154	0.018	1.13
HadGEM2-A	0.173	0.199	0.026	1.15
INM-CM4	0.155	0.187	0.032	1.21
IPSL-CM5A-LR	0.163	0.200	0.037	1.23
IPSL-CM5B-LR	0.167	0.203	0.037	1.22
MIROC5	0.184	0.226	0.043	1.23
MPI-ESM-LR	0.191	0.230	0.039	1.20
MPI-ESM-MR	0.190	0.225	0.035	1.19
MRI-AGCM3.2H	0.131	0.157	0.026	1.20
MRI-AGCM3.2S	0.130	0.159	0.029	1.22
MRI-CGCM3	0.171	0.209	0.038	1.22
NorESM1-M	0.174	0.213	0.039	1.22
Model Average	0.164	0.196	0.032	1.19

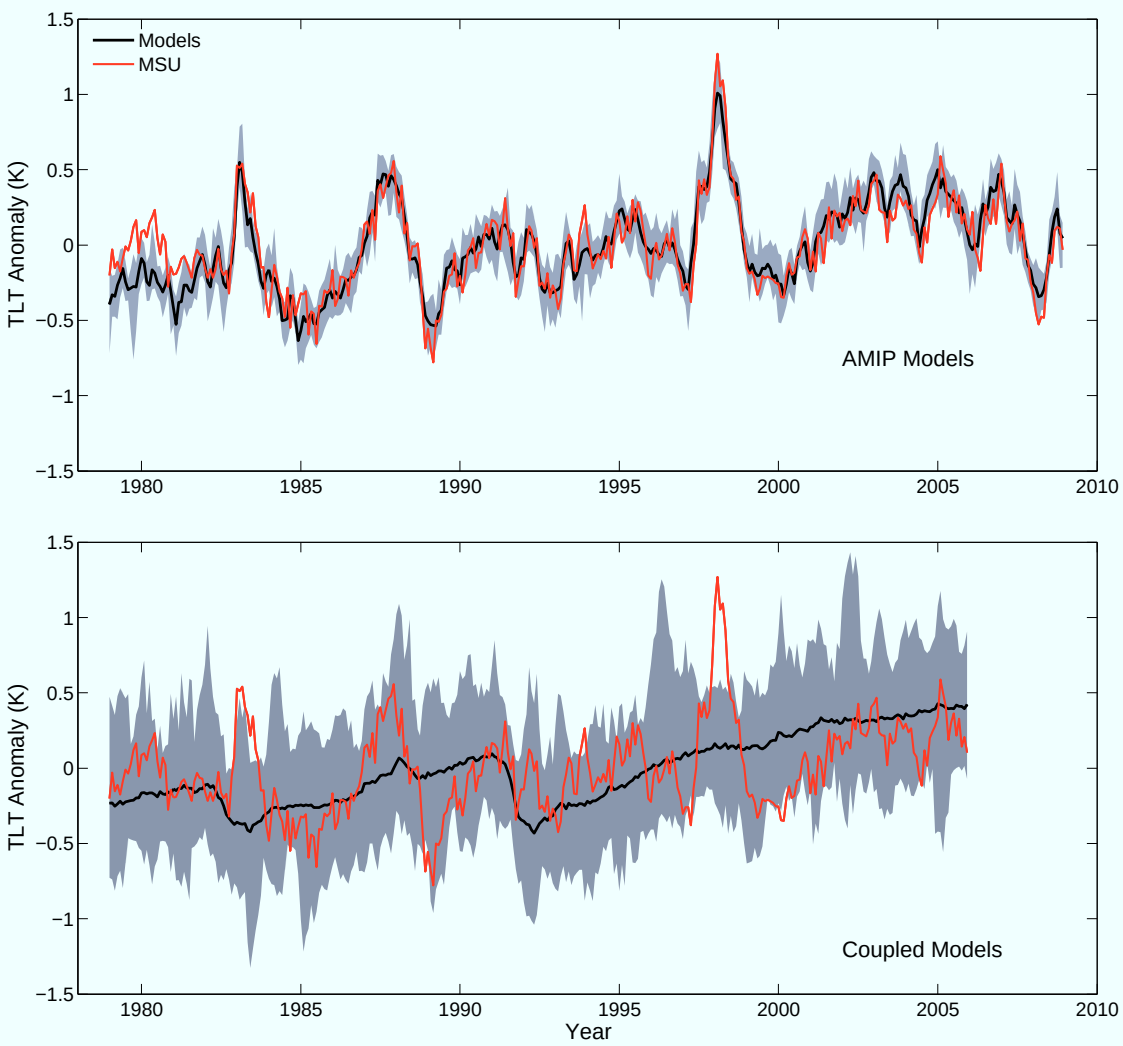

Figure 3.1: Times series of TLT monthly temperature anomalies in the tropics (20^{o}S-20^{o}N) for AMIP GCM (top) and coupled GCM (bottom) simulations (black) and the average of RSS and UAH (red). The model spread is shaded. Many coupled GCM simulations only extended to 2005, while AMIP runs included here extended to 2008.

Since AMIP models are constrained to observed SSTs, it is possible to closely compare the time evolution of the tropospheric amplification between models and observations using the difference between the upper and lower-middle tropospheric temperatures (i.e., T24 - TLT). Figure 3.2 shows the trend of the difference time series between AMIP simulations

and observations for T24 - TLT (i.e., (T24 - TLT)AMIP - (T24 - TLT)MSU). In Figure 3.2 we see that most models have T24 - TLT trends that significantly exceed MSU observations. Every model (in the ensemble mean) has larger trends in T24 - TLT than observations. 17 (12) of the 19 models have significantly (95% confidence) greater T24 - TLT trends than UAH (RSS). Even though GCMs generally have larger T24 - TLT trends than MSU observations, the GFDL and GISS models show good agreement with RSS.

We also repeated the analysis from Fu et al. [2011], comparing the T24 - TLT trends in coupled GCMs and satellites over the period 1981 - 2005. As can be seen from Table 3.1, the results shown here for CMIP5 coupled models are similar to the CMIP3 models in Fu et al. [2011]. The upper tropospheric warming relative to the lower-middle troposphere is consistently larger in models than observations. The T24 - TLT ensemble mean trends from all CMIP5 coupled GCMs except one are significantly larger than zero while T24 - TLT trends from both UAH and RSS are insignificantly different from zero (not shown here). Even though we use a shorter time period than Fu et al. [2011], 27 (15) of the 44 models have trends that are significantly different from UAH (RSS) using two-sample t-tests [Lanzante, 2005]. The reason that a smaller fraction of models are significantly different from MSU data compared to Fu et al. [2011] is because we use seven less years and models often overestimate the interannual variability. No model has an ensemble mean T24 - TLT trend as low as RSS or UAH (Table 3.1). We also note from Table 3.1 that CMIP5 coupled GCMs overestimate tropical warming just as CMIP3 models overestimated tropical warming [Fu et al., 2011].

Figure 3.3 demonstrates the distribution of tropical upper tropospheric amplification factors in AMIP simulations as measured by the ratio of T24 to TLT trends. AMIP models consistently have larger amplification factors than MSU observations, even though the models are constrained with observed SSTs. Of the 55 ensemble members, all have trend amplification exceeding that of UAH and only two ensemble members have less amplification than RSS. In the coupled historical simulations only 5 (6) ensemble members of 185 have amplification ratios less than UAH (RSS) and they are all from the same model (CNRM-CM5). Figure 3.4 demonstrates the scaling of monthly and decadal amplification for the tropical

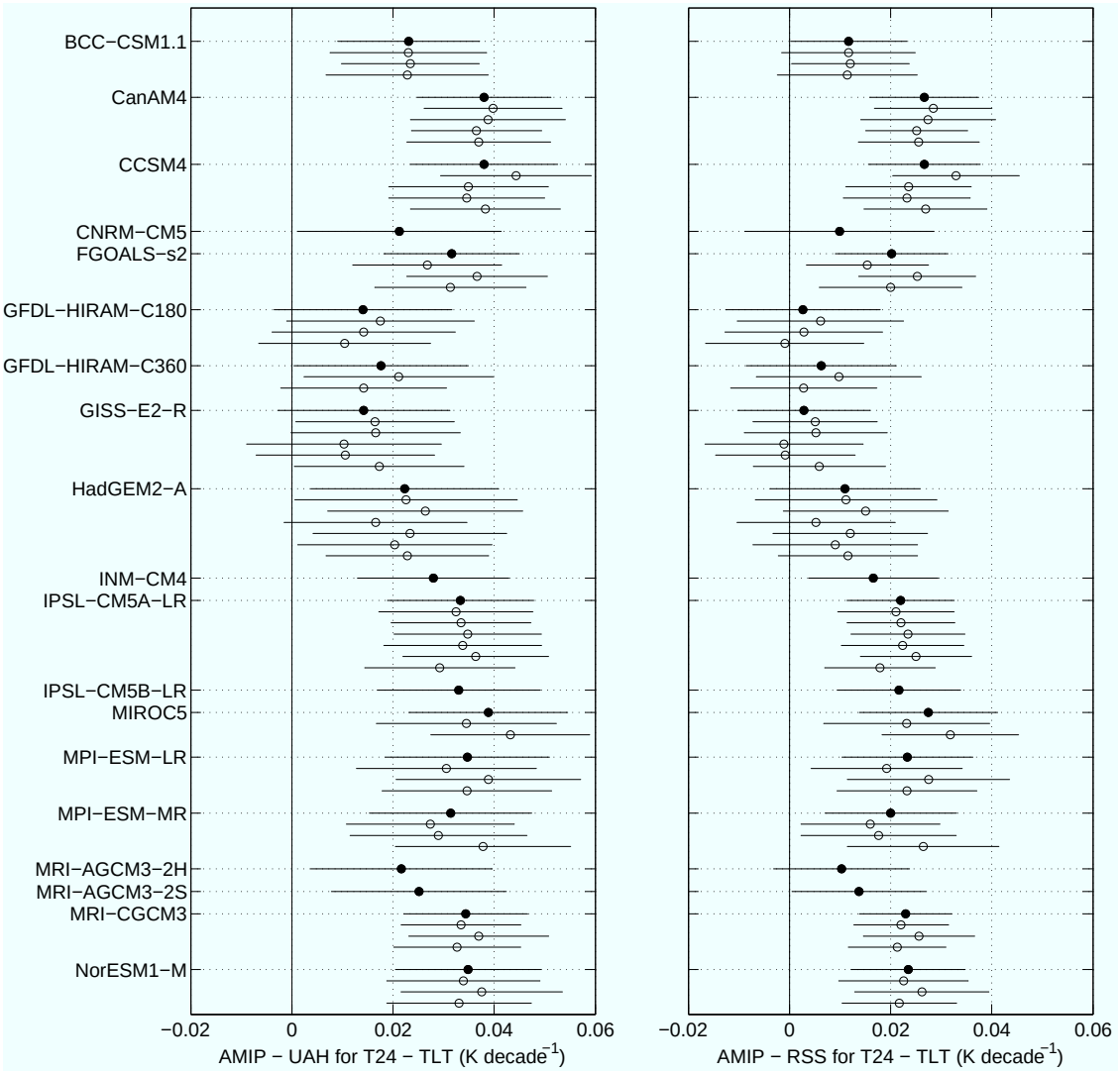

Figure 3.2: Trend of the differences between AMIP simulations and observations for T24-TLT over 1981 - 2008 for UAH (left) and RSS (right) in the tropics (20^{o}S-20^{o}N) (i.e., the trend of (T24-TLT)$_{AMIP}$ - (T24-TLT)$_{MSU}$). Open circles are individual AMIP model ensemble members and the solid circles represent the ensemble mean for a given AMIP model. The error bars represent the 95% confidence interval, including the effects of autocorrelation.

upper troposphere to the lower-middle troposphere. This is similar to monthly and decadal scaling ratios demonstrated in Santer et al. [2005] for the tropical lower-middle troposphere

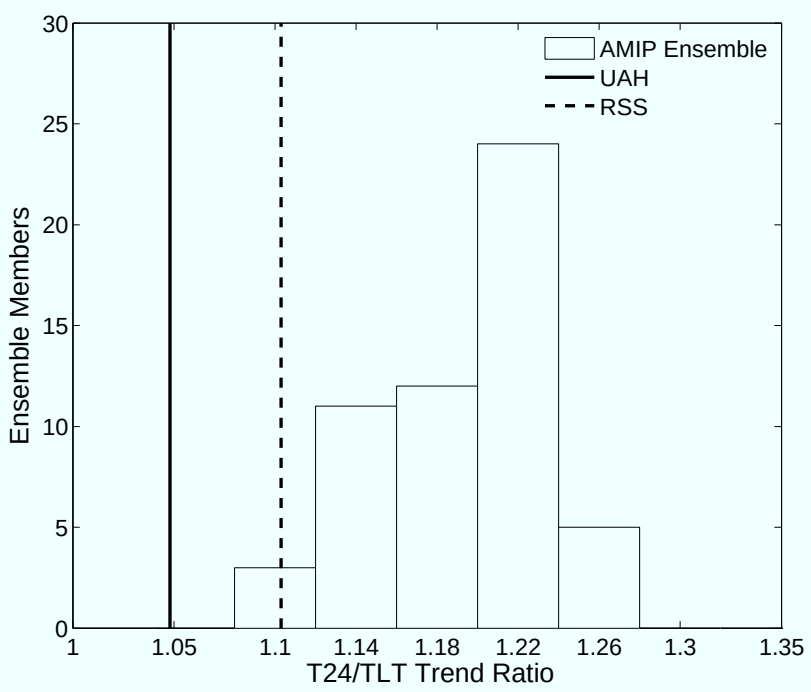

Figure 3.3: Histogram of the ratio of the T24 trend to the TLT trend over 1981 - 2008 from AMIP ensemble members in the tropics (20°S - 20°N). The T24 to TLT trend ratios for RSS and UAH are shown for comparison. The T24/TLT trend ratios under the histogram bins represent the bin center values.

to the surface. It is striking that even though the interannual amplification of UAH and RSS is in the range of the model results, both show less decadal amplification compared to the models. Figures 3.3 and 3.4 suggest that models consistently have greater tropical upper tropospheric amplification compared to satellite MSU observations.

3.5 Discussion and Conclusions

We have demonstrated that GCMs typically exhibit greater tropical upper to lower-middle tropospheric amplification compared to satellite-borne deep layer temperature observations for both coupled historical simulations and simulations constrained with historical SSTs. For most of the AMIP models, the relative warming in the upper troposphere compared to the

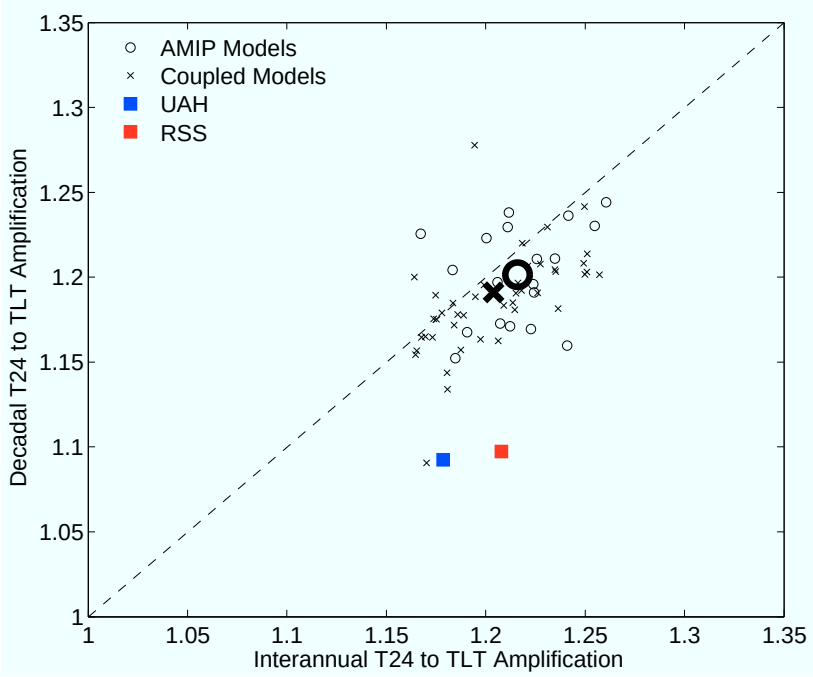

Figure 3.4: Decadal versus interannual amplification of T24 to TLT from both AMIP and coupled GCM simulations and MSU observations in the tropics (20°S-20°N) between 1981 and 2005. The decadal amplification is defined as the T24 trend divided by the TLT trend. The interannual amplification is defined as the standard deviation of the de-trended monthly T24 anomaly time series divided by the standard deviation of the de-trended monthly TLT anomaly values. Each cross or circle represents the ensemble mean for each model. The mean of all models is given by the bold symbols. Note that the MIROC-ESM-CHEM model is not contained in this plot as it has a relatively large decadal amplification value (Table 3.1), likely related to biases after the Mt. Pinatubo eruption in 1991 [Watanabe et al., 2011].

lower-middle troposphere is significantly (95% confidence) larger than both RSS and UAH, but some models such as GFDL and GISS demonstrate good agreement with RSS.

There are a number of explanations for the differences in tropical upper to lower-middle tropospheric temperature trend amplification. The consistent positive bias (though not always significant) in models for T24-TLT trends or T24/TLT trend ratios as compared to both UAH and RSS indicates that this might be a common problem amongst models. In

figure 3.5, we demonstrate the relationship between the annual mean tropical (20°S-20°N) T24 temperature and the ratio of the T24 trend to the TLT trend for AMIP models over 1981-2008. The relationship between the annual mean T24 temperature and the upper- to lower-middle tropospheric amplification is significant (95% confidence) with r = 0.56. One possible explanation for this relationship is that model parameterizations (i.e., deep convection, radiation or clouds) that influence the mean state climate have important implications for the changes in the static stability between the tropical upper and lower-middle troposphere. John and Soden [2007] found that CMIP3 coupled models tended to have cold biases in the tropical troposphere relative to reanalysis models and AIRS satellite infrared measurements and that these biases had little effect on the model's lapse rate and water vapor feedbacks on a global scale. Note that John and Soden [2007] considered the total column global mean tropospheric biases. Although measuring climatological biases are outside the scope of this work, figure 3.5 does suggest that the simulated tropical upper-middle tropospheric climatology may be related to the magnitude of lapse rate changes between the tropical upper and lower-middle troposphere.

The effect of internal uncertainty related to the dataset construction is large [Christy et al., 2003, Zou et al., 2009, Mears et al., 2011], so it is possible that the differences between GCMs and observations are byproducts of the merging procedure for satellite observations. It is unclear why the interannual amplification ratio should be different from the decadal amplification ratio, but MSU observations show less amplification on decadal time scales (figure 3.4). We also note that NOAA T24 has larger upper-middle tropospheric warming compared to RSS and UAH (tables 1 and 2) and that other analyses that use temperature trends derived from wind measurements have found that historical tropical tropospheric warming is largely consistent with GCM results (Allen and Sherwood 2008). In general, the comparison of model and observed trends over a relatively short time period has large uncertainties, so some of the discrepancy noted here may also be related to the length of the datasets.

A further possibility is that the agreement with some models (i.e., GFDL and GISS)

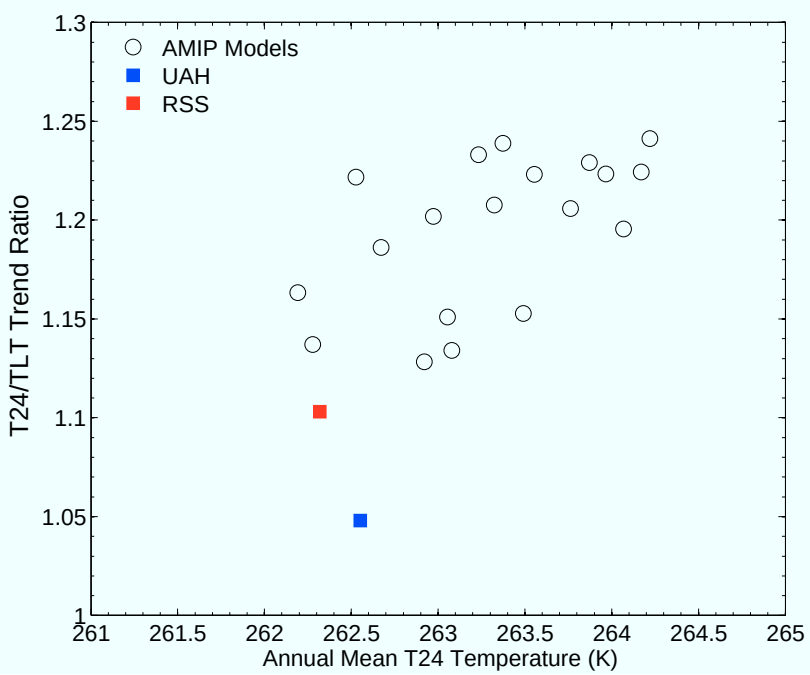

Figure 3.5: Decadal amplification (as in figure 3.4) versus the annual mean T24 temperature over 1981 - 2008 for AMIP models. The relationship is statistically significant (95% confidence) and the r-value is 0.56. The annual mean T24 temperatures are also presented for RSS and UAH for reference. Note that much of the focus for MSU groups has been on relative changes and not on absolute temperature calibration [Mears et al., 2011].

is real and the remaining models exhibit too much tropical upper to lower-middle tropospheric temperature trend amplification. With the latter possibility, it would be important to investigate why some atmospheric models exhibit too much amplification given that the representation of tropical upper to lower- middle tropospheric temperature trend amplification may be important to model estimates of climate sensitivity. In at least two cases, modeling groups (GFDL and NASA GISS) used a different SST dataset (HadISST, Rayner et al. [2003]) than that recommended for the CMIP5 AMIP experiment (Held and Ken Lo 2012). The GFDL and GISS models have tropical upper to lower-middle tropospheric amplification characteristics that agree much better with observations, indicating that the SST

dataset may explain some of the discrepancies noted here.

Given the importance of both models and observations, it will be important to continue to investigate this discrepancy between models and observations. The representation of upper tropospheric warming in models is important to climate sensitivity and thus future projections of anthropogenic global warming. Furthermore, SST and MSU datasets represent vital records for monitoring climate change, so it is important to resolve any errors that may exist. At a more basic level, understanding the discrepancies highlighted here may help advance our understanding of the coupling between the atmosphere and the ocean in the tropics and its implications for climate change.

3.6 Acknowledgments

We would like to thank Drs. Isaac Held, Syukuro Manabe and Dian Seidel for motivating and commenting on this research. We also thank two anonymous reviewers for their helpful comments. This work was supported by the National Science Foundation Graduate Research Fellowship (DGE-0718124), National Science Foundation of China under grant 41275070, DOE grant DE-FG02-09ER64769, and NOAA NESDIS-NESDISPO-2009-2001589 (SDS-09-15). We acknowledge the World Climate Research Programme's Working Group on Coupled Modeling, which is responsible for CMIP, and we thank the climate modeling groups (models are listed in tables 1 and 2 of this paper) for producing and making available their model output. For CMIP, the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

Chapter 4

CONTRIBUTIONS TO DIFFERENCES IN UPPER TROPICAL TROPOSPHERIC WARMING BETWEEN MODELS AND OBSERVATIONS

4.1 Abstract

The vertical profile of tropical tropospheric warming is important to changes in atmospheric circulation and regulates cloud, water vapor, and temperature feedbacks. Past analysis of the change in dry static stability in the tropical troposphere – the ratio of warming between the upper and lower-middle troposphere – have demonstrated that there are significant discrepancies between general circulation models (GCMs) and satellite-borne microwave observations. Via a set of sensitivity experiments, it is shown that these differences are unlikely due to deficiencies in the ozone or stratospheric aerosol forcing used in models. On the other hand, uncertainty in the pattern of sea-surface temperature changes may have an important influence on the vertical structure of atmospheric temperature changes in GCMs. Using estimates of observational uncertainty, approximately half of CMIP5 AMIP ensemble members lie outside the observational estimate for vertical temperature amplification, as measured by the ratio of T24 and TLT trends. The T24/TLT trend ratio is a relatively weak predictor (r=0.61) of the tropical lapse rate feedback in GCMs, though the ratio of T24 and near surface warming is strongly predictive (r=0.98) of the tropical lapse rate feedback. Tropical vertical temperature amplification can be modulated by the formulation of model microphysics and convection schemes over long time scales, though the pattern of SST variability also influences amplification over interannual and decadal timescales. Even though model physics plays an important role, model amplification from simulations over the satellite era only weakly scales with model behavior in CO_2 quadrupling experiments.

4.2 Introduction

Changes in the vertical temperature structure of the tropical troposphere is important to the atmospheric circulation [e.g., Lin and Fu, 2013], the height of clouds [e.g., Hartmann and Larson, 2002], changes in tropical cyclone activity [e.g., Vecchi et al., 2013], and climate feedbacks [e.g., Soden and Held, 2006]. General circulation models (GCMs) robustly exhibit enhanced warming in the tropical upper troposphere relative to the surface, with maximum trends near 200 hPa [e.g., Solomon et al., 2007]. This "vertical amplification" in the troposphere reduces the lapse rate in the tropics leading to a strong negative longwave climate feedback (i.e., the lapse rate feedback) [Soden and Held, 2006]. A number of studies over the last decade have investigated the fidelity in which general circulation models (GCMs) simulate the rate and structure of tropical atmospheric temperature change compared to observational datasets. Some studies have suggested that models have overly pronounced vertical amplification [Christy et al., 2007, Douglass et al., 2008, Christy et al., 2010], though the preponderance of evidence suggests that such claims are overstated [Thorne et al., 2007, Santer et al., 2008] or likely due to observational biases [Fu et al., 2004, Santer et al., 2005, Po-Chedley et al., 2015].

Recent attention has shifted to a more nuanced issue. Although the tropical troposphere broadly exhibits enhanced warming compared to the surface in both models and observations, is the vertical structure of warming in models realistic compared to observations? Fu et al. [2011] demonstrated that even when discrepancies in tropical surface warming were taken into account, coupled atmosphere-ocean GCMs exhibited significantly more warming in the upper troposphere relative to the lower-middle troposphere. This study utilized satellite-borne microwave observations of temperature in broad atmospheric layers, including the upper troposphere (T24) and the lower-middle troposphere (TLT). These measurements, available since late-1978, were compared to synthetic brightness temperatures simulated from model temperature output. In similar work, Seidel et al. [2012] demonstrated that model vertical amplification between 700 and 200 hPa is often significantly larger in GCMs

compared to radiosonde observations (though discrepancies are smaller between 700 and 300 hPa). Both of these studies utilized coupled atmosphere-ocean GCMs from Phase 3 of the Coupled Model Intercomparison Project (CMIP3).

The pattern of surface warming can influence the vertical profile of temperature changes in the tropics [e.g., Flannaghan et al., 2014]. It is therefore possible that model-observational disagreement may arise from internal climate variability over the observational record. In order to minimize the effect of internal climate variability, Po-Chedley and Fu [2012a] compared microwave temperature observations to GCM simulations forced with prescribed sea-surface temperatures (SSTs) from the historical record. This study also found significant discrepancies between models and observations in the warming ratio of the upper and lower-middle troposphere. This study demonstrated that model-observational warming discrepancies were present in the latest set of GCMs (from the CMIP5 archive) and that such discrepancies were not resolved even when models were tightly constrained using observed SSTs. Comparisons using radiosondes and CMIP5 prescribed-SST experiments similarly found that models tended to warm the upper troposphere (near 200 hPa) more than the observational record supports [Mitchell et al., 2013a]. These studies were not able to clearly attribute model-observational discrepancies to models or observations [Fu et al., 2011, Seidel et al., 2012, Po-Chedley and Fu, 2012a, Mitchell et al., 2013a].

A number of factors may contribute to differences between GCMs and satellite observations. One possibility for differences in the vertical pattern of warming between models and satellite microwave observations may result from biases in the prescribed forcing. Recent work has demonstrated that observational forcing datasets have large discrepancies in the derived lower stratospheric ozone trend [Solomon et al., 2013, Hassler et al., 2013, Young et al., 2014], that volcanic aerosols not represented in GCM simulations [Solomon et al., 2011, Vernier et al., 2011] have had an important influence on surface and tropospheric warming [Fyfe et al., 2013, Haywood et al., 2014, Santer et al., 2014], and that uncertainty in the observational record of SSTs is important to the tropical circulation [Tokinaga et al., 2012] and tropospheric warming [Flannaghan et al., 2014, Fueglistaler et al., 2015]. Another possible

driver of model-observational discrepancies in tropical tropospheric amplification could be related to issues simulating tropical convection [e.g., Mitchell et al., 2013a, Mauritsen and Stevens, 2015]. From the observational side, microwave satellite datasets have recently undergone a number of corrections [e.g., Mears and Wentz, 2016], and therefore observational biases may also be important.

This study explores factors that may contribute to the model-observational discrepancies noted above and the influence of inter-model differences in the vertical profile of warming on the lapse rate feedback. In Section 4.3, we outline our experimental design, the datasets used, and the methodological details. In Section 4.4, we estimate the sensitivity of the vertical profile of warming to alternative model forcings and boundary conditions including ozone, volcanic aerosols, and prescribed SSTs. In this section, we also consider the effect of perturbed model physics and estimates of observational uncertainty. In Section 4.5, we look at the connection between satellite-derived warming metrics and the lapse rate feedback. We discuss our findings and conclude in Section 4.6.

4.3 Experimental Design, Methods, and Data

One goal of this research is to understand how various sources of uncertainty contribute to model and observational differences in the profile of tropical atmospheric warming. To understand how GCMs may contribute to such discrepancies, we consider several sensitivity experiments broadly related to model forcings and boundary conditions and model physics. To explore the role of observational biases, we briefly review recent advances in the homogenization of microwave temperature measurements and study an ensemble reconstruction of observations.

4.3.1 Observational Datasets

We utilize satellite-borne microwave brightness temperatures as measures of the temperature of deep atmospheric layers. These microwave radiometer observations principally measure the radiance of microwave emissions from atmospheric oxygen, which can be related to at-

mospheric temperature. These data come from over a dozen polar-orbiting satellites that carry microwave sounding instruments and have provided continuous atmospheric temperature information since 1979. While the original observations were taken by the microwave sounding unit (MSU), this instrument has been replaced by the advanced microwave sounding unit (AMSU) beginning in 1998. A number of groups have homogenized these measurements removing known biases that result from issues such as orbital decay [Wentz and Schabel, 1998], inconsistent diurnal sampling [Mears and Wentz, 2005], and instrument body changes [Christy et al., 2000]. Tropospheric temperature information relies on MSU channel 2 and AMSU channel 5, which are combined to form TMT, the mid-tropospheric temperature [Christy et al., 2000, Mears et al., 2003]. TMT "mid-tropospheric" measurements are partially contaminated by stratospheric temperature changes. In order to create an unbiased measure of tropospheric temperatures, we use a derived tropospheric channel (T24), which uses lower stratospheric temperature channels (TLS from MSU channel 4 and AMSU channel 9) to remove stratospheric contamination from TMT [Fu et al., 2004, Fu and Johanson, 2004]. This derived channel is a weighted average of TMT and TLS such that $T24 = a \cdot TMT + b \cdot TLS$, where $a = 1.1$ and $b = -0.1$ in the tropics [Fu and Johanson, 2004]. We also utilize the lower-middle tropospheric temperature product (TLT), which represents a combination of view angles from TMT. These two metrics, TLT and T24, can be used to provide an estimate of static stability changes in the middle-upper tropical troposphere [Fu et al., 2011, Po-Chedley and Fu, 2012a].

The Remote Sensing Systems (RSS) and University of Alabama in Huntsville (UAH) research groups provide the TLT, TMT, and TLS data needed (T24 is derived from TMT and TLS). The National Oceanic and Atmospheric Administration (NOAA) also produces homogenized MSU/AMSU products [Zou et al., 2006, 2009, Zou and Wang, 2010, 2011], but does not produce a TLT dataset and is therefore not considered here. RSS recently updated their diurnal drift correction in their TMT product [Mears and Wentz, 2016], but this update has not yet been applied to their TLT data. As a result, we will use RSS v3.3 [Mears and Wentz, 2009a,b], though we will also list the T24 trend for the newest version (v4.0) [Mears

and Wentz, 2016]. We also use UAH v5.4 as in Po-Chedley and Fu [2012a], though we also include their latest dataset version, UAH v6.0 beta 5.

In order to assess the observational uncertainty, we utilize an ensemble of observational realizations produced by RSS [Mears et al., 2011]. Mears et al. [2011] used a Monte-Carlo approach to assess the uncertainty in TLT, TMT, and TLS temperature products. Briefly, their approach was to randomly perturb bias corrections related to sampling, the diurnal correction, and instrument calibration and re-merge (400 times) the microwave time series taken from over a dozen satellites. We use this 400 member ensemble of observations to assess the uncertainty in trends in the upper and lower troposphere. To compute T24 trends, we randomly pair TMT and TLS trends together. In pairing upper (T24) and lower (TLT) tropospheric trends to compute the T24/TLT trend ratio, we randomly group T24 and TLT trends in decile (i.e., 10%) bins. For example, we take the largest ten percent (40 ensemble members) of TLT and T24 trends and randomly match them together. This method eliminates physically unrealistic comparisons (e.g., a near-zero TLT trend and a large T24 trend). Pairing the data in pentiles (20% bins) does not significantly alter the results.

4.3.2 Model Experiments

Sensitivity to post-2000 volcanic aerosols

Solomon et al. [2011] demonstrated that non-negligible volcanic aerosols persist in the stratosphere even in the absence of large volcanic eruptions. This is important because these aerosols have a non-zero radiative effect and CMIP5 models are forced only with background stratospheric aerosol levels after the year 2000 [Solomon et al., 2011, Haywood et al., 2014]. Several studies demonstrated that the inclusion of more realistic volcanic forcings in GCMs decreases the surface [Fyfe et al., 2013, Haywood et al., 2014] and tropospheric temperature [Santer et al., 2014] after 2000. We analyze two sets of atmosphere-ocean GCM simulations that use a stratospheric aerosol forcing that has been updated to include realistic volcanic

aerosol changes after 2000. These model simulations are described in Santer et al. [2014] and include coupled atmosphere-ocean simulations from the GISS-E2-R and CanESM2 model. The GISS simulations include control experiments in which stratospheric aerosol optical depth decayed to zero at the end of the 20th century and simulations with observed aerosol optical depths that extend past the year 2000 based on an updated version of the Sato et al. [1993] dataset. Similarly, the CanESM2 model experiments (described in Fyfe et al. [2013]) include control simulations with stratospheric aerosol optical depths that decay to background levels near 2000 and simulations that include observed optical depths after 2000 (from Vernier et al. [2011]). Each model includes five control simulations (with near-zero stratospheric aerosol optical depth after 2000) and five simulations with updated stratospheric volcanic aerosol forcing based on observations. In each case, synthetic satellite-borne microwave brightness temperatures are computed for the models [Santer et al., 2014]. Brightness temperature estimates are made for the lower troposphere (TLT), the mid-troposphere (TMT), and the lower stratosphere (TLS). From TMT and TLS, we compute an upper tropospheric temperature (T24). We then compare the vertical warming differences in the control and updated volcanic simulations using T24 and TLT. We consider the warming trends over 1981 - 2008 as in Po-Chedley and Fu [2012a]. This time period is used throughout this analysis because it avoids time periods in which SST datasets and microwave observations have large inconsistencies [Po-Chedley and Fu, 2012a, Flannaghan et al., 2014].

Sensitivity to varied ozone changes

There are large differences in observed ozone trends in the lower tropical stratosphere [Hassler et al., 2013], which influence model-derived stratospheric and upper tropospheric temperature trends [Solomon et al., 2013, Eyring et al., 2013]. To estimate how much CMIP5 model simulations might differ based on different observational estimates of ozone concentration trends, we perform a set of time-slice experiments [as in Polvani and Solomon, 2012, Solomon et al., 2013]. We use CESM version 1.0.5 with prescribed year 2000 atmospheric forcings (other than ozone) and sea-surface temperatures. This version of the model utilizes version

4 of the Community Atmosphere Model (CAM4). The model is run with a spatial resolution of approximately 2° and has 26 vertical levels with a model top at 2.2 hPa. We prescribe the tropical stratospheric ozone concentration change using the Binary DataBase of Profiles (BDBP) Ozone Dataset [Bodeker et al., 2013] and the Stratospheric Processes and their Role in Climate (SPARC) dataset [Cionni et al., 2011]. The stratospheric (above 100 hPa) year 2000 ozone forcing is perturbed by scaling the zonally averaged ozone concentration to match the fractional ozone loss between the period 1979 - 1981 and 2005 - 2007. We do not use another prominent ozone dataset produced by Randel and Wu [2007], because the tropical stratospheric ozone trends and temperature response are similar to that in the SPARC dataset [Solomon et al., 2013, Hassler et al., 2013]. We integrate the control and each ozone loss experiment (BDBP and SPARC) for 40 years, but only consider the last 30 years from each simulation. We calculate monthly MSU brightness temperatures for each channel using static weighting functions at each gridpoint. Any weight below the model surface (e.g., in regions of high terrain) is assigned to the surface. The weighting functions were provided by Remote Sensing Systems.

Sensitivity to varied sea-surface temperature changes

It has long been known that the pattern of tropical warming can influence the vertical atmospheric response. Deep convection preferentially occurs over the warmest SSTs [e.g., Wallace, 1992, Johnson and Xie, 2010]. Increasing SST in the warmest waters therefore has the effect of heating the mid- and upper troposphere. This tropospheric warming is quickly distributed throughout the tropics, which do not maintain strong horizontal temperature gradients [e.g., Bretherton and Sobel, 2003]. Thus temperature changes in the upper troposphere are influenced by temperature trends over the warmest SSTs. Several studies have recently noted that the differences in the pattern of warming across observational SST datasets are large enough to significantly influence tropospheric temperature trends [Po-Chedley and Fu, 2012a, Flannaghan et al., 2014, Fueglistaler et al., 2015]. It has also been noted that such differences can also have an important effect on the tropical circulation [Tokinaga et al., 2012]. To under-

stand the sensitivity of the vertical profile of tropical warming to different SST datasets, we prescribe three SST datasets to CESM 1.0.5. This model is setup in the same way as in the ozone sensitivity experiments, except that we allow SSTs to evolve (we use year 2000 concentrations of atmospheric constituents). We utilize the SST dataset that was intended for CMIP5 Atmospheric Model Intercomparison Project (AMIP) experiments (Hurrell) [Hurrell et al., 2008], the Hadley Centre Sea Ice and Sea Surface Temperature dataset (HadISST1.1) [Rayner et al., 2003], and a derived dataset that uses nighttime marine air temperatures (NMAT) [based on Tokinaga et al., 2012]. The Hurrell and HadISST datasets are simply re-gridded and prescribed to the atmospheric component of CESM.

NMAT data is based on the Hadley Centre National Oceanography Centre Night Marine Air dataset [Kent et al., 2013]. Nighttime marine air temperatures closely track SST and have been shown to be more consistent with observed sea level pressure changes compared to other SST datasets [Tokinaga et al., 2012, Kent et al., 2013]. These observations are also attractive because some of the corrections applied to SST measurements do not apply [Tokinaga et al., 2012], though we caution that important and uncertain corrections are still applied to these data [Kent et al., 2013]. Since nighttime marine air temperature observations are taken from ships, there is substantial missing data. As a result, we blend the derived NMAT dataset into the Hurrell SST data. The SST time series can be decomposed into the long-term linear trend and the variability about the trend:

$$SST = \overline{SST} + SST_{trend} + SST' \tag{4.1}$$

where $\overline{SST}$ is the climatological mean temperature, SST_{trend} represents the long-term least-squares linear trend and SST' represents the monthly variability. To create a complete SST dataset that relies on NMAT trends, we retain the variability from the Hurrell dataset, but use trends computed using the NMAT data:

$$SST_{NMAT} = \overline{SST_{Hurrell}} + SST_{trend,NMAT} + SST'_{Hurrell} \tag{4.2}$$

Within 20°N-S, the SST trends are forced to match NMAT data, over 20 - 30°N-S the trends are a simple average of the NMAT and Hurrell data, and outside 30°N/S, the trends are the same as in the original Hurrell SST dataset. We view the inclusion of nighttime marine air temperatures as a plausible sensitivity experiment for issues related to amplification, not as a substitute for carefully homogenized SST datasets. We run these alternate SST experiments using 1981 - 2008 SSTs and MSU synthetic temperatures are calculated as in the ozone perturbation experiments. The simulations are referred to by the SST dataset used to force the models, Hurrell, HadISST, and NMAT.

Model Physics Sensitivity

Aside from prescribed forcings and SSTs, model vertical amplification may also be affected by model parameterizations for unresolved, sub-grid scale processes [e.g., Po-Chedley and Fu, 2012a, Mitchell et al., 2013a]. We consider an alternative model with changes made to the model parameterization for convection. The perturbed model is based on the Goddard Institute for Space Studies (GISS) Model E2 that was used as part of the CMIP5 experiment [Field et al., 2014]. The original (control simulation) model [Schmidt et al., 2014] had low cloud cover that peaked unrealistically close the surface and was unable to develop a Madden-Julien Oscillation (MJO). The model was modified such that the effective entrainment rate and convection scheme sensitivity to moisture was increased [e.g., Kim et al., 2011]. This inhibited "zero entrainment rate" convection events that had inhibited large-scale organization of convection, which dramatically improved the simulation of the MJO [Kim et al., 2012]. Changes were also made to the turbulence parameterization, which resulted in a more realistic vertical distribution of clouds [Yao and Cheng, 2012]. Further changes were made to constrain model parameters, especially the model entrainment rate, using the horizontal distribution of atmospheric water isotope fractionation [Field et al., 2014]. We refer to these experiments as the GISS control and "MJO" experiments. These experiments used the HadISST SST dataset in place of a coupled ocean. The control and perturbation simulation were run over 1981 - 2008 and MSU brightness temperatures were computed as in the ozone

perturbation experiments.

In Section 4.5, we consider two ensembles of perturbed physics models that have varied tropical lapse rate feedbacks. One ensemble (9 ensemble members) is derived from a developmental version of a NOAA GFDL model, with changes made to the parameterizations for the conversion of cloud water to rain. This ensemble is described in Zhao et al. [2016], which noted that the cloud feedback is strongly modulated by decisions in formulating the model's microphysics scheme. It was shown that the model climatologies in each ensemble member are realistic, though it was not possible to determine which set of model physics is most realistic based on existing understanding of microphysics. The runs include a control simulation with prescribed SSTs and a companion simulation with SSTs increased by 2 K. We use the difference of these monthly climatologies to estimate the atmospheric response to warming. The second perturbed physics ensemble was based on the CESM 1.0.5 (as in the ozone and SST perturbation experiments). To perturb the model we used the model default parameterization and then alter the convective timescale (multiplying it by 50, 125, 150, 175, and 200%) and, separately, a scaling factor for deep convection lateral entrainment (multiplying it by 50, 75, and 200%). This resulted in 9 separate model ensemble members. We used the default SST dataset and then re-ran the simulations with 4 K added to SSTs to determine the annual mean atmospheric response.

In Section 4.5, we also compare the perturbed physics ensembles with the CMIP5 archive model responses (using 28 models) and temperature change profiles that result from moist adiabatic warming. The CMIP5 model results are based on the "historical" simulation in which coupled models are prescribed 20th century forcings. We estimate the model responses over 1981 - 2005 based on least-squares linear temperature trends. We also estimate quasi-equilibrium GCM responses by comparing model simulations of abrupt CO_2 quadrupling (120 - 150 years after CO_2 quadrupling) with pre-industrial (i.e., "piControl") control simulations (100 year averages). We use the difference of these two experiments to estimate the forced model response. See Section 5.3 for more details on these simulations. In the perturbed physics experiments as well as the CMIP5 experiments we calculated the synthetic

satellite brightness temperatures as in the ozone experiments. The lapse rate feedback was calculated using radiative kernels (see Section 5.3). Last, to estimate the atmospheric response to moist adiabatic warming, we follow Flannaghan et al. [2014], though we assume a dry adiabatic response from the surface to 850hPa and a moist adiabatic response above. We launch our parcel from a surface pressure of 1013 hPa with a baseline surface temperature of 298 K. We find the response by differencing the baseline temperature profile with that after increasing the surface temperature. We calculate the lapse rate feedback with tropical average (20°S-N) radiative kernels assuming a tropopause height of 100 hPa. The synthetic MSU brightness temperatures for moist adiabatic warming profiles are calculated using static weighting functions assuming a stratospheric cooling (above 100 hPa) of -0.4 K.

4.4 On the differences between models and observations

For orientation, we show the weighting functions used in this study in Fig. 4.1. The T24 weighting function peaks near 300 hPa, which emphasizes changes in the upper tropospheric temperature. The TLT weighting function peaks in the lower-middle troposphere, near 650 hPa. Both weighting functions represent broad layers and so are not solely upper or lower tropospheric measurements. Several studies have used the ratio and difference of each synthetic channel's trends as proxies for changes in dry static stability [Fu et al., 2011, Po-Chedley and Fu, 2012a]. Models tend to warm the upper (T24) troposphere relative to the lower-middle troposphere (TLT) significantly faster than observations support.

Although models and observations diverge in their estimates of vertical warming amplification, it is unclear to what degree model or observational biases may contribute to this discrepancy. Since models are forced with historical ozone, stratospheric aerosols, and SSTs and details of these estimates vary across observational estimates, we investigate the sensitivity of models to varied ozone decline, stratospheric aerosol changes, and SST trends. We then turn to estimates of uncertainty in the microwave temperature record and discuss their role in model-observational warming differences.

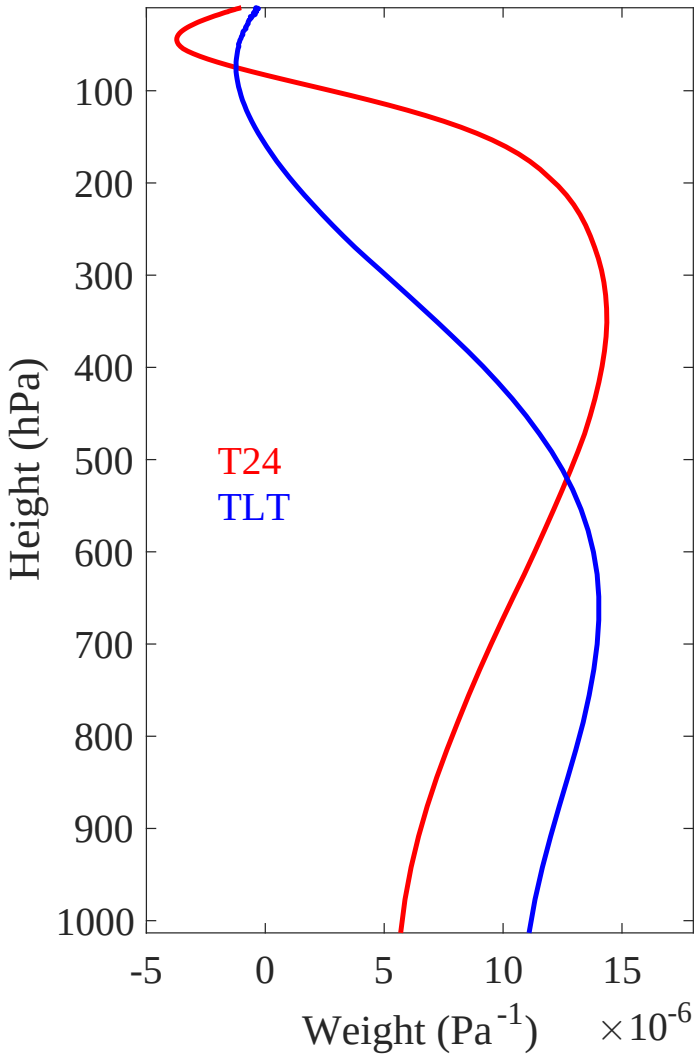

Figure 4.1: Weighting function for the upper (T24) and lower-middle tropospheric microwave sounding data used in this study. The weighting functions are provided from RSS. The surface weight for TLT (T24) is 0.14 (0.06).

4.4.1 The role of forcing discrepancies

Ozone forcing

There are large differences in the loss of tropical stratospheric ozone across datasets [Hassler et al., 2013], which results in large differences in model response in the stratosphere and smaller difference in the troposphere [Polvani and Solomon, 2012, Solomon et al., 2013]. We consider differences in GCM response based on the BDBP and SPARC datasets.

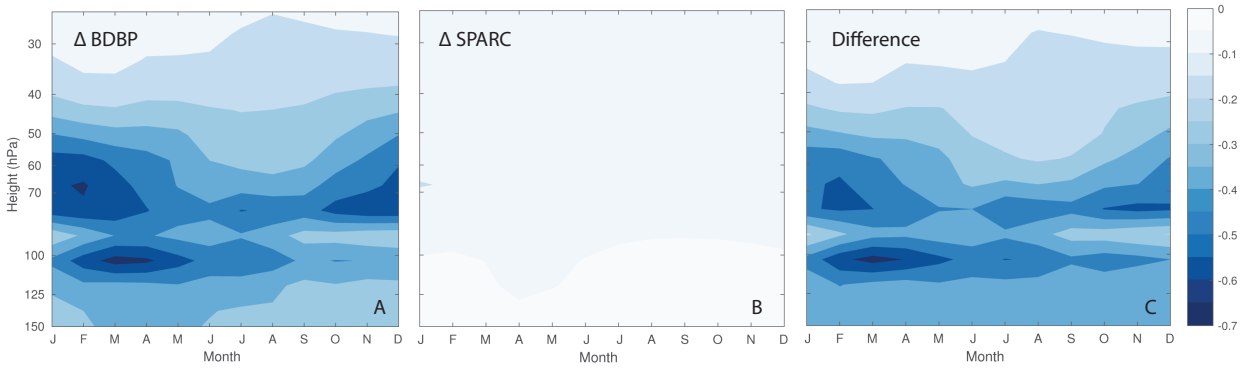

Figure 4.2: Seasonal cycle and vertical structure of fractional ozone changes when comparing 2006 - 2007 minus 1979 - 1980 for the A) BDBP and B) SPARC dataset. We also show the differences in ozone loss (BDBP - SPARC) in (C). In each case, we used the zonal mean ozone changes over 20°S - 20°N.

In Fig. 4.2 we show the loss of tropical stratospheric ozone in the BDBP and SPARC dataset over 1979 - 2007. The tropical lower stratospheric ozone loss in the BDBP dataset exceeds 40% for much of the seasonal cycle, whereas the SPARC dataset shows more modest ozone loss (less than 10%). We prescribe the stratospheric ozone loss as realized in each ozone dataset to the CAM4 atmospheric model and compare the response to an experiment with unchanged ozone concentrations (described in Section 4.3).

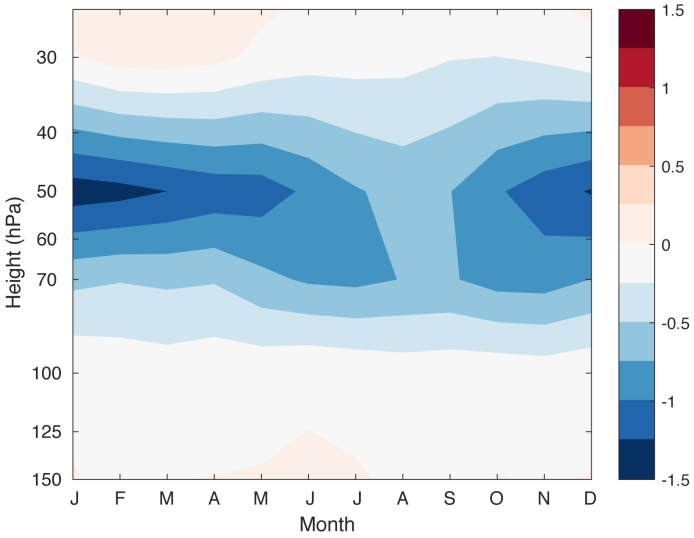

Figure 4.3: Seasonal and vertical differences in the temperature response (K) to different ozone forcing changes (BDBP response minus SPARC response) over 20°S - 20°N.

The difference in the model's tropical temperature response is shown in Fig. 4.3. CAM4 forced with the BDBP dataset has much stronger cooling in the mid and lower stratosphere, particularly in boreal winter and spring, though the differences are near-zero in the tropical troposphere and upper stratosphere. The T24 temperature response is quite small, -0.002 and 0.007 K decade^{-1} for SPARC and BDBP, respectively. The TLT response is similarly small, with trends of -0.003 (SPARC) and 0.003 K decade^{-1} (BDBP). For a typical model (e.g., from the AMIP model mean T24 and TLT trend in Table 4.1), the effect of using the BDBP dataset in place of the SPARC dataset would be small: the T24/TLT trend amplification would change from 1.195 to 1.205. This effect is essentially negligible for tropical tropospheric amplification, which indicates that even though there is some residual weight in the stratosphere (see Fig. 4.1) and a small upper tropospheric temperature response (see Fig. 4.3), the uncertainty in ozone trends in the tropical stratosphere is unimportant for tropical vertical temperature amplification in the troposphere. Although the model temperature responses in the troposphere are similar for varied stratospheric ozone forcing, the lower stratospheric trend differences are substantial. The lower stratospheric trend as measured by the lower stratospheric channel in microwave sounders (TLS) is -0.469 and -0.100 K decade^{-1} for the BDBP and SPARC dataset, respectively. Although not the focus of this analysis, these trend differences are quite large, which suggests that observational uncertainty in ozone trends may explain some of the discrepancies between models and observations for lower-stratospheric temperature trends [Thompson et al., 2012].

Volcanic forcing

In most GCM simulations in the CMIP5 archive, model simulations over the satellite era included stratospheric aerosols that decayed to near-zero after the year 2000 [e.g., Fyfe et al., 2013]. In reality, stratospheric aerosols persisted in radiatively significant concentrations [Vernier et al., 2011, Solomon et al., 2011]. Had this forcing been correctly included, surface [Fyfe et al., 2013, Haywood et al., 2014] and tropospheric temperature trends [Santer et al., 2014] would have been slightly smaller. We consider two model experiments with

updated stratospheric aerosol optical depth forcing to see if this forcing difference could have influenced tropospheric amplification.

The model's simulated MSU/AMSU temperature trend for TLT and T24 is listed in Table 4.1. We also include these values for homogenized microwave observations (from RSS and UAH) and values for CMIP5 AMIP and historical simulations for reference. Although the experiments with updated volcanic forcing (New Volcanic) for the GISS-E2-R and CanESM2 model experiments show reduced tropospheric warming compared to the control simulations (Old Volcanic) (trends reduced $\sim$0.02 - 0.05 K decade^{-1}), the T24/TLT trend ratio stays the same in both models. We can therefore conclude that while biases in the prescribed stratospheric aerosols may help explain model-observational discrepancies in the rate of warming, such biases do not appear to influence differences in the structure of atmospheric warming.

SST boundary conditions

Different SST datasets lead to large differences in the atmospheric temperature response when prescribed to GCMs, which suggests that the pattern of SST warming is important to the tropospheric temperature response [Po-Chedley and Fu, 2012a, Flannaghan et al., 2014]. In particular, trends over the warm pool are particularly important for the tropical mean tropospheric response. Weighting the SST changes by precipitation has the effect of weighting regions of deep convection and is strongly predictive of bulk tropospheric temperature changes [Sobel et al., 2002, Flannaghan et al., 2014, Fueglistaler et al., 2015]. We illustrate this in Fig. 4.4, which shows that annual upper tropospheric temperature (T24) differences between model's forced with different observational SST datasets are relatively weakly related to differences in the tropical average temperature, but strongly related to differences in the precipitation-weighted near surface temperature over ocean.

Dataset	TLT Trend	T24 Trend	Difference	Ratio
UAH v6b5 (v5.4)	0.118 (0.082)	0.114 (0.086)	-0.004 (0.004)	0.966 (1.05)
RSS v3.3 (v4.0)	0.146	0.160 (0.207)	0.015	1.10
CMIP5 AMIP Average	0.164	0.196	0.032	1.19
CMIP5 AOGCM Average	0.271	0.322	0.051	1.19
GISS-E2-R (Old Volcanic)	0.313 (0.278 - 0.340)	0.369 (0.333 - 0.398)	0.056 (0.053 - 0.058)	1.18 (1.16 - 1.20)
GISS-E2-R (New Volcanic)	0.286 (0.252 - 0.329)	0.336 (0.297 - 0.389)	0.050 (0.044 - 0.060)	1.18 (1.17 - 1.18)
CanESM (Old Volcanic)	0.363 (0.265 - 0.407)	0.424 (0.314 - 0.478)	0.061 (0.049 - 0.071)	1.17 (1.15 - 1.18)
CanESM (New Volcanic)	0.317 (0.264 - 0.396)	0.370 (0.306 - 0.468)	0.053 (0.042 - 0.071)	1.17 (1.16 - 1.18)
CESM (Hurrell)	0.159	0.194	0.035	1.22
CESM (HadISST)	0.124	0.149	0.025	1.20
CESM (NMAT)	0.089	0.102	0.013	1.15
GISS MJO Control	0.157	0.179	0.022	1.15
GISS MJO	0.156	0.169	0.013	1.08

Table 4.1: TLT and T24 tropical (20°N-S) temperature trends (K decade^{-1}) for observational datasets, CMIP5 multi-model averages, and model experiments considered in this study. We also include the difference and ratio of T24 and TLT trends. The satellite datasets include UAH v5.4 and RSS v3.3 (as in Po-Chedley and Fu [2012a]). We also include the most recent UAH dataset, v6 beta 5 (in parenthesis). The CMIP5 multi-model mean atmosphere-ocean coupled (CMIP5 AOGCM) and prescribed SST (AMIP) values are from Po-Chedley and Fu [2012a]. We also compare model responses to the standard (Old Volcanic) and updated (New Volcanic) volcanic forcings for the GISS-E2-R and CanESM model. The parenthesis represent the range in model responses over five ensemble members. The response of the CAM4 atmospheric model to the Hurrell, HadISST, and NMAT SST datasets are listed. Last, we also include trends from the GISS control model (GISS MJO Control) and the GISS model with improved physics (GISS MJO). All trends are over 1981 - 2008, except the CMIP5 AOGCM values, which are over 1979 - 2005.

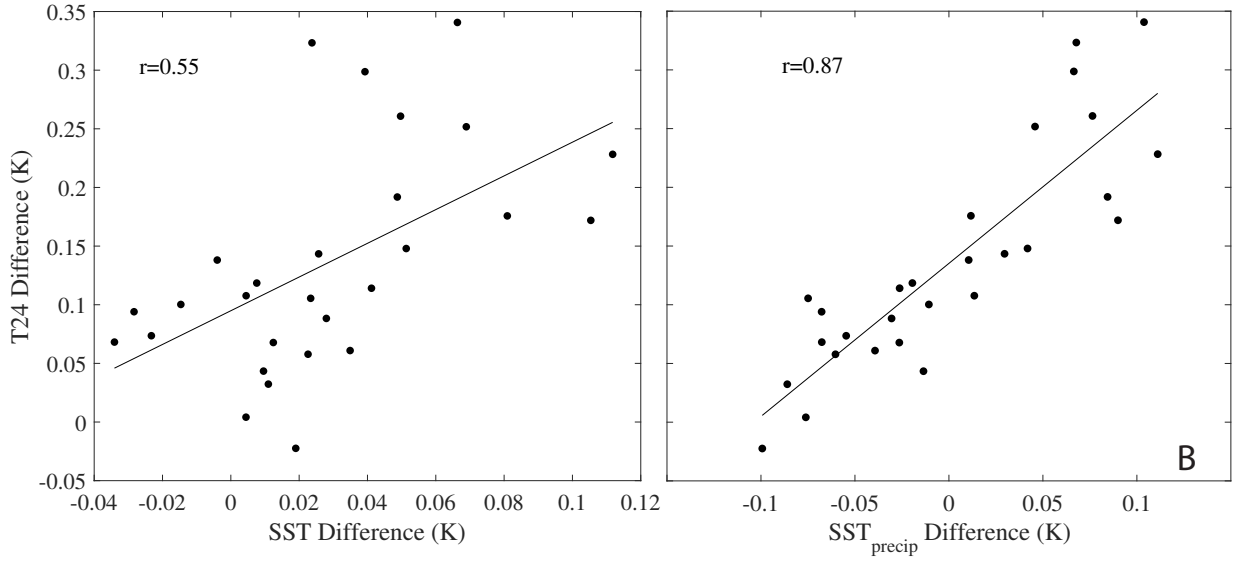

Figure 4.4: Regression of tropical average (20°N-S) annual T24 differences for CAM4 forced with different SST datasets (Hurrell minus HadISST) versus A) the tropical average annual SST difference (Hurrell minus HadISST) and B) the precipitation-weighted annual surface temperature differences over the tropical ocean (Hurrell minus HadISST). Each point represents one year of data. Precipitation used in the weighting comes from model simulation output.

The pattern of surface temperature changes could also affect satellite metrics for amplification. Upper tropospheric warming throughout the tropics (as measured by T24) is particularly sensitive to SST changes over the hottest SSTs, which links T24 to the tropical warm pool. TLT, on the other hand, which has more surface weight and peaks lower in the atmosphere (approximately one-third of the temperature information comes from the surface to 850hPa) is more sensitive to local SST changes. In Fig. 4.5, we show the spatial pattern of SST trends for the Hurrell dataset and the deviation from this pattern for the HadISST dataset and our derived NMAT dataset. Trends in the HadISST dataset are reduced relative to Hurrell SSTs throughout much of the western Pacific warm pool. Warm pool trends are further reduced in the NMAT dataset, though the NMAT dataset suggests that there may be more warming along the equatorial eastern Pacific.

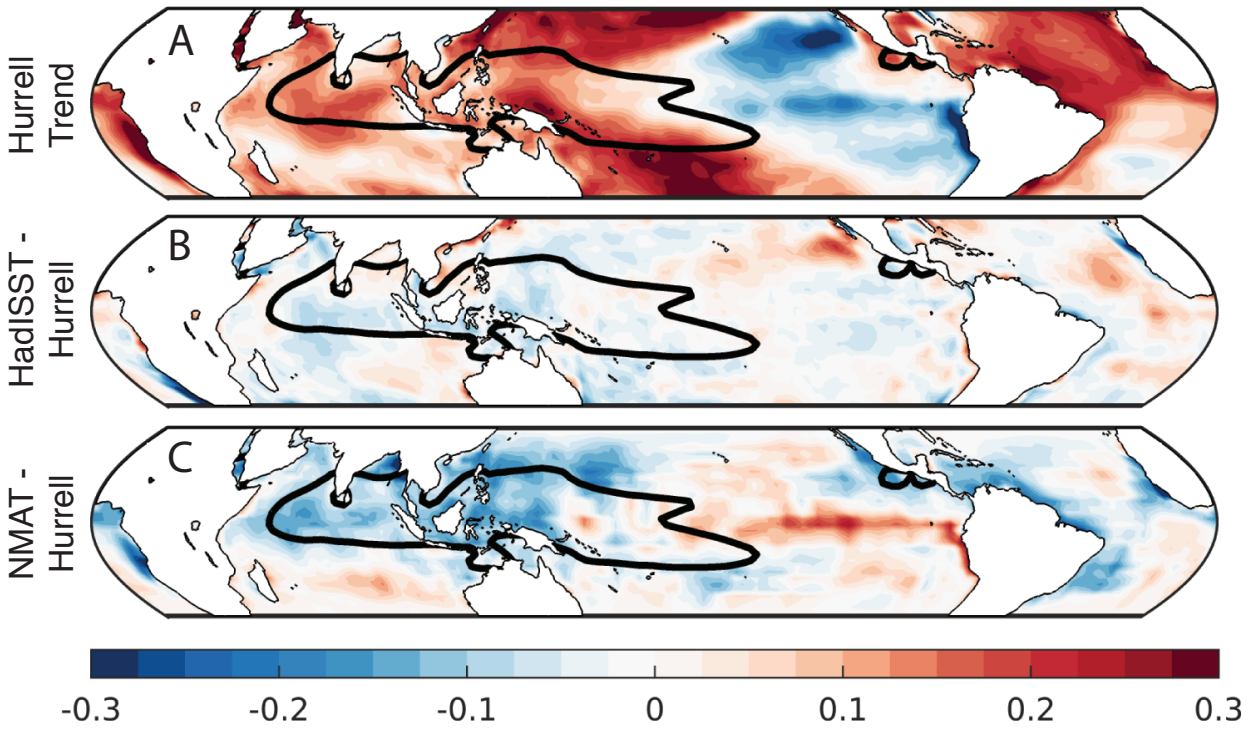

Figure 4.5: A) SST trends for the Hurrell dataset (K decade^{-1}) over 30°N-S (1981 - 2008) and trend difference for B) HadISST - Hurrell and C) NMAT - Hurrell. The black contour in each figure shows the annual average 28.5°C isotherm for reference.

We also show the SST trends as a function of the annual mean SST along with the average and precipitation-weighted SST trends in Fig. 4.6. This figure shows that there are large discrepancies in SST datasets over the warmest SSTs, though the NMAT dataset also has much lower trends over moderate SSTs. These differences in the spatial pattern influence both the tropical average and the precipitation-weighted SST trends. For both metrics, the Hurrell dataset has the largest trends followed by the HadISST dataset and then the NMAT data.

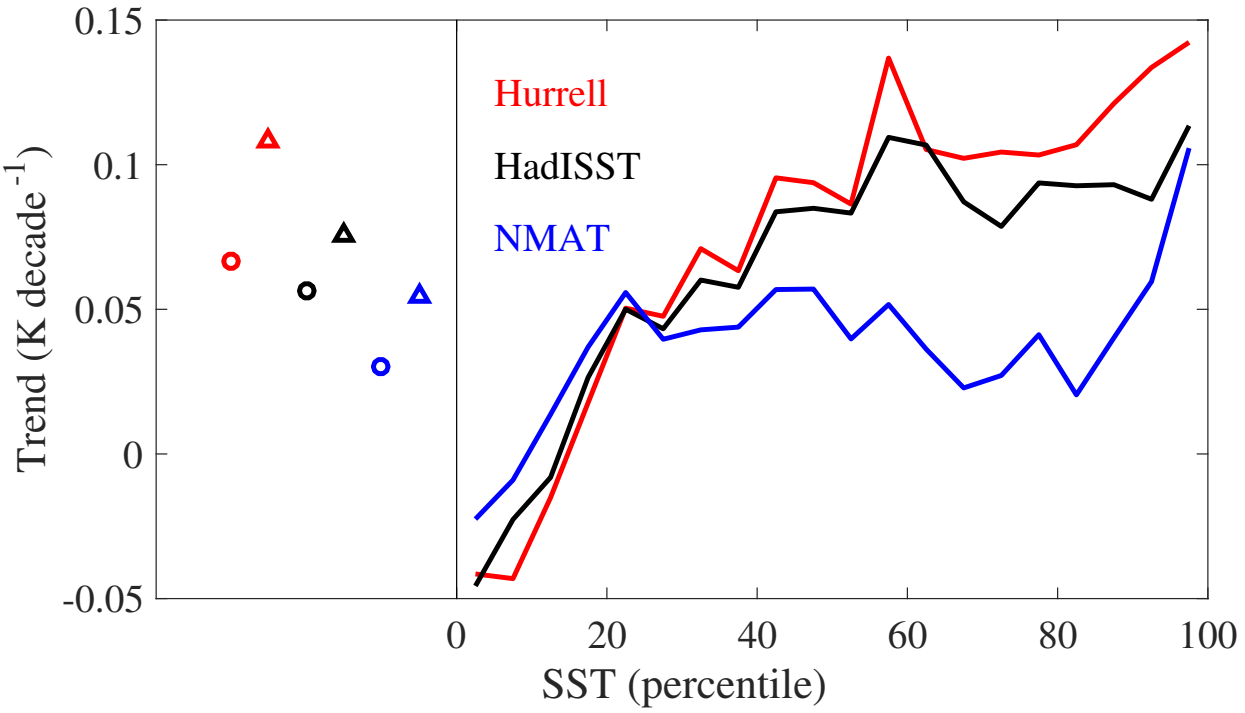

Figure 4.6: Left: Tropical (20°N-S) SST trends (circles) and precipitation weighted SST trends (triangles) for the Hurrell datasets (red), the HadISST dataset (black), and the NMAT dataset (blue). Right: temperature trends as a function of annual mean SST temperature percentile (by constructing a distribution using the long-term mean SST at each gridpoint). The trends corresponding to the coldest waters are on the left and the trends over the warmest waters are on the right hand side (as in Flannaghan et al. [2014]).

When these SST datasets are prescribed in CAM4, the magnitude of the model's tropical tropospheric warming response varies widely. In Fig. 4.7, we show that the atmospheric trend response scales with the precipitation-weighted SST trends; the atmospheric warming in response to the Hurrell dataset is largest, followed by the HadISST dataset, and then the NMAT dataset. In each experiment, the upper troposphere is most variable, which results because model temperature changes follow a moist adiabatic warming profile [Santer et al., 2005]. The differences in SST warming thus have a differential effect on the upper and lower troposphere, with trend differences amplified aloft.

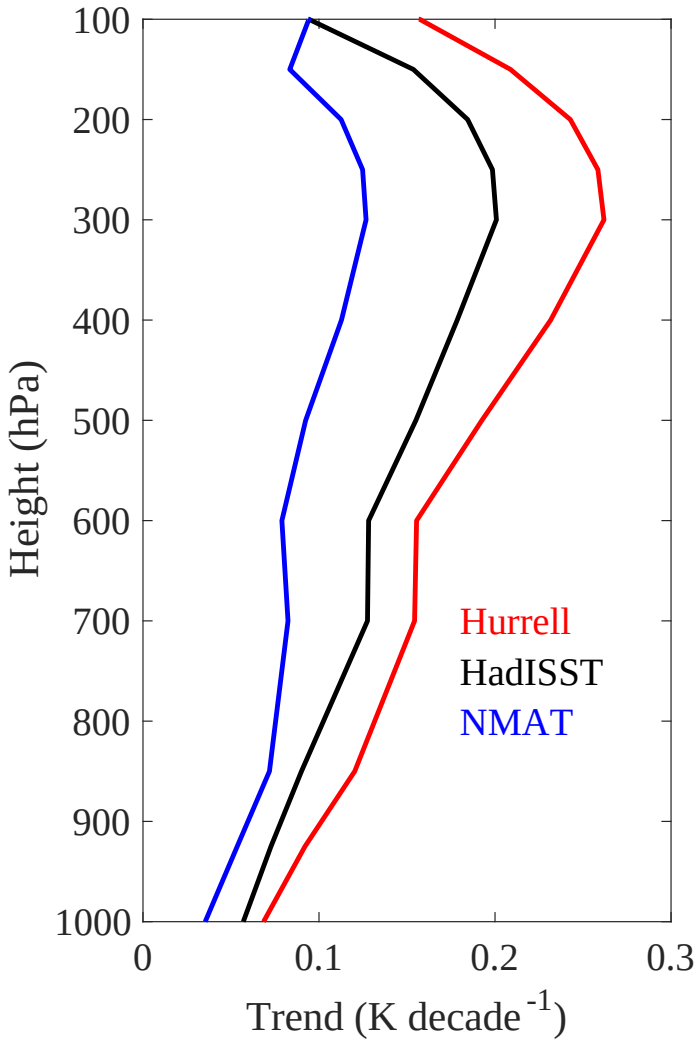

Figure 4.7: CAM4 tropical (20°N-S) average tropospheric trend response (1981 - 2008) to different SST datasets.

These differences in the pattern of surface warming influence the T24/TLT amplification ratio and T24 - TLT trend differences as demonstrated in Table 4.1. Across the datasets, the spread in the precipitation-weighted trends is larger than the average SST trends, which amplifies the differences in T24 compared to TLT. For example, going from the Hurrell dataset to the NMAT dataset the average SST trend goes down by 0.033 K decade^{-1}, but the precipitation-weighted trend is reduced by 0.054 K decade^{-1} (150% larger trend reduction). Since the precipitation-weighted SST trends have more influence on the upper troposphere (T24) and the average SSTs have a larger effect on the lower troposphere (TLT), this causes a differential effect with T24 trends reduced more than TLT trends. As a result the T24/TLT

trend ratio and differences are reduced with the HadISST dataset and even further reduced using the NMAT data.

In addition to the prescribed SST experiments in CAM4, we also consider the CMIP5 results from Chapter 3. In Table 4.2, we show the T24/TLT amplification ratio for the coupled and atmosphere-only version of a number of CMIP5 models. From Table 4.2, we see that 11 of the 15 models have a greater amplification ratio in the AMIP experiment compared to the CMIP experiment and the multi-model average ratio is larger in the AMIP experiment. This likely occurs because, over the time period considered, the SST trends were largest over the Western Hemisphere Warm Pool, which acts to increase the T24/TLT trend ratio in prescribed SST (AMIP) experiments. As we show in Section 4.5.2, coupled models typically have spatially uniform increases in SST.

From these prescribed SST experiments and the comparison of CMIP5 atmosphere-only and coupled models, we see that the spatial pattern of warming does have some influence on the model amplification. We also note that several of the datasets in Po-Chedley and Fu [2012a] that had amplification ratios close to the satellite-derived values were forced with the HadISST dataset rather than the Hurrell dataset (e.g., NASA GISS and GFDL models). We conclude that amplification is sensitive to the pattern of SST trends and observational biases in SST datasets can be important. In Section 4.4.2, we show that model physics is even more important than the pattern of SST changes.

4.4.2 The role of model physics

Another possibility for differences between observations and models is that model parameterizations for convection may not accurately simulate sub-grid scale processes. To understand the role that parameterizations for convection may play in GCM amplification behavior, we examine model experiments from the GISS E2 model.

A version of the GISS E2 model was updated so that the model would have a more realistic distribution of clouds and moisture and would be able to simulate a realistic MJO [Del Genio et al., 2012, Kim et al., 2012, Yao and Cheng, 2012, Field et al., 2014]. Underlying changes to

Model	AMIP	CMIP
BCC-CSM1.1	1.16	1.16
CanAM4/CanCM4	1.24	1.19
CCSM4	1.24	1.20
CNRM-CM5	1.14	1.09
FGOALS-s2	1.21	1.16
GISS-E2-R	1.13	1.19
HadGEM2-A/HadGEM2-AO	1.15	1.16
INM-CM4	1.21	1.20
IPSL-CM5A-LR	1.23	1.21
IPSL-CM5B-LR	1.22	1.24
MIROC5	1.23	1.20
MPI-ESM-LR	1.20	1.17
MPI-ESM-MR	1.19	1.18
MRI-CGCM3	1.22	1.18
NorESM1-M	1.22	1.18
Model Average	1.20	1.18

Table 4.2: The ratio of the T24 trend divided by the TLT trend for each model (ensemble average). The trends are for the tropics (20°S-20°N) and we include the values for both the atmosphere-only models (AMIP experiment) and the coupled models (historical CMIP experiment). When the atmosphere-only and coupled model have different names we display both. Values are reproduced from Tables 3.1 - 3.2.

the model convection parameterization effectively increased the fractional entrainment rate, which decreased the depth of deep convection in the model [Kim et al., 2012]. We refer to this experiment as GISS MJO.

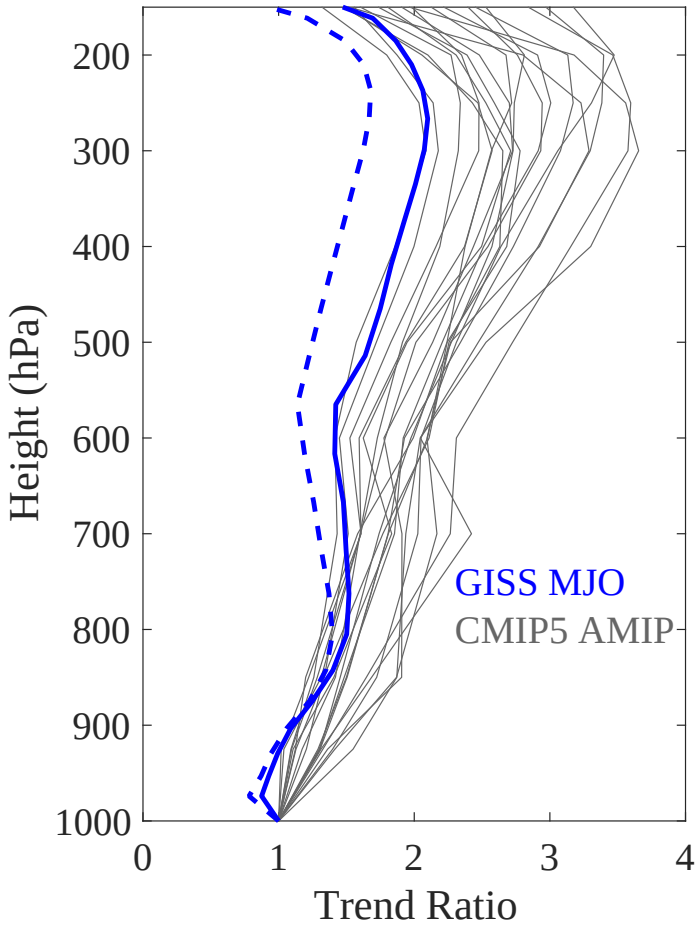

Figure 4.8: Average tropical (20°N-S) average tropospheric trends (1981 - 2008) normalized by the surface temperature trend in CMIP5 models (gray) and in the GISS MJO experiments (blue). The solid blue line represents the control and the dashed blue line represents the response with updated physical parameterizations.

In Fig. 4.8, we show the GISS MJO model responses to warming over 1981 - 2008. We also show the model warming in the first ensemble member for 19 CMIP5 AMIP model simulations for reference (gray lines). In the updated version of the GISS model (dashed blue line) there is clearly less upper tropospheric warming compared to the control simulation (solid blue line). In Table 4.1 we see that the updated GISS model has reduced vertical amplification relative to the control experiment, and is in close agreement to the RSS observational estimates (it is also closer to the UAH observational estimates). This indicates that

if observational estimates of vertical warming amplification can be trusted, biases in model parameterizations may play a role in reconciling existing discrepancies. We also note that the GISS model had a relatively small amplification ratio prior to parameterization changes, which suggests that the HadISST SST dataset that drove the model (instead of the Hurrell dataset) may also be important [Po-Chedley and Fu, 2012a].

4.4.3 The role of observational uncertainty

An alternative explanation for model simulations that are inconsistent with the observational record is that the observations contain unresolved biases. Studies that have explored differences in vertical amplification between models and observations have shown that homogenized datasets created by independent research groups consistently show that the upper troposphere appears to be warming too quickly in models (relative to the lower troposphere) [Fu et al., 2011, Po-Chedley and Fu, 2012a, Seidel et al., 2012]. Even though the rate of upper (T24) and lower (TLT) tropospheric warming differs in the RSS and UAH data, the ratio of T24 and TLT warming is consistent across both datasets. In this way, both RSS and UAH data are in close agreement, suggesting that the ratio of upper and lower tropospheric warming in models diverges from observations. Despite similar T24/TLT warming ratios in RSS and UAH data, observational uncertainty in this amplification metric is not negligible. The RSS v4.0 upper tropospheric T24 trend has increased dramatically compared to v3.3 (Table 4.1), which is a demonstration that observational uncertainty is substantial and that the amplification ratio may be larger (it is not known because RSS has not yet produced a v4.0 TLT trend). UAH dataset updates (v6.0 beta 5 - see Table 5.1) have reduced the UAH T24/TLT trend ratio, which also suggests this metric is subject to considerable uncertainty [Po-Chedley and Fu, 2012b, 2013, Po-Chedley et al., 2015].

In order to consider the role of observational uncertainty in model-observational mismatches, we utilize an ensemble of realizations of the RSS MSU/AMSU dataset [Mears et al., 2011]. This ensemble uses a Monte Carlo technique to iteratively perturb bias corrections that result from sampling, instrument calibration, and diurnal drift. This produces

400 plausible temperature time series, which can be used to assess observational uncertainty. For example, using this ensemble we can estimate the mean trend and uncertainty ($\pm 2\sigma$) for the TLT (lower troposphere), TMT (mid-troposphere), TLS (lower stratosphere), and T24 (full troposphere) channels in the tropics over 1981 - 2008: 0.142 ± 0.046, 0.117 ± 0.053, -0.315 ± 0.067, and 0.160 ± 0.058 K decade^{-1}, respectively. In estimating the T24 trends, we assumed that the TMT and TLS ensemble members are independent of one another, though this only increases the uncertainty estimate slightly (from 0.052 - 0.058 K decade^{-1}).

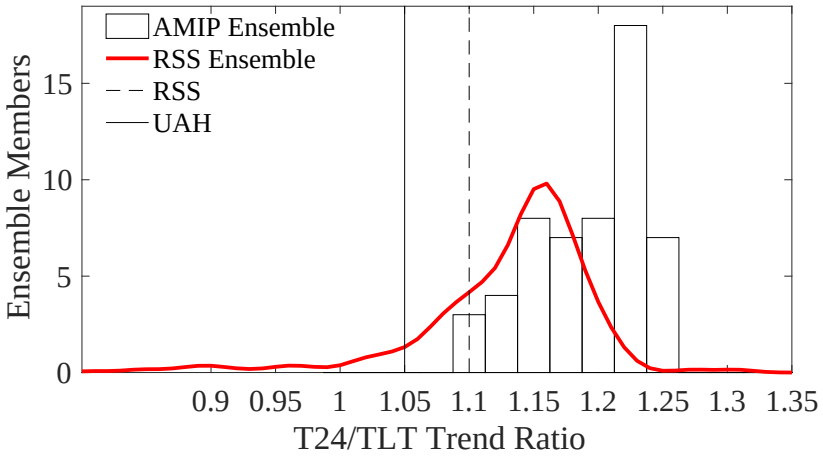

Figure 4.9: Probability Distribution Function of RSS T24/TLT tropical (20°N-S) trend (1981 - 2008) amplification ratios (red line) and a histogram of model ensemble member amplification ratios for AMIP experiments (from Po-Chedley and Fu [2012a]). We also plot vertical lines to represent the amplification ratios for UAH v5.4 and RSS v.3.3.

In Fig. 4.9, we show the probability distribution of T24/TLT trend ratios from the RSS ensemble. For comparison, we have included the histogram of model ensemble member amplification ratios from prescribed SST (AMIP) experiments. Note that the amplification ratio from the official RSS dataset is at the lower end of the ensemble with a ratio of 1.10 over 1981 - 2008. The median value in the RSS ensemble is 1.14 and the mean is 1.13. This difference occurs because the official dataset is just one possible temperature evolution and is not representative of the mean or median of the RSS ensemble. This figure helps put model observational amplification discrepancies into perspective in that the bulk of the histogram is

consistent with a number of model ensemble members. The probability of an RSS T24/TLT trend ratio above 1.21 is less than 5%, which indicates that a number of models (45% of the CMIP5 AMIP ensemble members) are still inconsistent with observations over this time period. Similarly, the RSS data is inconsistent with the newer UAH dataset; the probability of an RSS amplification ratio below 1 is also less than 5%.

We also consider the stability of satellite-derived amplification in Fig. 4.10. In previous assessments models and observations were compared over 26 to 31 year periods [Fu et al., 2011, Po-Chedley and Fu, 2012a], which corresponded to the maximum time periods in which data for both models and observations were available. In extending the ratio of T24 and TLT trends through 2015, we see that the RSS amplification ratio has increased substantially and is now 1.19 (1979 - 2015). The UAH v6.0b5 time series has a similar trend and interannual variability compared to the RSS time series, but the entire time series is offset by $\sim$-0.1. This drift in the the amplification ratio suggests that the observational record may have unresolved biases and that observations may be concordant with AMIP model experiments (the CMIP5 AMIP model mean amplification ratio is 1.19) [Santer et al., 2016]. In Fig. 4.10 we also show the GISS control experiment (blue solid line) and the GISS MJO experiment with perturbed physics (blue dashed line). The control simulation has a very stable amplification ratio over 1997 - 2011. In contrast, the GISS MJO simulation (with an updated convection parameterization) tracks the interannual variability in the RSS time series quite well. Given that the improved model simulation reproduces many of the features in the observational record, it is more likely that observational changes in the amplification between the upper and lower troposphere are not an artifact. It is difficult to interpret this information as being squarely in favor of either model or observational amplification behavior, though extending the GISS MJO time series forward may be helpful in understanding the variability in amplification (and perhaps shedding light on the merit of the model's convection scheme).

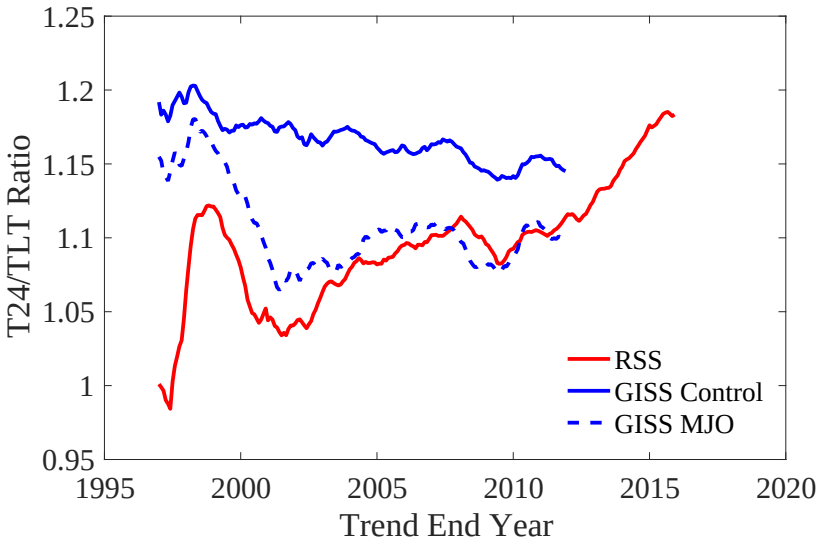

Figure 4.10: Tropical (20°N-S) trend ratios for T24 versus TLT as a function of time (starting in 1981). We show the ratios for the RSS dataset (red), the GISS control simulation (from Section 4.4, blue solid line), and the modified GISS ensemble member (blue dashed line).

In all, even when observational uncertainty is considered, it appears that many models exhibit increases in dry static stability that are larger than observations support. Importantly, microwave temperature datasets are still being updated [Mears and Wentz, 2016] and critically assessed [Po-Chedley and Fu, 2012b, Christy and Spencer, 2013, Po-Chedley and Fu, 2013, Po-Chedley et al., 2015] more than 25 years after the first attempt at intersatellite homogenization [Spencer and Christy, 1990]. Important issues remain, including puzzling inconsistencies between MSU and AMSU records [Po-Chedley et al., 2015]. Updated datasets from RSS and UAH have also diverged in their estimates of amplification, which deserves more scrutiny and may imply that observational uncertainty is larger than what is shown in this assessment.

4.5 Vertical amplification and the lapse rate feedback

A principal motivation for investigating the structure of warming in the tropics is that the vertical pattern of warming controls a fundamental climate feedback, the lapse rate feedback. In this section, we seek to understand how satellite metrics of warming relate to this feedback and attempt to shed some light on what controls the magnitude of this feedback.

4.5.1 Satellite metrics for the lapse rate feedback

MSU/AMSU observations measure microwave emissions from oxygen that correspond to temperature over deep atmospheric layers. Layers of interest include the upper troposphere, represented by T24, and the lower troposphere, represented by TLT. As shown in Fig. 4.1, T24 is also influenced by the lower troposphere and TLT includes some information from the upper troposphere. The lapse rate feedback represents the radiative response to non-uniform vertical warming. To measure this feedback, detailed information is needed about the structure of tropospheric temperature changes. Given that satellite-borne microwave measurements represent broad atmospheric layers, we assess the degree to which these measures can be used to assess the lapse rate feedback.

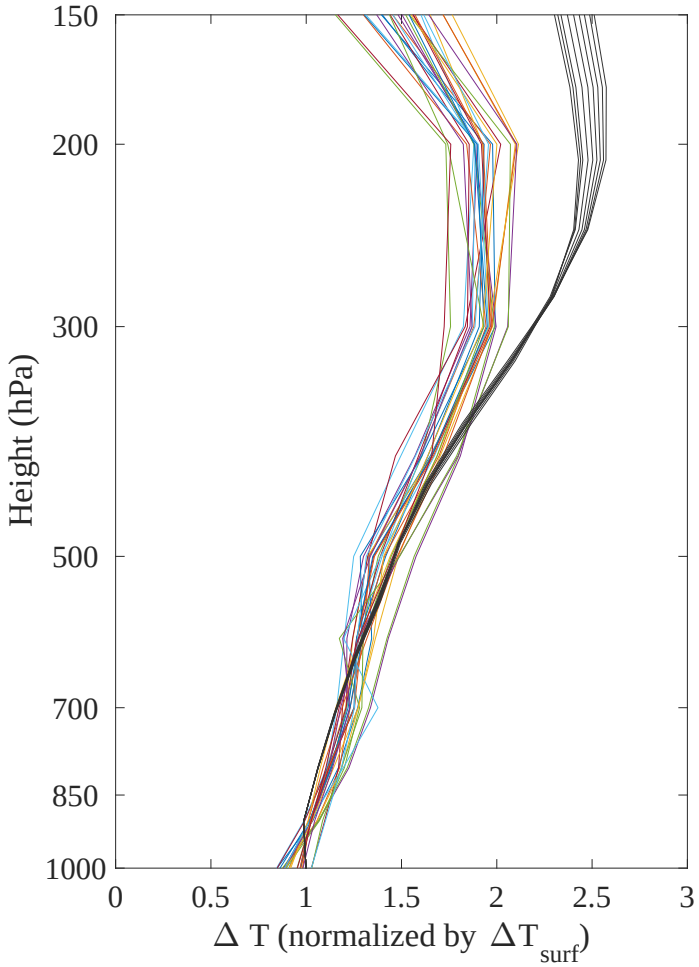

Figure 4.11: CMIP5 model tropical (20°N-S) tropospheric warming profiles from CO_2 quadrupling experiments (colored lines). The black lines show moist adiabatic warming profiles above 850hPa (assumed as dry adiabatic below 850hPa with a surface temperature of 298 K and surface warming ranging from 2 - 6 K). In both cases the trend profiles are normalized by the surface temperature change.

In Fig. 4.11 we show tropical warming profiles for 28 coupled atmosphere ocean models from the CMIP5 archive (colored lines). These models were forced with an abrupt quadrupling of CO_2 and the profiles represent the average warming 120 - 150 years after quadrupling (for more details see Section 5.3). Since models have varied rates of warming, we normalized the profiles by the surface temperature change. In Fig. 4.11 we also show profiles of temperature change assuming that warming follows a moist adiabat (black lines). The profiles represent surface warming from 2 to 6 K. Moist adiabatic warming is a good approximation for much of the troposphere, though it tends to overestimate model warming in the upper tropical troposphere [Santer et al., 2005]. Enhanced warming in the upper troposphere relative to the lower troposphere leads to a negative lapse rate feedback (and a ratio of T24 to TLT warming of ∼1.2 as shown in Table 4.1).

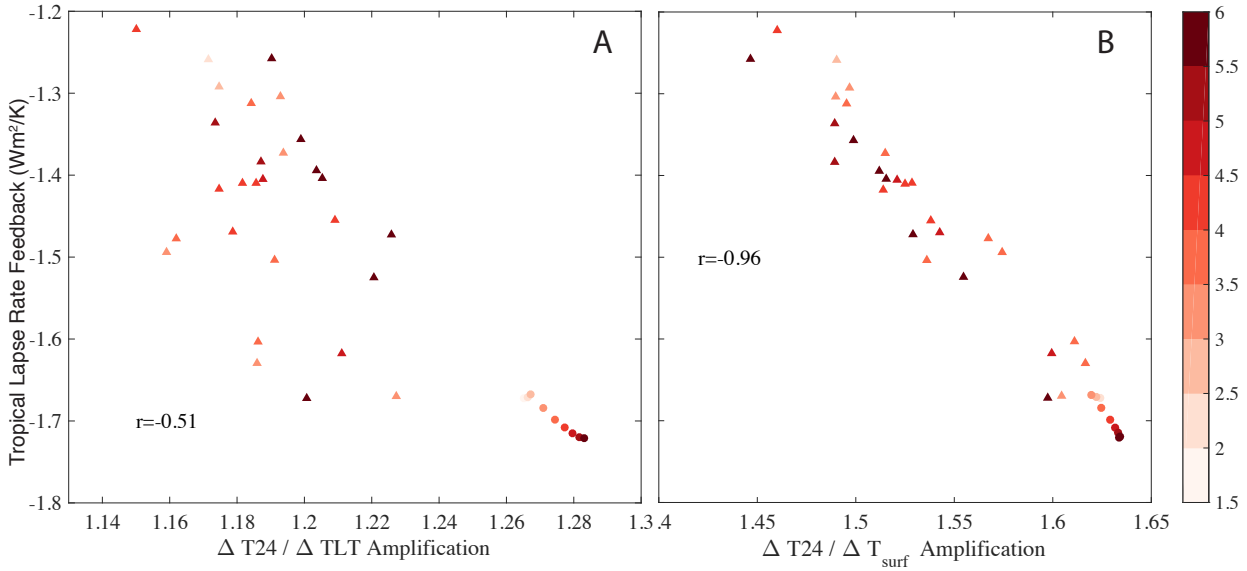

Figure 4.12: Regression between the tropical (20°N-S) lapse rate feedback versus A) the ratio of T24 warming relative to TLT warming and B) the ratio of T24 warming relative to surface warming. This is shown for CMIP5 models (triangles) and moist adiabatic scaling (circles). The r-values correspond only to the linear fit for CMIP5 model values. The colors represents the magnitude of surface warming.

In Fig. 4.12A we show the relationship between the tropical T24/TLT warming ratio and the tropical lapse rate feedback. While moist adiabatic warming tends to increase stability (i.e., larger T24/TLT values) with enhanced surface warming (represented by the color of the dots), GCMs do not exhibit the same behavior, demonstrating that model changes do not purely reflect a moist adiabatic response. Even though the T24/TLT metric is intended to assess dry static stability changes between the upper and lower troposphere, it does not closely scale with the lapse rate feedback. In contrast, we show the ratio of T24 warming relative to surface warming versus the tropical lapse rate feedback in Fig. 4.12B, which is a much better measure of the tropical lapse rate feedback. Upper tropospheric (i.e., T24) warming relative to the surface projects strongly onto the lapse rate feedback, because top-of-atmosphere (TOA) longwave radiative fluxes are strongly modulated by temperature changes near 300 hPa [Soden and Held, 2006], which also corresponds to the peak in the T24 weighting function. Furthermore, in computing the lapse rate feedback these TOA fluxes are normalized by surface temperature change. As a result, the relative warming in T24 and at the surface are excellent proxies for the lapse rate feedback. T24/TLT warming ratios are attractive in that they are derived independent of other datasets and this ratio may still be a good metric for the bulk behavior of stability changes in the tropical troposphere, but it is an inferior proxy for the lapse rate feedback.

4.5.2 Factors influencing the tropical lapse rate feedback

In Section 4.4, we showed that tropical SST warming over the warmest waters (as indexed by precipitation-weighted SSTs) is an excellent predictor of the interannual variability in T24 [see Fig. 4.4, Sobel et al., 2002, Flannaghan et al., 2014]. Given that T24 is an important component of the lapse rate feedback, we investigate the pattern of SST warming in models responding to CO_2 quadrupling. Consider a model with preferential warming over convective (warm SSTs) regions and reduced warming elsewhere. A model with this hypothetical behavior would also produce larger T24 trends, even if the average surface temperature change is the same as uniform SST warming. In turn, the pattern of SST warming, which influences

the ratio of T24 and surface warming, would modulate the strength of the lapse rate feedback (see Fig. 4.12). Across CMIP5 GCMs, such warming discrepancies between the warmest and average SSTs are small. In Fig. 4.13A we show the changes in SST as a function of the climatological SST percentile (as in Fig. 4.6). Models warm the tropical sea-surface relatively uniformly, which indicates that changes in the precipitation-weighted SSTs scale with changes in the average tropical SST. In Fig. 4.13B, we see that this is indeed the case – the average tropical SST increases nearly identically to the precipitation-weighted SST. Tropical T24 warming across models, in turn, scale quite closely with both tropical average SST warming and precipitation-weighted SST warming (r=0.99 in both cases). Since models warm the tropical SST uniformly, the precipitation-weighted SST is not a strong indicator of the tropical lapse rate feedback (r=-0.24).

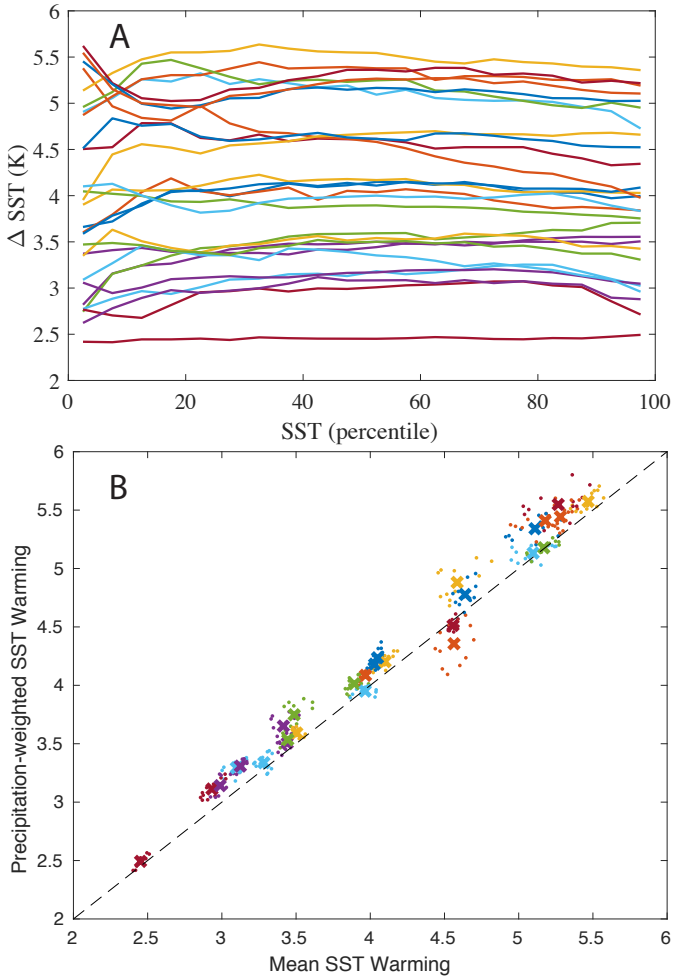

Figure 4.13: A) Changes in tropical (20°N-S) SSTs for different climatological SSTs (in percentile space) with smaller percentiles representing cooler waters and larger percentiles representing warmer waters. B) Regression of the precipitation-weighted SST warming versus the average SST warming over the tropics. Each color represents a different model with individual months represented by small dots and the large X representing the annual average. In both (A) and (B), the temperature changes are derived from model experiments quadrupling CO_2.

While the pattern of SST changes does not appear to be a factor in shaping the tropical lapse rate feedback, differences in model parameterizations for microphysics and convection can modulate the strength of the lapse rate feedback. In Fig. 4.14, we reproduce the relationship between the tropical lapse rate feedback and the T24/TLT amplification ratio for CMIP5 models (black dots). We also show a perturbed physics ensemble (based on a developmental GFDL model) with changes to the microphysics that influences the conversion of cloud water to rain (blue squares) [Zhao et al., 2016]. This model is perturbed such that the model's climatological mean state in each ensemble member is realistic and none of the ensemble's parameterizations can be ruled out based on existing understanding of cloud

microphysics. Surprisingly, the lapse rate feedback in this set of 9 ensemble members varies widely, spanning roughly half the range of lapse rate feedbacks seen across the entire CMIP5 archive. This demonstrates that decisions in formulation model physical parameterizations has a tremendous effect on differences in the lapse rate feedback across models. We reproduce this behavior in CAM4 by altering the deep convection parameterization (red triangles). In each ensemble member, we modify either the timescale for deep convection or the lateral entrainment rate. Although the lapse rate feedback in CAM4 is less sensitive to convective parameterizations compared to the microphysical changes made in the GFDL model, the CAM4 ensemble members still span roughly a quarter of the range of CMIP5 model tropical lapse rate feedbacks.

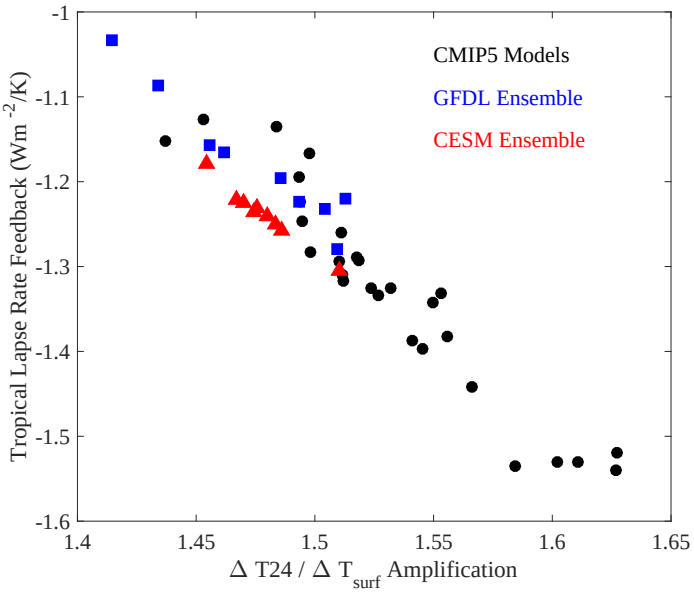

Figure 4.14: The tropical (30°N-S) lapse rate feedback versus the ratio of T24 warming relative to surface warming for CMIP5 GCMs (black dots), GFDL perturbed physics ensemble members (blue squares), and CESM perturbed physics ensemble members (red triangles).

Since some of the spread in the tropical lapse rate feedback appears to be inherent to model physics, we test whether transient vertical amplification rates across models scale with amplification in larger forcing experiments. In Fig. 4.15A we show the relationship between each model's T24/TLT warming ratio for CO_2 quadrupling versus the ratio for historical warming experiments. Figure 4.15B is similar, except that it is for the T24/$T_{surface}$ warming ratio. In each case, the correlation is modest, which demonstrates that differences in model

amplification in transient warming has limited predictive power of amplification in longer term and larger forcing scenarios. We also found that differences in interannual variability of amplification across models is also unable to predict long-term amplification behavior. Given that the lapse rate feedback is well defined by the amplification between the surface and T24, this result implies that MSU/AMSU observations may provide limited insight into long-term tropical lapse rate feedback behavior. We also note that although the tropical lapse rate is strongly negative across models in the tropics, the tropical lapse rate feedback only explains less than half of the variance (r^2=0.48) in the global lapse rate feedback across models.

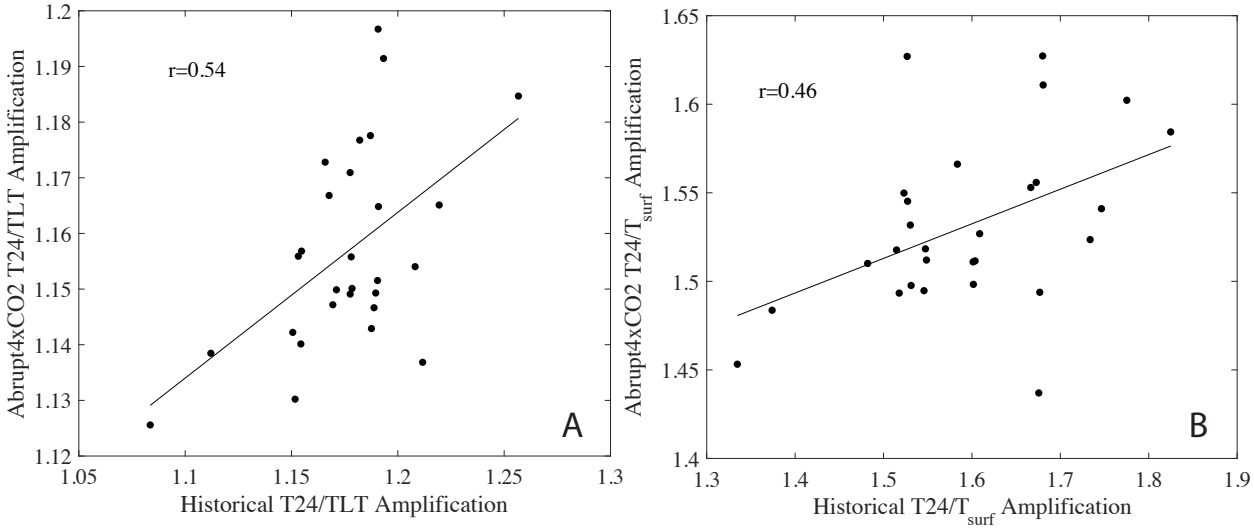

Figure 4.15: Regression of the ratio of tropical (30°N-S) warming for CO_2 quadrupling (120 years after abrupt CO_2 quadrupling) versus the ratio of warming in historical (1979 - 2005) experiments for A) the ratio of T24 to TLT trends and B) the ratio of T24 to surface temperature trends.

4.6 Conclusion

This chapter assessed changes in tropical mid-tropospheric dry static stability in GCMs and observations using the ratio of upper (T24) and lower (TLT) tropospheric trends. Specifically, we sought to understand mechanisms which might contribute to discrepancies between

models and observations in the vertical profile of tropospheric warming. We found that although the model forcing for ozone and stratospheric aerosols may have deficiencies, tropical tropospheric amplification in models is largely insensitive to changes in these atmospheric constituents. However, discordant rates of stratospheric cooling in GCM simulations and observations may be reconciled with improved treatment of ozone forcing in models. Model amplification behavior was influenced by the pattern of SST warming and reducing uncertainties in SST trends may help improve model simulations of the vertical pattern of warming. We note, however, that the dataset that best matched the observational record was based on ship measurements of nighttime marine air temperature, which may not be realistic. Taking observational uncertainty into account demonstrates that a number of models have amplification ratios that are consistent with observations, though approximately half of the CMIP5 AMIP ensemble members have amplification ratios that are outside of the observational range. Even still, it is possible that unresolved biases in the microwave record may contribute to the remaining model-observational disagreement in tropical tropospheric amplification.

Although model-observational discrepancies in the T24/TLT trend ratio may be important, the T24/TLT trend ratio does not strongly co-vary with the tropical lapse rate feedback. We find that an improved predictor for the lapse rate feedback is the ratio of T24 warming relative to near surface warming. Using two perturbed physics model ensembles, we show that the formulation of model parameterizations for deep convection and microphysics are critical in shaping model vertical amplification and the lapse rate feedback. Unfortunately, model simulations of amplification over the historical period are not strongly predictive of their behavior in quasi-equilibrium warming. Furthermore, SSTs in coupled atmosphere-ocean models largely warm uniformly, so the pattern of SST changes does not have a strong influence on changes in static stability.

This chapter was unable to reconcile differences between all model-based and observational estimates of tropical tropospheric amplification, but this work points to a number of promising approaches. Given that ensembles of SST reconstructions now exist [Morice et al.,

2012], it would be possible to quantify the uncertainty in model amplification due to differences in the pattern of SST warming by forcing an atmospheric model with the full range of observational SST realizations. The ability of the updated GISS model (MJO) to reproduce the variability in the T24/TLT ratio over time also suggests that this sort of analysis may be useful in discriminating bulk behavior of model convection and microphysics schemes, assuming observations of this behavior can be trusted. Regardless, this does suggest that tropospheric amplification is sensitive to interannual and decadal variability. As a result, existing observations of the profile of warming may not be representative of the long-term climate response, as demonstrated in comparing model transient and quasi-equilibrium behavior. While ozone plays little role in tropospheric temperature changes, it has a large effect in the lower stratosphere; updating model ozone forcing as suggested in several studies [e.g., Bodeker et al., 2013, Damadeo et al., 2014] may help reconcile trend differences noted in other studies [Thompson et al., 2012]. Discrepancies noted here should be revisited as microwave observations are updated in the near-future, because some of the datasets have undergone relatively large changes in order to remove biases related to diurnal sampling [Mears and Wentz, 2016].

In this thesis, two measures of amplification have been employed: the ratio of T24 and surface temperature trends (Chapter 2) and the ratio of T24 and TLT trends (Chapters 3 – 4). The T24/$T_{surface}$ trend ratio is a bulk indicator of amplification, expressing the degree to which the full troposphere warms relative to the surface. In Chapter 2, we were able to show that observations and models agree on the magnitude of T24/$T_{surface}$ amplification. The T24/TLT trend ratio expresses the amplification between the lower-middle and upper troposphere, which can be thought of as the vertical extent of vertical amplification. We showed in Chapters 3 – 4 that many models overestimate this ratio, though accounting for all observational uncertainty in the satellite record, improving SST datasets used to force models, and improving model convection and microphysics schemes may reduce model-observational discrepancies.

Chapter 5

ON THE INTERMODEL SPREAD OF THE GLOBAL LAPSE RATE FEEDBACK

5.1 Abstract

The lapse rate feedback is the Earth's largest negative (non-Planck) climate feedback and is poorly constrained across general circulation models. Understanding the spread in this temperature feedback is important for understanding aspects of the Earth's response to increased greenhouse gases. Model differences in the local lapse rate feedback are particularly large in the southern extratropics, and the feedback in this region is influenced by both tropical and Antarctic warming. Greater tropical surface warming across models is associated with greater warming in the extratropical troposphere producing a negative feedback in the southern extratropics because surface warming is muted over the Southern Ocean. Larger surface warming in the southern hemisphere extratropics can overwhelm the non-local influence of the tropics to create a positive local lapse rate feedback. The behavior of the southern hemisphere extratropical lapse rate feedback, as it spans from a strong negative to weak positive feedback, is a critical component of the global average lapse rate feedback. The balance of tropical and polar warming is thus not only important for its role in partitioning low latitude (negative) and high latitude (positive) lapse rate feedbacks, but also because the warming pattern shapes the meridional pattern of local feedbacks. Part of the model spread in the southern hemisphere meridional warming pattern is traced to differences in climatological Antarctic sea ice extent, which in turn contributes to as much as 30% of the variability in the global average lapse rate feedback even though Antarctic sea ice covers less than 4% of the Earth's surface.

5.2 Introduction

The Earth's response to an increase in carbon dioxide (CO_2) is modulated by feedback processes that damp or amplify the initial greenhouse gas induced radiative forcing. These feedbacks include the albedo, water vapor, lapse rate, Planck, and cloud feedbacks [e.g., Soden and Held, 2006]. The current generation of general circulation models (GCMs) exhibit a wide range of feedback strengths, particularly for the shortwave cloud and lapse rate feedback [Caldwell et al., 2016]. Understanding processes that control the spread in individual feedback processes is important, because climate feedbacks ultimately control the spread in equilibrium climate sensitivity (ECS) – the Earth's global temperature response to a doubling of the concentration of atmospheric CO_2 at equilibrium [e.g., Colman, 2003]. Although a number of studies have demonstrated that the shortwave cloud feedback contributes the most toward the spread in ECS [e.g., Cess et al., 1990, 1996, Soden and Held, 2006, Soden et al., 2008, Dufresne and Bony, 2008, Caldwell et al., 2016], other feedbacks also contribute to inter-model differences in ECS [Crook et al., 2011, Caldwell et al., 2016] and are also important to the magnitude and timing of regional warming [Armour et al., 2013] and the response of the atmospheric circulation to warming [Hwang and Frierson, 2010, Feldl and Bordoni, 2016].

In this analysis, we focus on a specific aspect of the Earth's temperature feedback. The temperature feedback can be decomposed into the radiative response due to uniform warming with height, the Planck feedback, and the response due to non-uniform warming with height, the lapse rate feedback. We focus on the latter feedback in this work. In the tropics, models and observations robustly show enhanced warming aloft compared to the surface [Manabe and Wetherald, 1975, Fu et al., 2004, Mitchell et al., 2013a, Po-Chedley et al., 2015, Sherwood and Nishant, 2015]. This vertical warming amplification results because tropical temperature changes resembles moist adiabatic warming [Santer et al., 2005]. As a result of vertical warming amplification in the tropics, the lapse rate is reduced, leading to a large negative radiative feedback. In the high latitudes, surface warming can exceed tropospheric warming,

which results in an increase in the lapse rate and a positive lapse rate feedback [Ramanathan, 1977]. This bottom-heavy warming is a product of stable atmospheric conditions that inhibits surface warming from being communicated to the upper troposphere [Payne et al., 2015]. It has been demonstrated that the global lapse rate is controlled by the meridional pattern of surface warming [Soden and Held, 2006]. Soden and Held [2006] explain that:

> The range in lapse rate feedbacks between models, in turn, stems from different meridional patterns of surface warming. Models with relatively larger warming at low latitudes have a greater reduction in lapse rate and thus a larger (more negative) lapse rate feedback. This behavior reflects the weaker coupling of the surface to the free troposphere at high latitudes compared to low latitudes where the model-simulated temperature response closely follows a moist adiabat regardless of the convection scheme [Santer et al., 2005].

Therefore, models with relatively enhanced tropical surface warming (compared to the global surface temperature change) tend to have a stronger (negative) lapse rate feedback. Soden and Held [2006] show that an excellent predictor of the lapse rate feedback is the ratio of tropical to global surface warming. On a global scale, the lapse rate feedback is negative in all of the most recent generation of coupled atmosphere-ocean models due to the large negative feedback throughout the tropics and into the mid-latitudes [Soden and Held, 2006, Bony et al., 2006, Caldwell et al., 2016].

Although the spread of the global lapse rate feedback across general circulation models is relatively large [e.g., Caldwell et al., 2016], it is strongly and inversely correlated with the water vapor feedback [Cess, 1975, Soden and Held, 2006]. Since models warm without large changes in relative humidity, tropospheric warming that produces a negative lapse rate feedback is also associated with an increases in tropospheric water vapor, which induces a positive radiative feedback. As a result of this physical connection, the water vapor and lapse rate feedbacks are typically combined [Held and Soden, 2000]. The sum of these two feedbacks is robustly positive, and the variance is much smaller than that of either feedback

alone [Colman, 2003, Soden and Held, 2006, Caldwell et al., 2016]. While understanding inter-model variability in the lapse rate feedback may not help explain the large spread in ECS because the lapse rate and water vapor feedback mostly cancel one another, the lapse rate feedback is important in other ways. The lapse rate feedback has been highlighted as an important contributor to polar amplification [Winton, 2006, Pithan and Mauritsen, 2014, Roe et al., 2015], can amplify the albedo feedback that is controlled by snow cover and sea ice loss [Ramanathan, 1977, Feldl et al., 2016], influences atmospheric heat transport [Feldl and Bordoni, 2016], and has acted as a globally-averaged positive feedback in past climates [Atwood et al., 2016]. In the tropics, the vertical structure of atmospheric warming can be important in determining the change in the Brewer-Dobson circulation [Lin and Fu, 2013], the height of anvil clouds [Hartmann and Larson, 2002], and to changes in tropical cyclone activity [Vecchi et al., 2013].

This work aims to understand controls on the magnitude of the lapse rate feedback across an ensemble of GCMs. In Section 5.3 we describe the data used in this analysis and our methodology for calculating the lapse rate feedback for each GCM. In Section 5.4 we look at the behavior of the multi-model mean lapse rate feedback and how regional modes of variability in the local lapse rate feedback across models correspond to model differences in the global mean lapse rate feedback. We show that model differences in the southern mid and high latitudes seem to be particularly important in understanding discrepancies in the global lapse rate feedback. Therefore, in section 5.5, we attempt to understand local controls on the lapse rate feedback in the southern hemisphere. We summarize our results in Section 5.6.

5.3 Data and methods

We focus on interpreting GCM data from the Coupled Model Intercomparison Project Phase 5 (CMIP5) [Taylor et al., 2012]. In particular, we use fields from the pre-industrial control experiment (piControl), which contains century to millennial scale model runs with unvarying pre-industrial atmospheric CO_2 concentrations. We consider data from each model's

first ensemble member (denoted "r1i1p1"). To evaluate each model's response to increased greenhouse gases, we use experiments in which models are forced with an abrupt quadrupling of atmospheric CO_2 (abrupt4xCO2). We consider data from 28 coupled atmosphere-ocean GCMs using climatologies from the last 100 years of the piControl simulations and years 120 - 149 after quadrupling of CO_2 in the abrupt4xCO2 experiments. The models used in this study are the same as those in Caldwell et al. [2016] and are listed in Table 5.1. We also use sea ice data from the same 28 models in a transient forcing simulation in which the atmospheric concentration of CO_2 is increased by 1% each year (1pctCO2). To assist in comparing model sea ice fields with present-day estimates of sea ice cover, we employ monthly satellite passive microwave sea ice concentration observations over 1979 - 2014 [Meier et al., 2013, Peng et al., 2013]. In a number of instances, we use the model sea ice extent as an index with which to compare models; we calculate this value by summing the product of grid cell area multiplied by the grid cell average sea ice concentration.

In order to calculate the lapse rate feedback in models, we used the difference in atmospheric temperature between the piControl and the abrupt4xCO2 climatologies. The feedback can be calculated using radiative kernels [e.g., Soden and Held, 2006, Shell et al., 2008, Caldwell et al., 2016], which represent the top-of-atmosphere (TOA) radiative response due to changes in temperature, water vapor, and albedo (e.g., $\partial R/\partial i$, where R is the TOA radiation change and i is the variable under consideration). For temperature, which is three-dimensional, the radiative response at each model gridpoint is an integral over the atmospheric column:

$$\Delta R_{temperature}(r) = \int_{p_s}^{p_t} \frac{\partial R(r,p)}{\partial T(r,p)} \cdot \Delta T(r,p) dp \tag{5.1}$$

where $\Delta R_{temperature}$ is the TOA radiative response due to temperature changes, r is the spatial point (in the horizontal direction), p is pressure, $\partial R(r,p)/\partial T(r,p)$ is the radiative kernel for temperature, and ΔT is the temperature response after quadrupling CO_2. We integrate from the surface, p_s, to the tropopause, p_t, following Reichler et al. [2003] to determine the tropopause pressure. This can be decomposed into the Planck response (Eq.

5.2), which results from uniform warming with height, and the lapse rate response, which is the TOA radiative result of deviations from uniform warming (Eq. 5.3):

$$\Delta R_{Pl}(r) = \int_{p_s}^{p_t} \frac{\partial R(r,p)}{\partial T(r,p)} \cdot \Delta T_{surface}(r) dp \tag{5.2}$$

$$\Delta R_{lr}(r) = \int_{p_s}^{p_t} \frac{\partial R(r,p)}{\partial T(r,p)} \cdot \Delta (T(r,p) - T_{surface}(r)) dp \tag{5.3}$$

We use all-sky radiative kernels with monthly climatological resolution from Shell et al. [2008] to calculate the radiative response due to lapse rate changes at each grid point for each of the 28 GCMs considered in this work. The lapse rate feedback at any point is typically calculated by dividing the radiative response due to lapse rate changes, $\Delta R_{lr}(r)$, by the global near-surface temperature change, $\overline{\Delta T_{surf}}$. The global average lapse rate feedback is defined as:

$$\overline{\lambda_{lr}} = \frac{\overline{\Delta R_{lr}}}{\overline{\Delta T_{surf}}} \tag{5.4}$$

where $\overline{\Delta R_{lr}}$ is the global average radiative change due to lapse rate changes. An alternative formulation of this feedback is to divide the radiative response due to lapse rate changes by the zonal temperature change to estimate the local lapse rate feedback at each point [Feldl and Roe, 2013, Armour et al., 2013, Feldl and Bordoni, 2016]:

$$\lambda_{lr,local}(r) = \frac{\Delta R_{lr}(r)}{\Delta T_{surf}(\phi)} \tag{5.5}$$

where ϕ is latitude and we denote the zonal average surface temperature change as $\Delta T_{surf}(\phi)$. The advantage of this metric is that it emphasizes the connection between the local temperature change and the local radiative response [Feldl and Roe, 2013]. In general, we will refer to the feedback as the "local" lapse rate feedback when the radiative change is divided by the zonal temperature change; when we divide by the global mean temperature change we will refer to this simply as the lapse rate feedback. The global average lapse rate feedback can also be calculated using the local lapse rate feedback definition (from Eq. 5.4 - 5.5):

$$\overline{\lambda_{lr}} = \frac{\int_{-\pi/2}^{\pi/2} \lambda_{lr,local}(\phi) \cdot \Delta T_{surf}(\phi) \cdot cos(\phi) d\phi}{\int_{-\pi/2}^{\pi/2} \Delta T_{surf}(\phi) \cdot cos(\phi) d\phi} \tag{5.6}$$

where the numerator is simply the global average radiative change due to changes in the lapse rate, $\overline{\Delta R_{lr}}$, and the denominator is the global mean surface warming, $\overline{\Delta T_{surf}}$, as in Eq. 5.4. Here we have used the zonal average local lapse rate feedback, $\lambda_{lr,local}(\phi)$.

Group	Model
Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia	ACCESS1-0 ACCESS1-3
College of Global Change and Earth System Science, Beijing Normal University	BNU-ESM
Beijing Climate Center, China Meteorological Administration	BCC-CSM1-1 BCC-CSM1-1-m
National Center for Atmospheric Research	CCSM4
Centre National de Recherches Météorologiques / Centre Europen de Recherche et Formation Avancée en Calcul Scientifique	CNRM-CM5
Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence	CSIRO-Mk3-6-0
Canadian Centre for Climate Modelling and Analysis	CanESM2
LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences and CESS,Tsinghua University	FGOALS-g2 FGOALS-s2
NOAA Geophysical Fluid Dynamics Laboratory	GFDL-CM3 GFDL-ESM2G GFDL-ESM2M
NASA Goddard Institute for Space Studies	GISS-E2-H GISS-E2-R
Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)	HadGEM2-ES
Institut Pierre-Simon Laplace Institute for Numerical Mathematics	INM-CM4
	IPSL-CM5A-LR IPSL-CM5A-MR IPSL-CM5B-LR
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 MIROC5
Max-Planck-Institut für Meteorologie (Max Planck Institute for Meteorology)	MPI-ESM-LR MPI-ESM-MR MPI-ESM-P
Meteorological Research Institute	MRI-CGCM3
Norwegian Climate Centre	NorESM1-M

Table 5.1: CMIP5 modelling groups and models considered in this study.

5.4 Variability of the Lapse Rate Feedback

The dominant features of the lapse rate feedback can be seen in Fig. 5.1. As described in Section 5.2, the lapse rate feedback is strongly negative in the tropics where the upper troposphere warms faster than the surface and becomes positive in the high latitudes (poleward of ∼60°) where near surface temperature inversions tend to produce more warming near the surface [e.g., Payne et al., 2015, Roe et al., 2015]. The lapse rate feedback tends to be more strongly positive in the northern hemisphere, which likely results because surface warming is more rapid in the Arctic compared to the Antarctic over the timescales considered here [e.g., Stouffer et al., 1989, Manabe et al., 1991, Collins and Knutti, 2013, Armour et al., 2016].

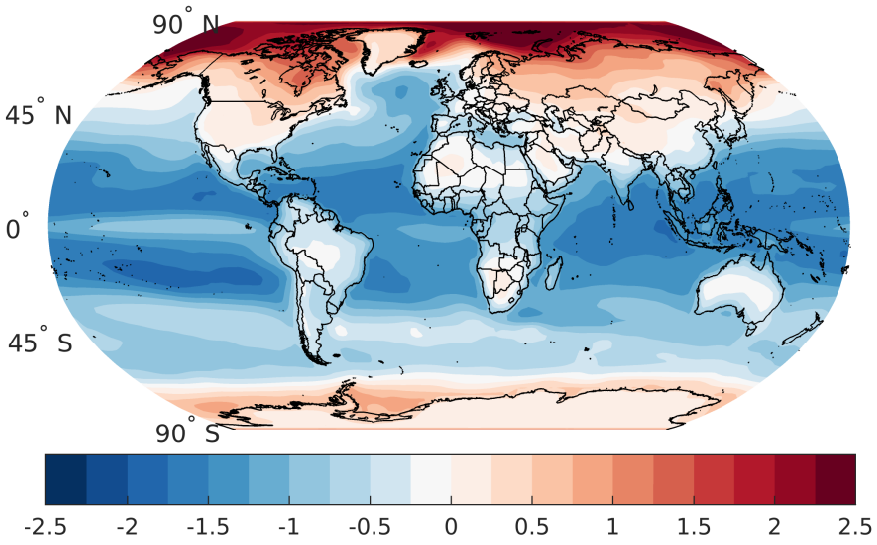

Figure 5.1: CMIP5 multi-model mean lapse rate feedback (Wm^{-2}/K). At each grid point the radiative response at the TOA due to lapse rate changes is normalized by the global average surface temperature for each model.

While Fig. 5.1 captures the mean behavior across models, Fig. 5.2 illustrates the differences between models from two perspectives. Fig. 5.2A shows the lapse rate feedback normalized by the zonal temperature change (i.e., the "local" lapse rate feedback) and Fig. 5.2B shows the lapse rate feedback normalized by the global mean surface temperature change for

each model (as in Fig. 5.1). The spread across models is reduced in the northern hemisphere as defined by the local feedbacks (Fig. 5.2A), which indicates that models have similar local radiative responses per unit local warming. In contrast, the local lapse rate feedbacks in the southern hemisphere tend to differ greatly across models, particularly near 60°S, which means that model radiative responses due to lapse rate changes do not consistently scale with local surface temperature change. This behavior has been noted in other studies [Feldl and Bordoni, 2016] and is not limited to the lapse rate feedback. The large spread in the local lapse rate feedback tends to be located in the same zonal band as a robust warming minimum as shown in Fig. 5.2C, which results from ocean heat uptake and transport in the Southern Ocean [e.g., Armour et al., 2016]. In fact, the zonally-averaged temperature change near 60°S is strongly correlated with the lapse rate feedback in this region across models ($r > 0.8$). This indicates that differences in warming in this region may be important for the local lapse rate feedback. The spread tends to be more uniform across latitudes in Fig. 5.2B, though this is more difficult to interpret because discrepancies between models could be a result of differences in the local lapse rate feedback or differences in the meridional warming pattern.

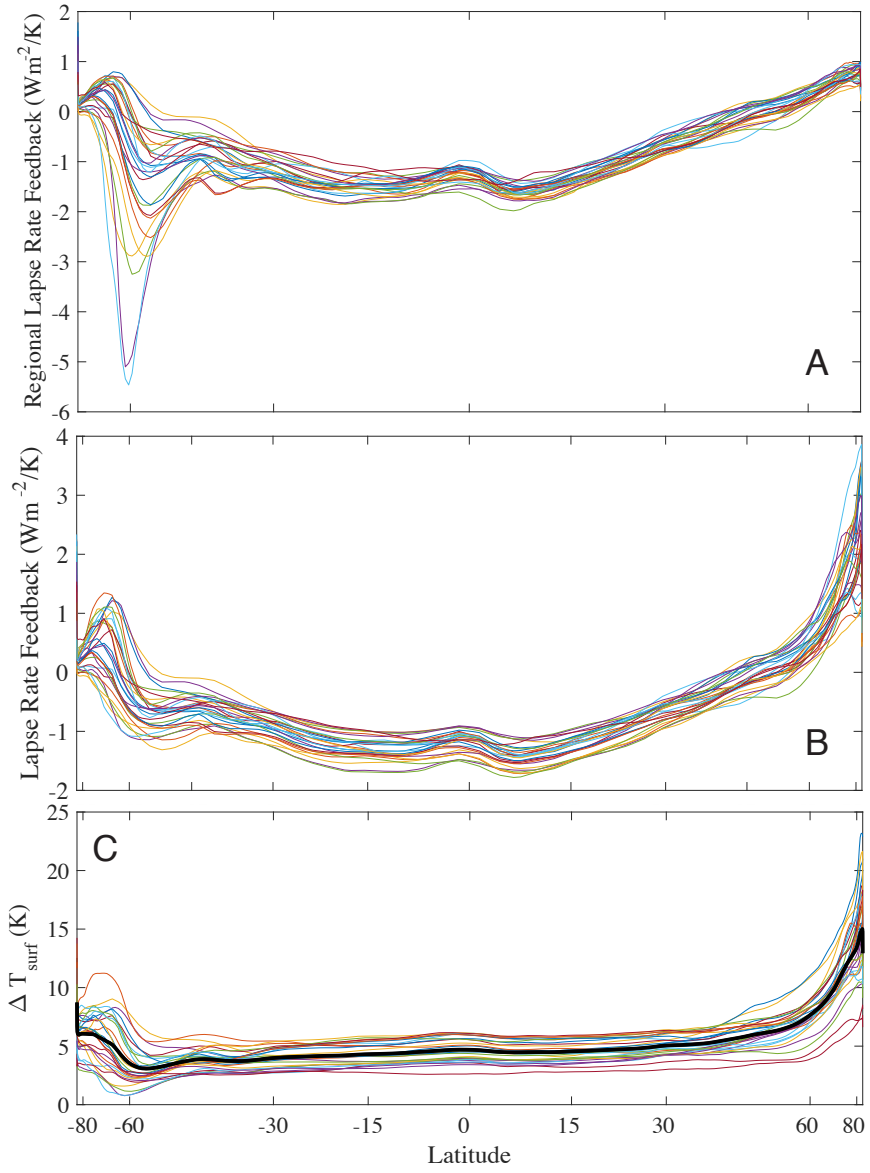

Figure 5.2: Zonal mean lapse rate feedback normalized by (A) zonal temperature change and (B) global temperature change. Zonal mean surface warming is displayed in C, with the multi-model mean denoted with a solid black line. The colored lines in each subplot represents a separate model.

In order to capture the leading mode of variability across models, we employ singular value decomposition (SVD) using the annual mean local lapse rate for each model. For this calculation we regrid each model to a common spatial domain and our analysis is performed

over the model and spatial dimension. This results in a set of empirical orthogonal functions (EOFs) that represent the main spatial modes of variance of the lapse rate feedback across models. The principal component (PC) of this analysis is a measure of how strongly the EOF projects onto each model. In Fig. 5.3 we show the leading EOFs that result from this analysis. These leading EOFs result from two separate calculations. In Fig. 5.3A, we solve for the EOF using the covariance matrix of the local lapse rate feedback data. This method tends to emphasize regions with enhanced variability and the EOF highlights the region near 60°S where there are large discrepancies in the local lapse rate feedback across models (see Fig. 5.2A). An alternative approach is to use the correlation matrix in which the local lapse rate feedback is normalized so that at each gridpoint the standard deviation across models is equal to one. This method de-emphasizes regions with large differences across models and highlights broader, coherent regions that may have less absolute variance. The leading EOF of the normalized local lapse rate feedback is shown in Fig. 5.3C. In Figs. 5.3B and 5.3D we also show the regression of the PCs for each leading EOF versus the global lapse rate feedback. Both PCs are significantly ($p < 0.05$) correlated with the global lapse rate feedback, which suggests that both patterns are important in determining the global lapse rate feedback, though the PC using normalized local lapse rate feedback values (Fig. 5.3D) has a stronger relationship. Given the relatively broad pattern in Fig. 5.3C (compared to Fig. 5.3A), it is unsurprising that the principal component is strongly related to the global average lapse rate feedback. We show both EOFs, because it is not clear *a priori* which approach is better, though we note that the PC from each EOF are strongly correlated with one another ($r = 0.86$). This suggests that the two approaches may be driven by common geophysical differences across models, except that the un-normalized analysis is sensitive to the large variance in the local lapse rate feedback at 60°S and thus misses more coherent regions of variability. From the EOF in Fig. 5.3C, we see that the mid-latitude region between the Antarctic and the deep tropics ($\sim$30° – 60°S) appears to be particularly important in understanding differences in the global lapse rate feedback across models.

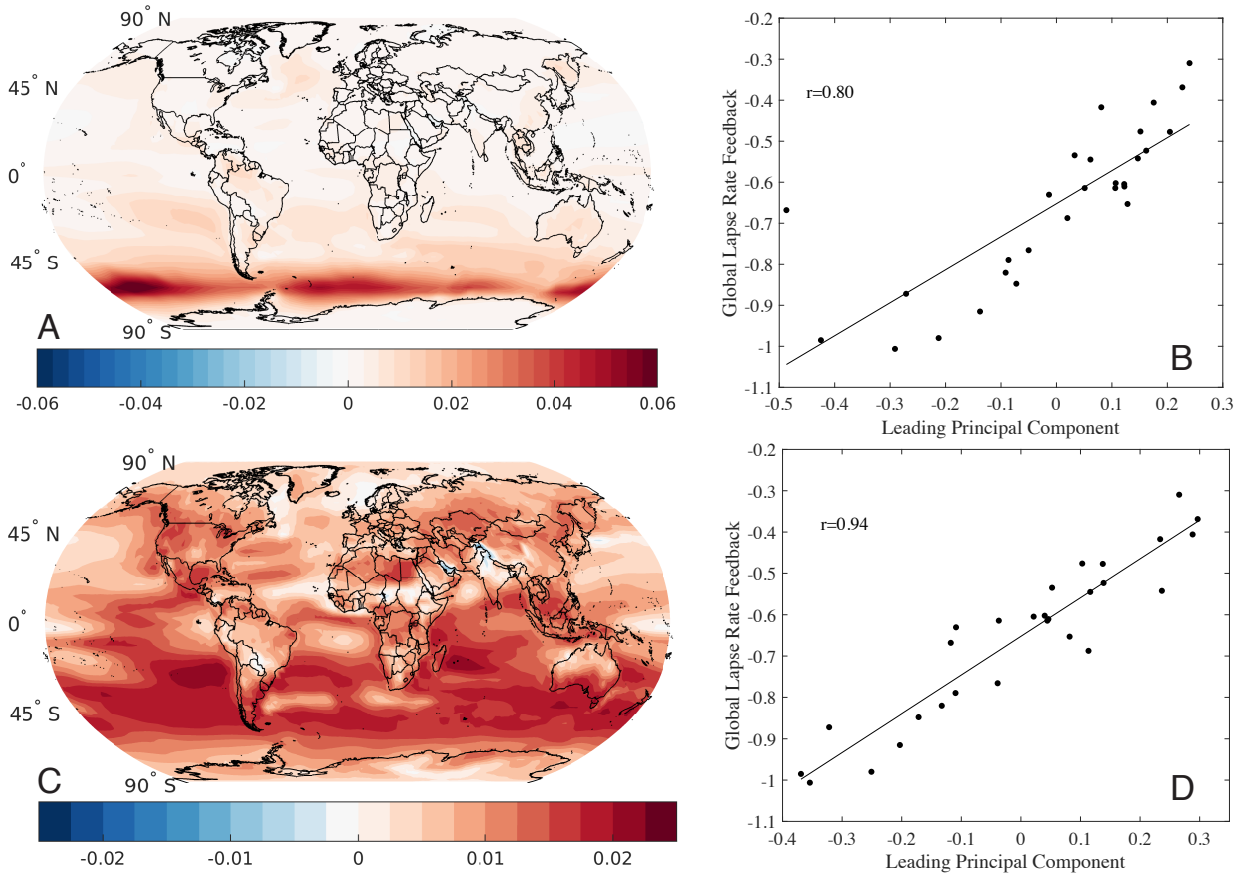

Figure 5.3: Principal component analysis of the local lapse rate feedback. A) The leading EOF of the local lapse rate feedback using the covariance matrix. This EOF explains 56% of the variance in the local lapse rate feedback across models. B) The principal component for each model from the EOF in (A) versus the global mean lapse rate feedback. C) As in (A) but for the correlation matrix of the local lapse rate feedback. This EOF explains 27% of the variance of the local lapse rate feedback across models. D) The principal component for each model from the EOF in (C) versus the global mean lapse rate feedback.

Soden and Held [2006] showed that the pattern of warming is a good predictor of the global lapse rate feedback. With this in mind, we also show the strength of the correlation between local warming and the global average lapse rate feedback in Fig. 5.4A. From this figure, we see that warming in the tropics tends to decrease the lapse rate feedback (i.e.,

more negative) and warming in the polar regions, especially in the Antarctic is associated with a more positive lapse rate feedback (i.e., closer to zero). This is consistent with Fig. 5.2, which shows that the local lapse rate feedback in the tropics is negative and so enhanced tropical warming should decrease the global lapse rate feedback. This is shown more clearly in Fig. 5.4B, which shows the correlation between the relative warming (e.g., local warming per global mean warming) versus the global lapse rate feedback. The pattern in Fig. 5.4A becomes stronger in Fig. 5.4B demonstrating that models that have *relatively* more warming over the tropics (poles) compared to the global average have a more negative (positive) global lapse rate feedback. This relationship is summarized in Fig. 5.4C (similar to Fig. 3 from Soden and Held [2006] for CMIP3 models). In line with the results demonstrating modes of variance in the lapse rate feedback across models (Fig. 5.3), the southern hemisphere tends to be more important compared to the northern hemisphere. In fact, the ratio of Antarctic warming to tropical warming is more strongly related to the global average lapse rate feedback ($r = 0.83$) than the ratio of Arctic to tropical warming ($r = 0.61$), even though the local lapse rate feedback tends to be stronger in the Arctic. Warming in the southern mid-latitudes and poles tends to be significantly related to the global lapse rate feedback, which is not true in the northern mid and high latitudes.

The results from the principal component and regression analysis highlight a few regions that are important for lapse rate feedback variability across models. From the principal component analysis, most of the variability in the local lapse rate feedback is in the southern hemisphere mid-latitudes and this variability is strongly correlated with the global mean lapse rate feedback. Similarly, from our regression analysis, we also see that model warming that is more concentrated in the tropics tends to create a stronger (negative) global mean lapse rate feedback. When warming is larger toward the poles, particularly in the southern hemisphere, the global average lapse rate feedback is more positive. Strangely, the lapse rate feedback and warming in the Arctic seem to be relatively unimportant in setting the global mean lapse rate feedback, even though this is where the local lapse rate feedback and warming is the strongest. Furthermore, although this analysis generally highlights the

importance of the southern hemisphere, regression analysis suggests that the balance of warming between the tropics and the poles is important, whereas the principal component analysis suggests that the strength of the local lapse rate feedback in between these regions is a controlling factor for the global average lapse rate feedback. In the next section, we seek to understand this statistical analysis and highlight important geophysical controls on the lapse rate feedback variability.

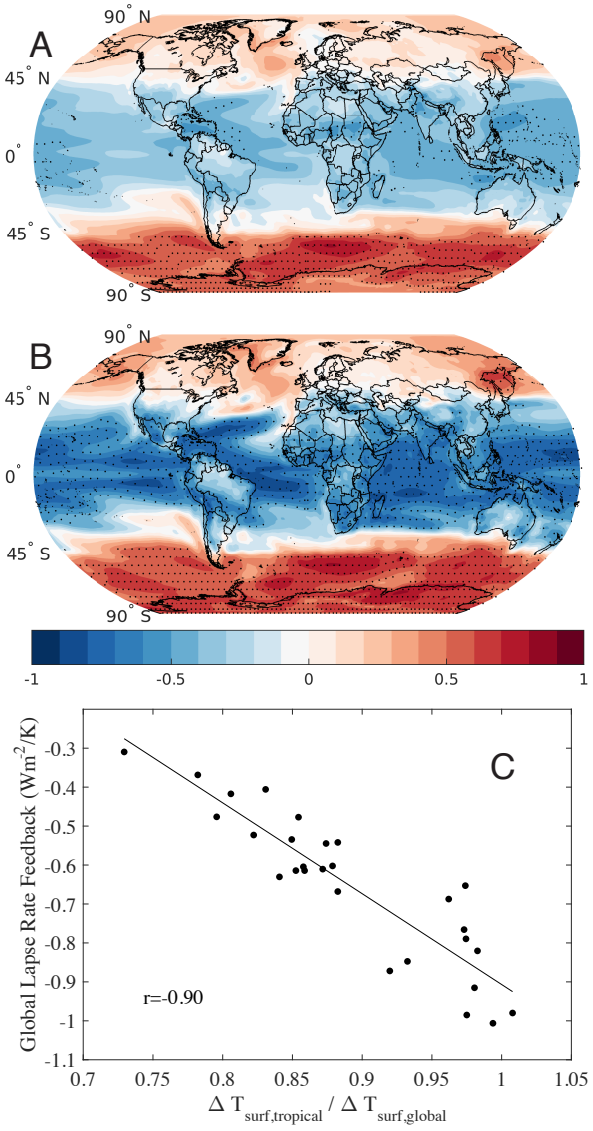

Figure 5.4: A) The correlation (across models) for the least squares linear fit between the surface temperature change at each point and the model's global mean lapse rate feedback. B) As in (A) except the local warming is normalized by the model's global mean warming (i.e., the amplitude of local warming relative to global average warming). C) The global average lapse rate feedback versus the ratio of tropical (30°S - 30°N) to global warming across GCMs. In (A) and (B) the stippling denotes that the correlation coefficient is significant at $p < 0.05$.

5.5 Understanding the variability and controls governing the lapse rate feedback

Results from the principal component and regression analysis in Section 5.4 demonstrate that the local lapse rate feedback in the southern mid-latitudes and the partitioning of warming between the tropics and the Antarctic are strongly related to each GCM's global average lapse rate feedback value. In this section, we begin by evaluating the extent to which these two results complement one another. We then demonstrate that differences in the sign and magnitude of southern hemisphere extratropical local lapse rate feedbacks are critical to the global average lapse rate feedback and demonstrate that these feedbacks have important non-local controls. This analysis elucidates the reasons behind the marked variability in the local lapse rate feedback near 60°S and why the local feedback is so stable across models in the northern hemisphere. We conclude our analysis by investigating high latitude Antarctic controls on polar amplification and their importance to the global average lapse rate feedback.

5.5.1 The pattern of meridional warming and local feedbacks

Fig. 5.5A shows the regression between the average surface warming over the south pole (60° – 90°S) and the zonal mean atmospheric warming across models. From this figure, we see that surface warming is largely confined to the lower troposphere poleward of 60°S, but is also significantly related to warming along the surface across the southern mid-latitudes (i.e., 30°S – 60°S). Fig. 5.5B shows the regression between tropical average surface warming and zonal atmospheric warming. The tropics also influence warming in the southern mid-latitudes, though in this case tropical surface warming is more strongly associated with warming aloft both in the tropics and the mid-latitudes. A linear combination of tropical (20°N–S) and Antarctic (60°S – 90°S) surface warming is predictive of surface warming across the southern mid-latitudes (r = 0.93), though warming poleward of 60°S alone is more strongly related to mid-latitude warming (r = 0.85) than tropical warming (r = 0.57). In the mid-latitude troposphere the situation is reversed, with tropical surface warming explaining nearly all of

the 300 hPa warming in the southern mid-latitudes (r = 0.95). In general, tropical warming tends to decrease the lapse rate in the southern hemisphere between 30° and 60° while polar warming tends to increase the lapse rate. Therefore, the balance of warming between the tropics and the south pole is an important control on the lapse rate feedback in the mid-latitudes. Consistent with this interpretation, the ratio of warming over the south pole relative to the tropics is strongly related to the local lapse rate feedback in the southern middle latitudes (r = 0.92). In this way, the pair of results in Section 5.4 that showed both that the local southern mid-latitude lapse rate feedback and the ratio of tropical to polar warming are important to the global lapse rate feedback are physically consistent with one another. This analysis is also consistent with Soden and Held [2006] that demonstrated that the meridional pattern of warming is important to the global average lapse rate feedback.

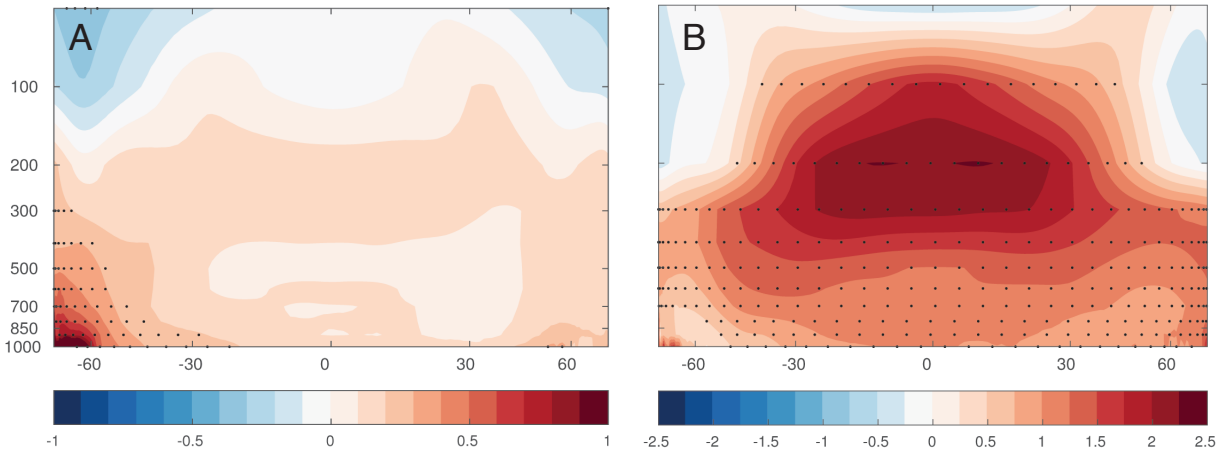

Figure 5.5: A) Slope of the regression between the surface temperature change poleward of 60°S and the change in the zonal atmospheric temperature change across models. B) As in (A), except using the tropical (30°S - 30°N) as a predictor. Regions where the slope is significant ($p < 0.05$) are stippled.

Although it is clear that the meridional pattern of warming controls the global lapse rate feedback, there are now two mechanisms that may be important. A direct effect of

the meridional profile of warming is that this pattern controls the balance of large negative feedbacks in the tropics and large positive feedbacks at the poles. On the other hand, tropical and Antarctic surface warming both influence the local lapse rate feedback in the mid-latitudes. Therefore, the pattern of warming's influence on local lapse rate feedbacks may also be important.

To assess whether the magnitude of the local lapse rate feedback is important in controlling the global lapse rate feedback value, we have estimated each model's feedback value making two different simplifications. To examine the role of the meridional pattern of warming in explaining the variance in the global lapse rate feedback, we compute the lapse rate feedback using the multi-model mean local lapse rate feedback, $\overline{\overline{\lambda_{lr,local}}}$. From Eq. 5.6, we have:

$$\overline{\lambda_{lr}} \approx \frac{\int_{-\pi/2}^{\pi/2} \overline{\overline{\lambda_{lr,local}(\phi)}} \cdot \Delta T_{surf}(\phi) \cdot cos(\phi) d\phi}{\int_{-\pi/2}^{\pi/2} \Delta T_{surf}(\phi) \cdot cos(\phi) d\phi} \tag{5.7}$$

where the double overbars denote that we are using the multi-model average in calculating each individual model's global average lapse rate feedback (single overbars have been used to denote spatial averages). Alternatively, we can examine the role of the meridional feedback pattern in explaining the global feedback value by using the multi-model mean meridional warming pattern, $\overline{\overline{\Delta T_{surf}}}$:

$$\overline{\lambda_{lr}} \approx \frac{\int_{-\pi/2}^{\pi/2} \lambda_{lr,local}(\phi) \cdot \overline{\overline{\Delta T_{surf}(\phi)}} \cdot cos(\phi) d\phi}{\int_{-\pi/2}^{\pi/2} \overline{\overline{\Delta T_{surf}(\phi)}} \cdot cos(\phi) d\phi} \tag{5.8}$$

We demonstrate the impact of these simplifications in Fig. 5.6. Perhaps surprisingly, the assumption that the meridional pattern of warming across models is invariant still captures 90% of the variance in the global lapse rate feedback (and also closely tracks the actual global lapse rate feedback values). On the other hand, assuming that the local lapse rate feedbacks are the same for each model is a relatively poor assumption. From Fig. 5.3C, we see that model differences in the lapse rate feedback in the southern mid-latitudes is the principal mode of variability. As such, the local lapse rate feedback over $30^\circ - 60^\circ$ is strongly predictive of the global lapse rate feedback (r = 0.92). This analysis of course does not imply that the

meridional pattern of warming is unimportant; from Fig. 5.5 we know that the meridional pattern of warming strongly modifies the local lapse rate feedback in this region. Since the southern mid-latitude local lapse rate feedback is a product of the competing influence of high and low latitude warming, the polar amplification in the southern hemisphere (ratio of tropical to Antarctic warming) is also predictive of the global lapse rate feedback (r = 0.83). Unlike in the northern hemisphere where Arctic surface warming is amplified relative to tropical surface warming, "Antarctic amplification" is less robust across models. We will refer to the relative Antarctic-tropical warming as Antarctic amplification, though in reality Antarctic surface warming is not consistently larger than tropical surface warming (see Fig. 5.2C). The warming pattern is thus important both because it balances the activation of positive high latitude and negative tropical feedbacks, but also influences the magnitude of the local feedbacks themselves.

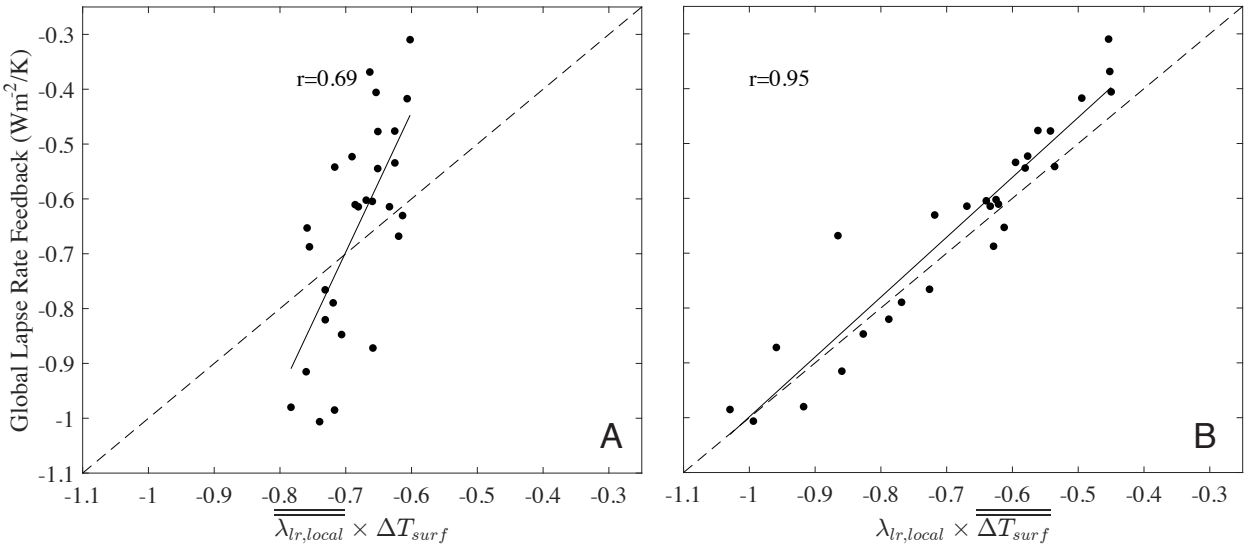

Figure 5.6: A) The global average lapse rate feedback versus the lapse rate feedback computed assuming that the meridional profile of the local lapse rate feedback is invariant across models. B) As in (A), except we instead assume that the meridional profile of warming is invariant across models. The dashed line in each figure represents the 1 : 1 line.

5.5.2 On the non-local influence on local feedbacks

Given that the meridional profile of the local lapse rate feedback is important in determining the global average lapse rate feedback for each model, we explore the stark differences in behavior between the northern and southern hemisphere local lapse rate feedback. From Fig. 5.2A, we see that models tend to have widely varying local lapse rate feedbacks poleward of ∼45°S, yet models tightly converge on the feedback in the northern hemisphere over the same latitude range. In Fig. 5.7A we see that the mid to high latitude TOA radiative response (ΔR_{lr}) due to the lapse rate scales linearly with warming in both the northern and southern hemisphere. This implies that surface warming in both hemispheres tends to be concentrated in the lower troposphere and that the radiative response increases with enhanced surface warming. Furthermore, the intercept is non-zero, which means that if there is little or no warming, the radiative changes due to lapse rate changes are actually negative (i.e., more radiation to space). This demonstrates that upper tropospheric warming is decoupled from the surface and has a non-local control. Indeed, the upper tropospheric warming (at 300 hPa) response is significantly related to tropical surface warming in both the northern (r = 0.92) and the southern hemispheres poleward of 45° (r = 0.92). A similar pairing of local and non-local controls was demonstrated to be important with local sea ice and SSTs largely controlling near surface warming and remote SSTs influencing temperature changes aloft [Screen et al., 2012]. This balance between locally driven surface warming and non-locally driven warming aloft has been characterized as "radiative-advective equilibrium," which can be seen as a framework to understand the lapse rate feedback in the high latitudes [Payne et al., 2015, Cronin and Jansen, 2016]. If a model has little surface warming in the mid to high latitudes, the radiative response from lapse rate changes is governed by tropical warming, which effectively warms the extratropical upper troposphere. As this region warms, the relatively stable atmosphere has a bottom-heavy warming profile and the local feedback switches to a positive feedback response. This behavior has important implications for the variability of the local feedbacks. In Fig. 5.7B, we show the local lapse rate feedback by

dividing the radiative response in Fig. 5.7A by the surface temperature change poleward of 45°S versus the surface warming over the same region (recall that $\lambda_{lr} = \Delta R_{lr}/\Delta T_{surf}$). In this figure we also show what the local lapse rate feedback would be if it followed the linear fit in Fig. 5.7A. The non-zero intercept value, which we attribute to non-local controls over the lapse rate feedback, has the effect of giving the local lapse rate feedback an apparent non-linearity when there is little surface warming.

This "non-linearity" refers to the curvature in Fig. 5.7B, though it arises from the *linear* combination atmospheric warming associated with tropical surface and Antarctic surface warming (see Fig. 5.5). Physical non-linear processes do exist, such as the transition from an atmosphere characterized by a near surface inversion (e.g., over sea ice) to a convective regime when sea ice is less extensive in a warmer climate [Payne et al., 2015, Cronin and Jansen, 2016], though we are not referring to this mechanism here. The non-local influence of tropical surface warming has important implications in the southern hemisphere where warming is suppressed relative to the northern hemisphere (see Fig. 5.2C) due to ocean heat uptake [e.g., Stouffer et al., 1989, Manabe et al., 1990, 1991, Armour et al., 2016]. The combined influence of tropically-driven upper tropospheric warming in the extratropics and the reduced surface warming in the southern hemisphere induce a highly variable local lapse rate feedback in the southern hemisphere, which is not present in the northern hemisphere because models all display robust polar amplification of surface temperature change in this region (see Fig. 5.2C).

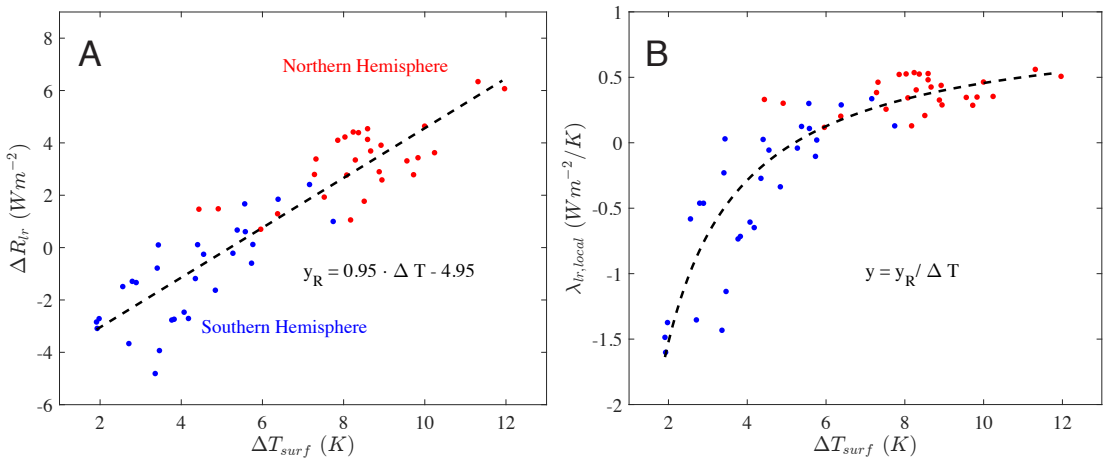

Figure 5.7: A) Average radiative flux at the TOA (positive down) due to lapse rate changes (ΔR_{lr}) versus the surface warming (ΔT_{surf}) poleward of 45° for each model in the northern hemisphere (red) and southern hemisphere (blue). The dashed line is the least-squares linear regression. B) The local lapse rate feedback ($\lambda_{lr,local}$) versus surface warming. The dashed line uses the linear fit from (A) and divides each point on the line by ΔT_{surf} in order to represent the local lapse rate feedback.

5.5.3 Controls on high latitude southern hemisphere warming

Given the role of the meridional pattern of warming in controlling the lapse rate feedback across models, it is important to understand the geophysical controls on the ratio of equator-to-pole warming. Since this analysis has highlighted the significance of the southern hemisphere and because the variance of Antarctic warming is roughly five times larger than the variance of tropical warming, we focus on mechanisms that influence surface warming over the south pole. In Fig. 5.8A we show the standard deviation of the warming across models at each gridpoint over the southern hemisphere poleward of 45°. We also show the annual mean sea ice extent (Fig. 5.8B). From this figure, we see that most of the spread in Antarctic warming is concentrated in regions that contain sea ice, which suggests that sea ice may be an important factor in controlling the warming over the southern hemisphere (60°S – 90°S).

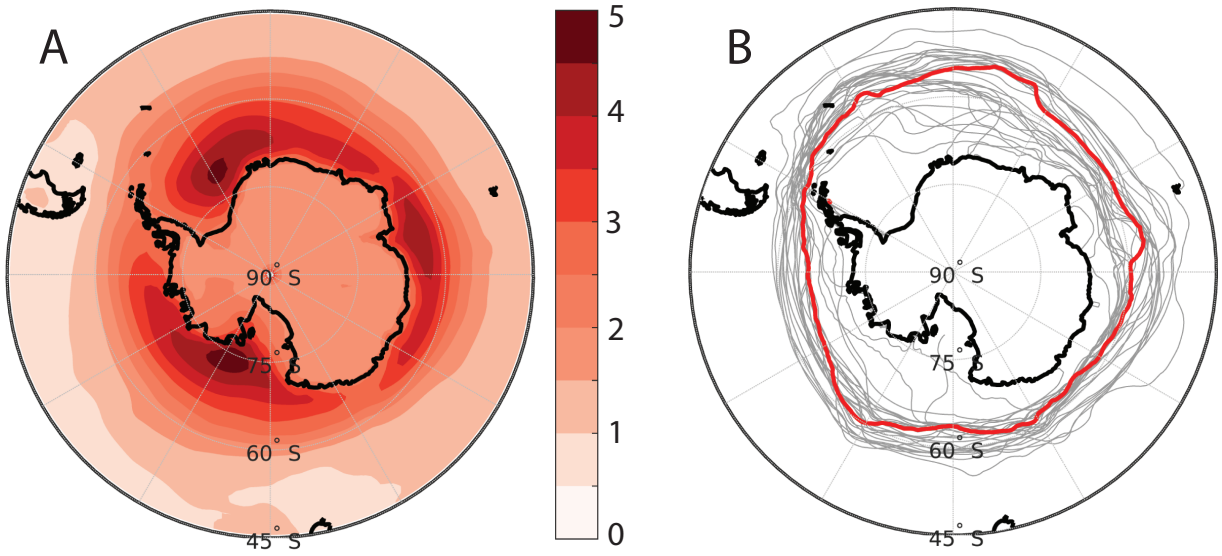

Figure 5.8: A) Standard deviation of warming across models at each gridpoint. B) Annual mean sea ice extent (contours denote 5% sea ice concentration) for each model's piControl simulation (gray lines) and observed sea ice extent (red thick line) over 1979 - 2014 [Peng et al., 2013, Meier et al., 2013].

Sea ice thickness and snow cover modulate the thermal conductivity of the ice pack, which in turn regulates surface heat fluxes [e.g., Lecomte et al., 2013, Lei et al., 2014]. Pavelsky et al. [2011] has shown that annually averaged sea ice concentration is a reasonable surrogate for the effective thermal conductivity of sea ice. At a basic level, sea ice insulates the ocean from the atmosphere, which reduces heat fluxes from a relatively warm ocean into the atmosphere, particularly in the wintertime. In Fig. 5.9A we show the wintertime difference between the surface skin temperature and the sea-surface temperature (SST). Over regions with large concentrations of sea ice, the SST (below the ice) is much larger than the surface temperature. Fig. 5.9B demonstrates that models with increased climatological sea ice cover have reduced wintertime upward heat fluxes as the model's relatively warm ocean is isolated from the atmosphere. The strong relationship between sea ice cover and surface heat fluxes across a range of models indicates that models with enhanced ice cover

loss may also have enhanced upward heat fluxes in warming scenarios. Models with large sea ice concentrations also have cooler sub-Antarctic surface temperatures, which produces enhanced near surface temperature inversions in agreement with observations that show the same relationship [Pavelsky et al., 2011]. Figure 5.9C demonstrates that models with greater sea ice cover have stronger inversions, which has been shown to influence longwave feedbacks in the high latitudes [e.g., Boé et al., 2009, Medeiros et al., 2011, Payne et al., 2015]. Strong inversions enhance near-surface warming [e.g., Payne et al., 2015]. It is therefore possible that aspects of a model's unforced climatology may affect the model response to greenhouse gas forcing. For example, Flato [2004] showed that initial sea ice extent is predictive of the Antarctic temperature change across CMIP1 and CMIP2 models. Since sea ice extent is able to explain the climatological variability across models for geophysical fields that are critical to model feedbacks, we will investigate the degree to which sea ice modulates the southern hemisphere high latitude response to greenhouse gas forcing.

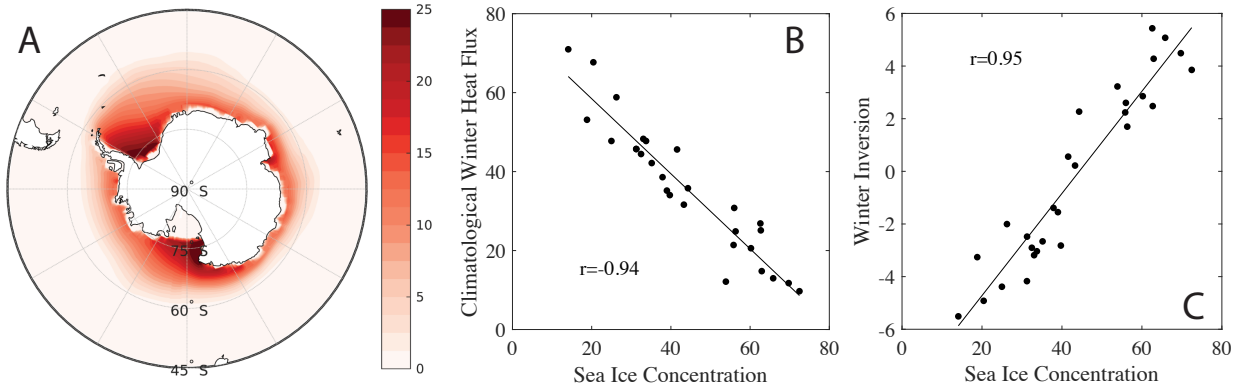

Figure 5.9: A) Wintertime (June - August) sea-surface temperature minus skin temperature for the CMIP5 ensemble average. Some regions with ice shelves have no data (white). BNU-ESM is not included due to an offset in the ocean grid. FGOALS-g2 and FGOALS-s2 are not included because these model fields are not available. B) Climatological annual mean sea ice concentration versus wintertime (June - August) climatological upward surface heat (sensible and latent, Wm^{-2}) flux for CMIP5 models. C) Climatological annual mean sea ice concentration versus the climatological wintertime inversion strength (defined as the temperature at 850hPa minus the near surface temperature). Climatologies in (B) and (C) are derived from model piControl simulations over the Antarctic ocean (60° - 90°S).

Because winter heat fluxes scale with sea ice cover across a range of individual model climatologies (see Fig. 5.9A), there is an expectation that the change in model ice extent may vary with the change in surface heat fluxes. Indeed, as can be seen in Fig. 5.10A, the change in sea ice extent is strongly related to the change in surface heat fluxes over the ocean. Large declines in sea ice also leads to enhanced radiative fluxes from changes in the lapse rate and albedo (Fig. 5.10B-C). These processes all work together to strongly shape the surface temperature change, which is closely tied to the loss of sea ice (Fig. 5.10D). Screen and Simmonds [2010] have similarly found that sea ice losses over the observational record are tied to Arctic warming and moistening. A number of studies have demonstrated that high latitude feedbacks tend to amplify one another, which means that the processes considered here work in tandem to enhance Antarctic temperature response differences between models

[Graversen et al., 2014, Feldl et al., 2016]. Finally, since we have considered experiments that are perturbed with a large ($4xCO_2$) forcing, much of the Antarctic sea ice melts by the end of the $4xCO_2$ run (150 years after quadrupling). As a result, the change in Antarctic sea ice is a function of the initial sea ice extent (Fig. 5.11A). Changes in winter surface heat fluxes and radiative fluxes from summer albedo and winter lapse rate changes then scale both with the change in sea ice extent and the model's preindustrial control sea ice climatology ($r > 0.80$ in all cases). In Fig. 5.11B, we see that Antarctic temperature change is related to the preindustrial sea ice extent, which has a significant impact on the polar amplification in the southern hemisphere (Fig. 5.11C). In all, the climatological inversion strength, changes in surface heat fluxes, albedo and lapse rate feedbacks are intimately tied to the climatology and response of Antarctic sea ice and act together to enhance the Antarctic surface temperature change.

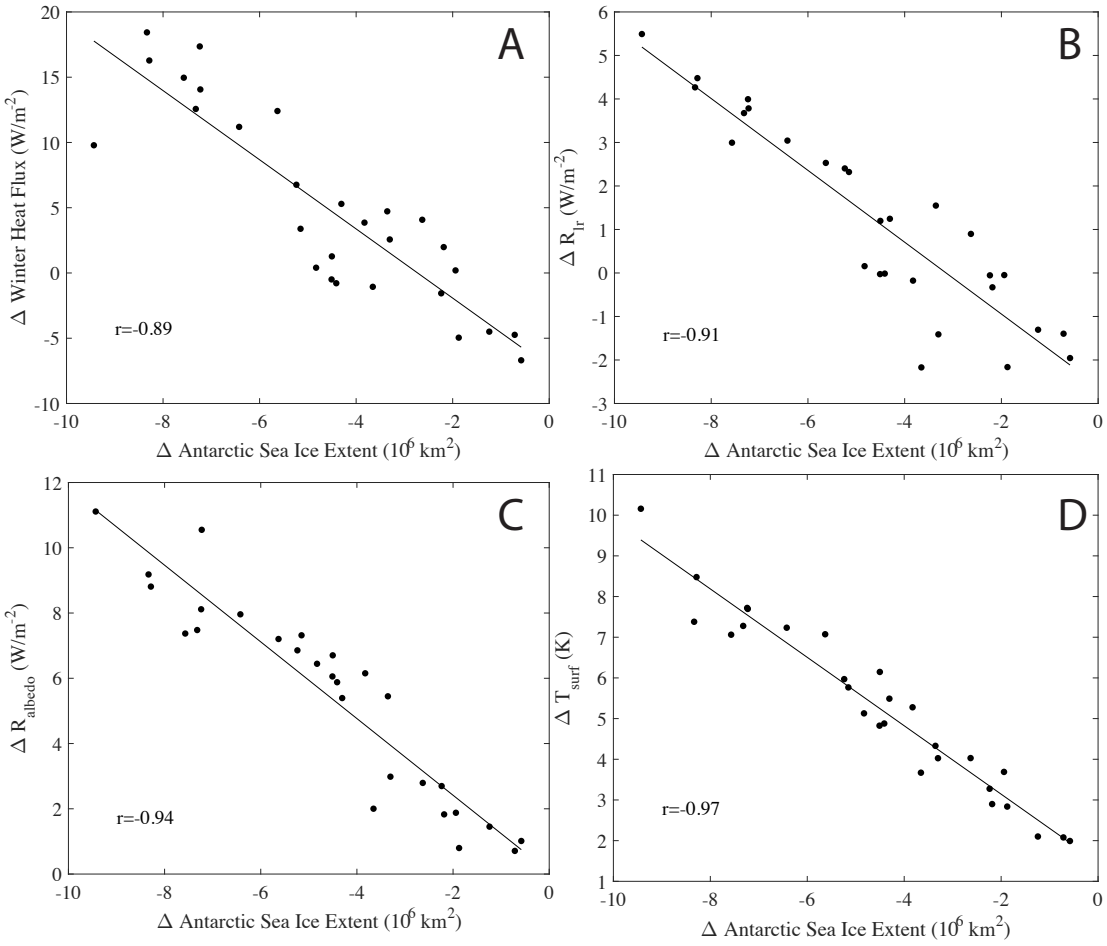

Figure 5.10: The change in Antarctic sea ice extent compared to A) change in upward wintertime surface heat (latent and sensible) flux, B) change in TOA radiation (positive is down) due to lapse rate changes, C) change in TOA radiation due to albedo changes, and D) change in near-surface temperature. All fields are computed over the ocean over 60° - 90°S.

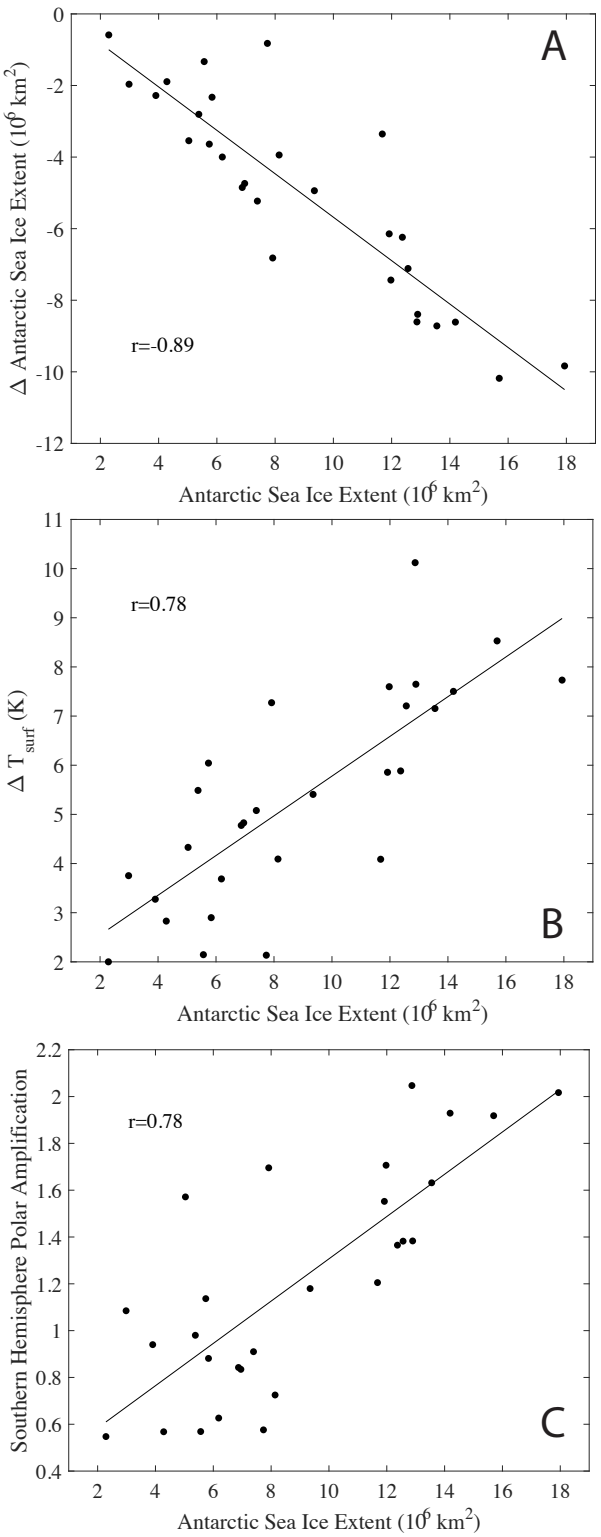

Figure 5.11: The annual mean sea ice extent in each model compared with the change in A) Antarctic sea ice extent B) change in near surface temperature poleward of 60°S and C) the ratio of southern high latitude (60° - 90°S) surface warming relative to tropical (30°S - 30°N) surface warming.

In Section 5.5.1 we showed that the meridional warming pattern is important to the global lapse rate feedback both because it modulates the ratio of positive (high latitude) to negative (low latitude) feedbacks and also influences the magnitude of the local lapse rate feedback in the southern mid-latitudes. As such, southern hemisphere polar amplification is a significant predictor of the global lapse rate feedback across GCMs. Antarctic sea ice thus plays a role in shaping the global lapse rate feedback. Models with more expansive Antarctic sea ice have larger near surface inversions, enhanced surface heat fluxes, and positive feedbacks that lead to greater Antarctic surface warming, which strengthens the southern hemisphere polar amplification. As a result, southern hemisphere sea ice loss (Fig. 5.12A) and extent (Fig. 5.12B) significantly covary with the global lapse rate feedback (r^2=0.30 and r^2=0.37, respectively).

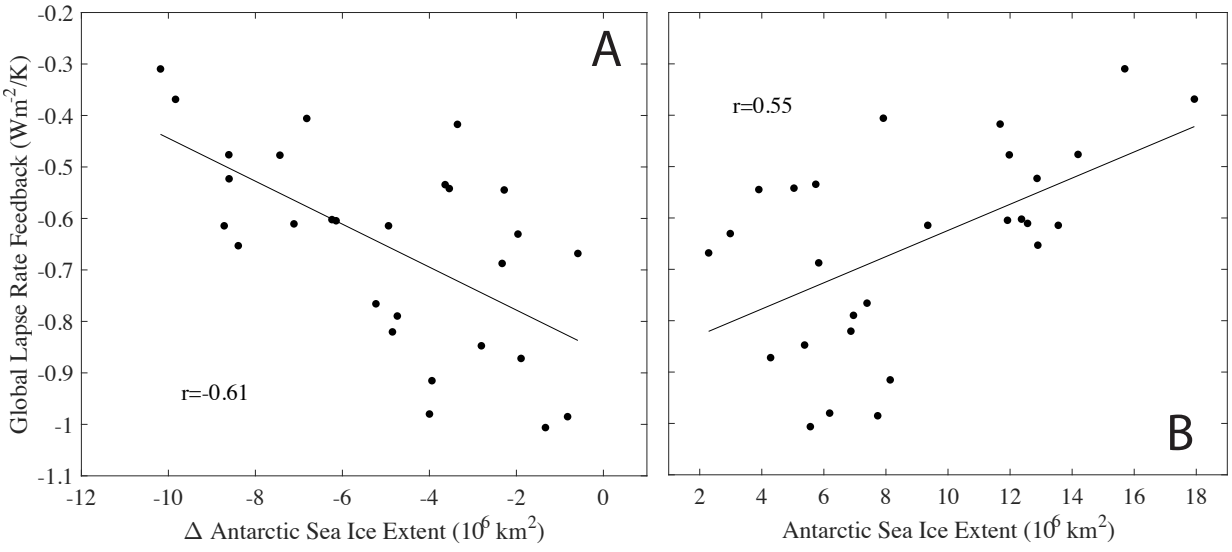

Figure 5.12: Global lapse rate feedback versus A) the change in the annual average Antarctic sea ice extent and B) the annual average Antarctic sea ice extent.

Although the relationship between climatological Antarctic sea ice and the temperature response is likely partly due to the large forcing in this set of model experiments, this relationship does not hold in the Arctic. In the northern hemisphere, the Arctic temperature

response and sea ice response is weakly related to the initial sea ice extent (r = -0.38 and r = -0.41, respectively). In contrast, the connection to Antarctic sea ice extent emerges within the first decade of CO_2 quadrupling in the southern hemisphere for both temperature change (r = 0.87) and change in sea ice cover (r = -0.84). This pattern emerges in transient forcing simulations as well – in CMIP5 experiments in which atmospheric CO_2 is increased at 1% per year, the loss of Antarctic sea ice is significantly related to the initial sea ice extent in the decade after CO_2 doubling (r = 0.74).

5.6 Discussion and Conclusions

This analysis focused on understanding differences in the lapse rate feedback across GCMs. This work shows that the variance in the local lapse rate feedback across models is concentrated in the southern hemisphere extratropics and that differences in the southern hemisphere lapse rate feedback, in particular, are strongly predictive of the global mean lapse rate feedback. Although it has long been known that the pattern of warming controls global average lapse rate changes [Soden and Held, 2006], this work highlights the southern hemisphere as a region of particular interest. We use a local feedback framework [Feldl and Roe, 2013, Armour et al., 2013] to identify and understand the role of non-local influence on feedbacks and to shed light on the large inter-model discrepancies in feedback strength in the southern hemisphere noted in other studies [e.g., Feldl and Bordoni, 2016]. We're able to show that the robust minimum in Southern Ocean warming [Stouffer et al., 1989, Manabe et al., 1990, 1991, Armour et al., 2016] and the non-local influence of tropical warming on the extratropical lapse rate feedback play an important role in explaining the large variability in the southern hemisphere local lapse rate feedback and, by extension, the global average lapse rate feedback. We find that much of the variance in southern hemisphere polar amplification is modulated by climatological sea ice. Since the meridional warming profile is important in balancing the influence of local and non-local influences on the lapse rate feedback, Antarctic sea ice plays a modest, but important role in shaping the global lapse rate feedback.

The region between 30° – 60°S is influenced by both tropical and polar warming, which

work together to control the local lapse rate feedback in this region. As a result, the local lapse rate feedback in the southern mid-latitudes is a proxy for the polar amplification over the southern hemisphere. Although other studies have noted the importance of the partitioning of tropical and high latitude warming [e.g., Soden and Held, 2006], we have shown that the surface polar amplification response in the southern hemisphere largely controls the global lapse rate feedback. Furthermore, the ratio of Antarctic to tropical warming is important not just for balancing positive (high-latitude) and negative (low-latitude) feedbacks, but actually modulates the local feedback strength in the southern hemisphere extratropics. The importance of the local lapse rate feedback in the southern extratropics is heightened relative to the northern hemisphere, because it is highly variable across models. The meridional pattern of warming in isolation is not sufficient to accurately predict the global lapse rate feedback. The spatial pattern of warming also invokes changes in the meridional pattern of the local lapse rate feedback, which is a key determinant of the global lapse rate feedback.

The variability in the southern hemisphere extratropical local lapse rate feedback is driven by the non-local influence of tropical surface warming, which tends to reduce the lapse rate over the southern extratropics, and the reduced warming over the Southern Ocean, which creates a highly variable local lapse rate feedback in the southern hemisphere. Models with relatively little warming over the southern hemisphere extratropical surface are dominated by upper tropospheric warming (that originates in the tropics) [e.g. Screen et al., 2012], which produces a negative local lapse rate feedback. In models with increased surface warming, changes in temperature are concentrated at the surface and the local lapse rate feedback quickly becomes positive. As a result, the southern extratropical local lapse rate feedback is highly variable as it transitions from a negative to positive local feedback. In the northern hemisphere, polar amplification is large enough that near surface warming dominates and the local lapse rate feedback is positive with markedly less variability. This helps shed light on the highly variable nature of local feedbacks in the southern hemisphere that has been noted in other studies [Feldl and Bordoni, 2016] and points to muted Southern Ocean warming as a key difference between the northern and southern hemisphere [Stouffer et al., 1989, Manabe

et al., 1990, 1991, Collins and Knutti, 2013, Armour et al., 2016].

The pattern of meridional warming over the southern hemisphere is strongly modulated by Antarctic warming, which has much greater variability across models compared to tropical warming. We do not attempt to separate the influence of the inversion strength, sea ice change, radiative feedbacks, and changes in surface heat fluxes, which all contribute to and are enhanced by Antarctic warming. As such, Antarctic warming over 60°S - 90°S is a metric that encapsulates a number of physical processes. An important factor that contributes to the model spread in warming over the Antarctic is the climatological sea ice cover, which is predictive of sea ice loss, inversion strength, changes in surface heat fluxes, and radiative feedbacks. These processes all act together such that model differences in climatological sea ice extent are predictive of Antarctic temperature change and southern hemisphere polar amplification. The tendency for models with larger sea ice extents to exhibit greater warming has been noted in previous model ensembles with ramped CO_2 forcing (r=0.48) [Flato, 2004], though we find a much stronger connection in this work (r=0.78). Furthermore, Flato [2004] do not find evidence that initial Antarctic sea ice extent influences the change in sea ice extent, whereas we find a strong connection within a decade of CO_2 quadrupling and in the decade following CO_2 doubling in an experiment where CO_2 is increased at 1% per year. We do not find strong evidence for an influence of climatological Arctic sea ice extent on the change in Arctic temperature or sea ice coverage. Antarctic sea ice influences the meridional warming profile in the southern hemisphere and, in turn, contributes to as much as 30% of the global lapse rate feedback variability across models, even though Antarctic sea ice represents less than 4% of the Earth's surface area. This suggests that the initial model climatological state is important for feedback mechanisms and contributes to biases in the model response to greenhouse gases.

Chapter 6

CONCLUSION

This dissertation first focuses on understanding large discrepancies in the satellite microwave temperature record (Chapter 2). Three independent efforts (from UAH, RSS, and NOAA STAR) to homogenize the mid-tropospheric temperature record, result in tropical mid-tropospheric trend estimates that vary by a factor of three. This is striking, because each team uses the same raw data. We investigated this issue by building a dataset from Level 1C data, applying varied diurnal corrections and warm target instrument calibrations. Corrections for drifts in diurnal sampling derived from GCMs leave residual trends in the differences between co-orbiting satellites. Because RSS and NOAA STAR use the diurnal cycle from a GCM to correct for diurnal sampling issues, their long-term trends are influenced by their diurnal correction methodology. We developed and applied an observationally-based diurnal cycle adjustment, which effectively removes residuals from co-orbiting satellites. This technique has since been adopted by RSS (version 4.0) [Mears and Wentz, 2016] and NOAA STAR (version 4.0), resulting in significantly larger mid-tropospheric temperature trends. In both of these datasets, the tropical (20°N-S) mid-tropospheric temperature (TMT) trends (1979 - 2015) increase by about 25%, from 0.095 to 0.132 (0.125 to 0.160) K decade^{-1} for RSS (NOAA). This improvement is important, because it demonstrates that tropospheric warming is amplified relative to the surface. Over 1979 - 2015, the trend amplification between the tropical troposphere (T24) and the surface (HadCRUT4) is 1.35 (1.62) for RSS (NOAA). Typical amplification in climate models is $\sim$1.4 - 1.6 [Santer et al., 2005, Fu et al., 2011].

In contrast to RSS and NOAA, UAH uses a different approach to the diurnal cycle correction, which relies on deriving a diurnal cycle from co-orbiting satellites. We found residual

errors in their diurnal cycle correction. This bias appears to have been addressed recently, acting to reduce the mid-tropospheric temperature trend in their new dataset (UAH v6.0 beta 5). Although we were unable to thoroughly examine the UAH diurnal cycle correction, we did find that longstanding issues with the calibration of the NOAA-9 satellite remain [Po-Chedley, 2012, Po-Chedley and Fu, 2013]. This issue is seemingly addressed in the new UAH dataset resulting in a modest net trend increase (from 0.052 to 0.066 K decade^{-1}) and bringing their data into closer accord with the RSS and NOAA datasets. Despite these improvements, the UAH tropical tropospheric dataset (T24) still exhibits less warming than the surface over 1979 - 2015. The UAH tropical T24 trend is 0.101 K decade^{-1}, while the HadCRUT4 surface trend is 0.125 K decade^{-1}. UAH data continues to be an outlier relative to complementary datasets, GCM simulations over the satellite record, and is in conflict with our basic understanding of the processes that control atmospheric temperature change in the tropics. It is possible that some of these differences relate to the treatment of microwave data from the NOAA-14 and NOAA-15 satellites [Mears and Wentz, 2016]. NOAA-14 microwave observations exhibit marked warming compared to observations from NOAA-15. UAH attributes this discrepancy to NOAA-14 observations and removes these data earlier than RSS and NOAA, which likely reduces the UAH trend. It will therefore be important to investigate this issue as it has important implications for estimates of tropospheric warming and resolving lingering dataset discrepancies. Potential causes for this intersatellite difference include instrument calibration issues and differences in the central frequency for NOAA-14 (MSU) and NOAA-15 (AMSU) [e.g., Zou and Wang, 2011, Lu and Bell, 2014].

In Chapter 3, we investigate model vertical temperature amplification between the tropical upper and lower-middle troposphere. Note that this amplification, between the upper and lower-middle troposphere, is different from the vertical amplification between the surface and the troposphere discussed in Chapter 2. Amplification between the upper (T24) and lower-middle (TLT) troposphere helps to isolate specific atmospheric layers, which essentially tests the depth of amplification in the troposphere. Studies using CMIP3 models had previously shown that coupled atmosphere-ocean models tend to exaggerate upper tropospheric trends

compared to satellite microwave observations [Fu et al., 2011]. We updated this analysis by utilizing the most recent generation of coupled-atmosphere ocean models from CMIP5 and by using model simulations with prescribed SSTs in order to reduce the influence of natural variability on model-observational differences. We found that even when models use observational SSTs as a boundary condition, they still tend to exhibit excessive increases in static stability. Other studies have found similar results using radiosondes [Seidel et al., 2012], even when using prescribed SST experiments from CMIP5 [Mitchell et al., 2013b]. Our results did point to a potential role of SSTs. CMIP5 model simulations were supposed to use the Hurrell et al. [2008] SST dataset, but several modeling centers used the HadISST dataset instead [Rayner et al., 2003]. Models using the HadISST dataset had mid-tropospheric amplification that was in accord to satellite observations. Flannaghan et al. [2014] investigated this issue further by applying both the HadISST and Hurrell SST datasets to individual models (GFDL and NCAR models). This research showed that the SST dataset can have an important influence over the tropical tropospheric temperature trends (as measured by T24). Differences in SST trends in regions of deep convection were highlighted as being important drivers of atmospheric temperature trend differences.

Chapter 4 looks closer at potential mechanisms that may drive model-observational discrepancies in tropical vertical amplification between the upper and lower-middle troposphere. Part of this discrepancy may result from issues with data used to force GCM simulations. Observations show a wide range in tropical lower stratospheric ozone trends. CMIP5 models were also forced with stratospheric volcanic aerosol datasets that underestimated the concentration of stratospheric volcanic aerosols after 2000. Furthermore, Flannaghan et al. [2014] showed that observational uncertainty in the pattern of SST warming can have a large influence on tropical tropospheric temperature trends. We undertake sensitivity studies and show that it is unlikely that model biases in the ozone or volcanic aerosol forcings influenced model-observational discrepancies in changes in mid-tropospheric static stability. On the other hand, we found that the ratio of T24 and TLT trends were sensitive to the SST dataset, which suggests that SST uncertainty needs to be taken into account when compar-

ing models and observations of tropical tropospheric amplification. Improvements to model physics are also important. The GISS E2 model was modified to better capture the MJO [Kim et al., 2012], the vertical distribution and radiative effects of clouds [Yao and Cheng, 2012], and the fractionation of atmospheric water vapor [Field et al., 2014]. These improvements also enhanced agreement with observations for the vertical structure of temperature trends and variability. This suggests that model vertical amplification behavior may also serve as a metric to improve model parameterizations. Using estimates of observational uncertainty, about half of CMIP5 model ensemble members exhibit significantly more tropical vertical amplification than observations. Taken as a whole, significant discrepancies still exist between observations and model simulations of tropical tropospheric warming, but SST observations may explain some of this difference. Similarly, improvements to the GISS E2 model demonstrate that improved model physics can also be important and may bring model behavior into closer accord with satellite observations of amplification. Last, given that large discrepancies still exist for co-orbiting satellite measurements and across the UAH, RSS, and NOAA STAR microwave temperature datasets, it is possible that observational structural uncertainty may still be responsible for some of the differences noted here.

An underlying motivation for assessing the vertical structure of tropical tropospheric temperature trends in models is that this behavior is responsible for variations in the lapse rate feedback. This fundamental feedback varies widely across models and significantly covaries with the water vapor feedback [Caldwell et al., 2016]. We find that the ratio of upper (T24) and lower-middle (TLT) tropical tropospheric warming is a relatively poor metric for the lapse rate feedback. On the other hand, an excellent metric for the lapse rate feedback is the ratio of upper tropospheric (T24) and surface warming. In contrast to observed multi-decadal warming, the tropical SST warms relatively uniformly in coupled atmosphere-ocean simulations forced with increased greenhouse gas concentrations. Thus the pattern of SST warming does not play a significant role in coupled atmosphere-ocean model simulations of the vertical structure of atmospheric warming. Differences in amplification, and thus differences in the lapse rate feedback, can be related to model physics. A perturbed physics

ensemble using the GFDL AM4 model results in a wide range of values for the lapse rate feedback by modifying the parameterization for the conversion of cloud water to rain. We then mimic this behavior using the NCAR CAM4 model by varying only parameters related to deep convection. Even though model parameterizations can dramatically alter the magnitude of the tropical lapse rate feedback, model simulations of amplification over the satellite record are only weakly predictive of the quasi-equilibrium amplification in experiments that quadruple the concentration of atmospheric CO_2. Furthermore, model differences in the tropical lapse rate feedback explain less than half of the variance in the global average lapse rate feedback, which suggests that warming in the extratropics is an important factor.

In Chapter 5, we expand the scope of this analysis to the entire globe, to explore the factors that influence variability in the global lapse rate feedback across models. We find that models with enhanced warming in the tropics tend to have a smaller global lapse rate feedback, whereas models with larger warming at the poles, especially over the Antarctic, tend to have a larger global lapse rate feedback. Using principal component analysis, we find that most of the intermodel variability in the local lapse rate feedback is over the southern hemisphere extratropics and that this variability projects strongly onto differences in the global average lapse rate feedback. We are able to relate these two analyses, showing that the relative warming between the tropics and the Antarctic strongly modifies the local lapse rate feedback over the southern hemisphere extratropics. Tropical surface warming tends to enhance upper tropospheric warming over the entire globe. In models with little southern hemisphere extratropical surface warming, the southern hemisphere lapse rate feedback is strongly negative. In models with larger Antarctic and sub-Antarctic warming, lower tropospheric temperature changes dominate and the lapse rate feedback becomes positive. This behavior of the southern hemisphere extratropical lapse rate feedback is non-linear, which we attribute to the muted warming over the southern ocean. This non-linear behavior leads to large model-to-model discrepancies in the local lapse rate feedback in the southern hemisphere, which is a critical component in the spread of the global lapse rate feedback. This same non-linear behavior is not exhibited in the northern hemisphere, because the Arctic

experiences robust polar amplification across models, leading to a consistent, positive northern extratropical lapse rate feedback. Reduced southern ocean warming thus is a critical component of the global lapse rate feedback. Given that the meridional pattern of warming over the southern hemisphere is important and that most of the variance in southern hemisphere warming is in the Antarctic, we explored factors that influence warming in this region. We found that sea ice change strongly influences the albedo feedback, the lapse rate feedback, wintertime surface heat fluxes, and surface temperature change. Further, in the large forcing ($4xCO_2$) simulations considered, the change in sea ice is strongly related to the initial Antarctic sea ice extent. We are then able to relate the Antarctic warming and the global lapse rate feedback to the model's climatological sea ice distribution. Although the model's Antarctic sea ice extent covers less than 4% of the Earth's surface area, it explains about 30% of the variance in the global lapse rate feedback.

This research was successful in providing insight into observational issues with satellite temperature data and in characterizing and interpreting differences in the structure of atmospheric warming between models and observations. Despite this progress, a number of important issues remain. Significant discrepancies between both co-orbiting satellite measurements and across microwave temperature datasets remain and must be addressed to reduce the uncertainty in estimates of long-term atmospheric temperature change. Differences in the vertical structure of warming between models and observations are still not fully understood; future work should assess the role of uncertainty in the pattern of SST warming and determine if interannual amplification behavior can act as a constraint on uncertain model physics. While a great deal of attention is paid to Arctic amplification, we have shown that muted warming over the southern ocean induces large discrepancies in feedbacks in the southern hemisphere. This suggests that we should continue to scrutinize model behavior in the southern ocean, particularly in model sea ice fields [e.g., Meehl et al., 2016], which are important to the surface temperature response in the sub-Antarctic. Continued work analyzing discrepancies across observational datasets and between models and observations has the potential to improve both model simulations and observations of global change.

BIBLIOGRAPHY

K. C. Armour, C. M. Bitz, and G. H. Roe. Time-Varying Climate Sensitivity from Regional Feedbacks. *Journal of Climate*, 26(13):4518–4534, 2013. ISSN 08948755. doi: 10.1175/JCLI-D-12-00544.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-12-00544.1`.

K. C. Armour, J. Marshall, J. R. Scott, A. Donohoe, and E. R. Newsom. Southern Ocean warming delayed by circumpolar upwelling and equatorward transport. *Nature Geoscience*, 9(7):549–554, 2016. ISSN 1752-0894. doi: 10.1038/ngeo2731. URL `http://www.nature.com/doifinder/10.1038/ngeo2731`.

A. R. Atwood, E. Wu, D. M. W. Frierson, D. S. Battisti, and J. P. Sachs. Quantifying climate forcings and feedbacks over the last millennium in the CMIP5-PMIP3 models. *Journal of Climate*, 29(3):1161–1178, 2016. ISSN 08948755. doi: 10.1175/JCLI-D-15-0063.1.

G. E. Bodeker, B. Hassler, P. J. Young, and R. W. Portmann. A vertically resolved, global, gap-free ozone database for assessing or constraining global climate model simulations. *Earth System Science Data*, 5(1):31–43, 2013. ISSN 1866-3516. doi: 10.5194/essd-5-31-2013. URL `http://www.earth-syst-sci-data.net/5/31/2013/`.

J. Boé, A. Hall, and X. Qu. Current GCMs ' Unrealistic Negative Feedback in the Arctic. *Journal of Climate*, 22:4682–4695, 2009. doi: 10.1175/2009JCLI2885.1.

S. Bony, R. Colman, V. M. Kattsov, R. P. Allan, C. S. Bretherton, J.-L. Dufresne, A. Hall, S. Hallegatte, M. M. Holland, W. Ingram, D. A. Randall, B. J. Soden, G. Tselioudis, M. J. Webb, S. Bony, R. Colman, V. M. Kattsov, R. P. Allan, C. S. Bretherton, J.-L. Dufresne, A. Hall, S. Hallegatte, M. M. Holland, W. Ingram, D. A. Randall, B. J.

Soden, G. Tselioudis, and M. J. Webb. How Well Do We Understand and Evaluate Climate Change Feedback Processes? *Journal of Climate*, 19(15):3445–3482, 2006. ISSN 0894-8755. doi: 10.1175/JCLI3819.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JCLI3819.1`.

C. S. Bretherton and A. H. Sobel. The Gill Model and the Weak Temperature Gradient Approximation. *Journal of the Atmospheric Sciences*, 60:451–460, 2003. ISSN 0022-4928.

P. M. Caldwell, M. D. Zelinka, K. E. Taylor, and K. Marvel. Quantifying the Sources of Intermodel Spread in Equilibrium Climate Sensitivity. *Journal of Climate*, 29:513–524, 2016. doi: 10.1175/JCLI-D-15-0352.1.

B. R. D. Cess. Global climate change : an investigation of atmospheric feedback mechanisms. *Tellus*, 27(2):193 – 198, 1975.

R. D. Cess, G. L. Potter, J. P. Blanchet, G. J. Boer, A. D. D. E. L. Genio, M. D. Qul, V. Dymnikov, V. Galin, W. L. Gates, S. J. Ghan, J. T. Kiehl, A. A. Lacis, H. L. E. Treut, Z. Li, X. Liang, B. J. Mcavaney, V. P. Meleshko, D. A. Randall, D. A. Sheinin, K. E. Taylor, W. M. Washington, R. T. Wetherald, I. Yagai, and M. Zhang. Intercomparison and Interpretation of Climate Feedback Processes in 19 Atmospheric General Circulation Models. *Journal of Geophysical Research*, 95:16601–16615, 1990.

R. D. Cess, M. H. Zhang, W. J. Ingram, G. L. Potter, V. Alekseev, D. Genio, M. R. Dix, V. Dymnikov, M. Esch, L. D. F. J. R. Fraser, V. Galin, W. L. Gates, J. J. Hack, J. T. Kiehl, H. L. Treut, K. K. Lo, K. E. Taylor, W. Wang, and R. T. Wetheraid. Cloud feedback in atmospheric general circulation models: An update. *Journal of Geophysical Research*, 101:791–794, 1996.

X. Chen and X. Zou. Postlaunch calibration and bias characterization of AMSU-A upper air sounding channels using GPS RO Data. *Journal of Geophysical Research*, pages 1–8, 2014. doi: 10.1002/2013JD021037.

J. Christy, J. Curry, W. Happer, M. Steyn, and D. Titley. Data or Dogma? Promoting Open Inquiry in the Debate over the Magnitude of Human Impact on Earth's Climate, 2015. URL http://www.commerce.senate.gov/public/index.cfm/hearings?ID=CA2ABC55-B1E8-4B7A-AF38-34821F6468F7.

J. R. Christy and W. B. Norris. Satellite and VIZ Radiosonde Intercomparisons for Diagnosis of. *Journal of Atmospheric and Oceanic Technology*, 23:1181–1194, 2006. doi: 10.1175/JTECH1937.1.

J. R. Christy and R. W. Spencer. Comments on A Bias in the Midtropospheric Channel Warm Target Factor on the NOAA-9 Microwave Sounding Unit. *Journal of Atmospheric and Oceanic Technology*, 30(5):1006–1013, 2013. ISSN 0739-0572. doi: 10.1175/JTECH-D-12-00107.1. URL http://journals.ametsoc.org/doi/abs/10.1175/JTECH-D-12-00107.1.

J. R. Christy, R. W. Spencer, and W. D. Braswell. How accurate are satellite thermometers'? *Nature*, 389(September):342–343, 1997.

J. R. Christy, R. W. Spencer, and E. S. Lobl. Analysis of the Merging Procedure for the MSU Daily Temperature Time Series. *Journal of Climate*, 11(8):2016–2041, 1998. ISSN 08948755. doi: 10.1175/1520-0442-11.8.2016.

J. R. Christy, R. W. Spencer, and W. D. Braswell. MSU tropospheric temperatures: Dataset construction and radiosonde comparisons. *Journal of Atmospheric and Oceanic Technology*, 17(9):1153–1170, 2000. ISSN 07390572.

J. R. Christy, R. W. Spencer, W. B. Norris, W. D. Braswell, and D. E. Parker. Error Estimates of Version 5.0 of MSUAMSU Bulk Atmospheric Temperatures. *Journal of Atmospheric and Oceanic Technology*, 20(5):613–629, 2003. ISSN 07390572.

J. R. Christy, W. B. Norris, R. W. Spencer, and J. J. Hnilo. Tropospheric temperature change since 1979 from tropical radiosonde and satellite measurements. *Journal*

of Geophysical Research: Atmospheres, 112(D06102):1–16, 2007. ISSN 01480227. doi: 10.1029/2005JD006881.

J. R. Christy, B. Herman, R. Pielke, P. Klotzbach, R. T. McNider, J. J. Hnilo, R. W. Spencer, T. Chase, and D. Douglass. What Do Observational Datasets Say about Modeled Tropospheric Temperature Trends since 1979? *Remote Sensing*, 2(9):2148–2169, 2010. ISSN 2072-4292. doi: 10.3390/rs2092148. URL `http://www.mdpi.com/2072-4292/2/9/2148/`.

J. R. Christy, R. W. Spencer, and W. B. Norris. International Journal of Remote The role of remote sensing in monitoring global bulk tropospheric temperatures. *International Journal of Remote Sensing*, 32(3):671 – 685, 2011. ISSN 0143-1161. doi: 10.1080/01431161.2010.517803.

I. Cionni, V. Eyring, J. F. Lamarque, W. J. Randel, D. S. Stevenson, F. Wu, G. E. Bodeker, T. G. Shepherd, D. T. Shindell, and D. W. Waugh. Ozone database in support of CMIP5 simulations: results and corresponding radiative forcing. *Atmos. Chem. Phys.*, 11(21):11267–11292, 2011. ISSN 16807316. doi: 10.5194/acp-11-11267-2011.

M. Collins and R. Knutti. Long-term Climate Change: Projections, Commitments and Irreversibility. In T. F. Stocker, D. Qin, K. Plattner, M. M. B. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, editors, *Climate Change 2013: The Physical Science Basis*, pages 1029 – 1136. IPCC, Cambridge Univ. Press, Cambridge, United Kingdon, 2013.

R. Colman. A comparison of climate feedbacks in general circulation models. *Climate Dynamics*, 20:865–873, 2003. doi: 10.1007/s00382-003-0310-z.

T. W. Cronin and M. F. Jansen. Analytic radiative-advective equilibrium as a model for high-latitude climate. *Geophysical Research Letters*, 43:449–457, 2016. doi: 10.1002/2015GL067172.Received.

J. A. Crook, P. M. Forster, and N. Stuber. Spatial Patterns of Modeled Climate Feedback and Contributions to Temperature Response and Polar Amplification. *Journal of Climate*, 24:3575–3592, 2011. doi: 10.1175/2011JCLI3863.1.

A. Dai and K. E. Trenberth. The diurnal cycle and its depiction in the community climate system model. *Journal of Climate*, 17(5):930–951, 2004. ISSN 08948755.

R. P. Damadeo, J. M. Zawodny, and L. W. Thomason. Reevaluation of stratospheric ozone trends from SAGE II data using a simultaneous temporal and spatial analysis. *Atmospheric Chemistry and Physics*, 14(24):13455–13470, 2014. ISSN 16807324. doi: 10.5194/acp-14-13455-2014.

D. P. Dee and S. Uppala. Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quarterly Journal of the Royal ...*, 135(October):1830–1841, 2009. ISSN 00359009. doi: 10.1002/qj. URL `http://onlinelibrary.wiley.com/doi/10.1002/qj.71/abstract`.

A. D. Del Genio, Y. Chen, D. Kim, and M. S. Yao. The MJO Transition from shallow to deep convection in cloudsat/CALIPSO data and GISS GCM simulations. *Journal of Climate*, 25(11):3755–3770, 2012. ISSN 08948755. doi: 10.1175/JCLI-D-11-00384.1.

D. H. Douglass, J. R. Christy, B. D. Pearson, and S. F. Singer. A comparison of tropical temperature trends with model predictions. *International Journal of Climatology*, 28:1693 – 1701, 2008. doi: 10.1002/joc.

J.-L. Dufresne and S. Bony. An Assessment of the Primary Sources of Spread of Global Warming Estimates from Coupled Atmosphere–Ocean Models. *Journal of Climate*, 21 (19):5135–5144, 2008. doi: 10.1175/2008JCLI2239.1.

V. Eyring, J. M. Arblaster, I. Cionni, J. Sedláček, J. Perlwitz, P. J. Young, S. Bekki, D. Bergmann, P. Cameron-Smith, W. J. Collins, G. Faluvegi, K. D. Gottschaldt, L. W. Horowitz, D. E. Kinnison, J. F. Lamarque, D. R. Marsh, D. Saint-Martin, D. T. Shindell,

K. Sudo, S. Szopa, and S. Watanabe. Long-term ozone changes and associated climate impacts in CMIP5 simulations. *Journal of Geophysical Research: Atmospheres*, 118:5029–5060, 2013. ISSN 21698996. doi: 10.1002/jgrd.50316.

N. Feldl and S. Bordoni. Characterizing the Hadley Circulation Response through Regional Climate Feedbacks. *Journal of Climate*, 29:613–622, 2016. doi: 10.1175/JCLI-D-15-0424.1.

N. Feldl and G. H. Roe. Four perspectives on climate feedbacks. *Geophysical Research Letters*, 40:4007–4011, 2013. doi: 10.1002/grl.50711.

N. Feldl, S. Bordoni, and T. M. Merlis. Coupled high-latitude climate feedbacks and their impact on atmospheric heat transport. *Journal of Climate*, submitted, 2016.

R. D. Field, D. Kim, A. N. LeGrande, J. Worden, M. Kelley, and G. A. Schmidt. Evaluating climate model performance in the tropics with retrievals of water isotopic composition from Aura TES. *Geophysical Research Letters*, pages 1–7, 2014. doi: 10.1002/2014GL060572.

T. J. Flannaghan, S. Fueglistaler, I. M. Held, S. Po-Chedley, B. Wyman, and M. Zhao. Tropical Temperature Trends in AMIP Simulations and the Impact of SST Uncertainties. *Journal of Geophysical Research — Atmospheres*, 119, 2014. doi: 10.1002/2014JD022365. URL `http://onlinelibrary.wiley.com/doi/10.1002/2014JD022365/abstract`.

G. M. Flato. Sea-ice and its response to CO 2 forcing as simulated by global climate models. *Climate Dynamics*, 23:229–241, 2004. doi: 10.1007/s00382-004-0436-7.

Q. Fu and C. M. Johanson. Stratospheric influences on MSU-derived tropospheric temperature trends: A direct error analysis. *Journal of Climate*, 17(24):4636–4640, 2004. ISSN 08948755. doi: 10.1175/JCLI-3267.1.

Q. Fu and C. M. Johanson. Satellite-derived vertical dependence of tropical tropospheric temperature trends. *Geophysical Research Letters*, 32(10):1–5, 2005. ISSN 00948276. doi: 10.1029/2004GL022266.

Q. Fu, C. M. Johanson, S. G. Warren, and D. J. Seidel. Contribution of stratospheric cooling to satellite-inferred tropospheric temperature trends. *Nature*, 429(6987):55–58, 2004. ISSN 0028-0836. doi: 10.1038/nature02524. URL http://www.nature.com/nature/journal/v429/n6987/full/nature02524.html.

Q. Fu, C. M. Johanson, J. M. Wallace, and T. Reichler. Enhanced mid-latitude tropospheric warming in satellite measurements. *Science (New York, N.Y.)*, 312(5777):1179, 2006. ISSN 0036-8075. doi: 10.1126/science.1125566. URL http://www.sciencemag.org/content/312/5777/1179.abstract.

Q. Fu, S. Manabe, and C. M. Johanson. On the warming in the tropical upper troposphere: Models versus observations. *Geophysical Research Letters*, 38:1–6, 2011. ISSN 00948276. doi: 10.1029/2012GL053850. URL http://onlinelibrary.wiley.com/doi/10.1029/2011GL048101/abstract.

S. Fueglistaler, C. Radley, and I. M. Held. The distribution of precipitation and the spread in tropical upper tropospheric temperature trends in CMIP5 / AMIP simulations. *Geophysical Research Letters*, 42:6000–6007, 2015. doi: 10.1002/2015GL064966.1. URL http://onlinelibrary.wiley.com/doi/10.1002/2015GL064966/abstract.

J. C. Fyfe, K. Von Salzen, J. N. S. Cole, N. P. Gillett, and J. P. Vernier. Surface response to stratospheric aerosol changes in a coupled atmosphere-ocean model. *Geophysical Research Letters*, 40(3):584–588, 2013. ISSN 00948276. doi: 10.1002/grl.50156.

C. I. Garfinkel, D. W. Waugh, L. D. Oman, L. Wang, and M. M. Hurwitz. Temperature trends in the tropical upper troposphere and lower stratosphere: Connections with sea surface temperatures and implications for water vapor and ozone. *Journal of Geophysical Research: Atmospheres*, 118(17):9658–9672, 2013. ISSN 2169897X. doi: 10.1002/jgrd.50772. URL http://doi.wiley.com/10.1002/jgrd.50772.

M. D. Goldberg, D. S. Crosby, and L. Zhou. The Limb Adjustment of AMSU-A Observations:

Methodology and Validation. *Journal of Applied Meteorology*, 40(1):70–83, 2001. ISSN 0894-8763.

R. G. Graversen, P. L. Langen, and T. Mauritsen. Polar amplification in CCSM4: Contributions from the lapse rate and surface albedo feedbacks. *Journal of Climate*, 27(12): 4433–4450, 2014. ISSN 08948755. doi: 10.1175/JCLI-D-13-00551.1.

N. C. Grody, K. Y. Vinnikov, M. D. Goldberg, J. T. Sullivan, and J. D. Tarpley. Calibration of multisatellite observations for climatic studies: Microwave Sounding Unit (MSU). *Journal of Geophysical Research D: Atmospheres*, 109:1–12, 2004. ISSN 01480227. doi: 10.1029/2004JD005079.

J. Hansen, M. Sato, and R. Ruedy. Long-term changes of the diurnal temperature cycle: implications about mechanisms of global climate change. *Atmospheric Research*, 37(1-3): 175–209, 1995. ISSN 01698095. doi: 10.1016/0169-8095(94)00077-Q.

D. L. Hartmann and K. Larson. An important constraint on tropical cloud - climate feedback. *Geophysical Research Letters*, 29(20):10–13, 2002. doi: 10.1029/2002GL015835.

B. Hassler, P. J. Young, R. W. Portmann, G. E. Bodeker, J. S. Daniel, K. H. Rosenlof, and S. Solomon. Comparison of three vertically resolved ozone data sets: Climatology, trends and radiative forcings. *Atmospheric Chemistry and Physics*, 13(11):5533–5550, 2013. ISSN 16807316. doi: 10.5194/acp-13-5533-2013.

J. M. Haywood, A. Jones, and G. S. Jones. The impact of volcanic eruptions in the period 2000-2013 on global mean temperature trends evaluated in the HadGEM2-ES climate model. *Atmospheric Science Letters*, 15(November 2013):92–96, 2014. ISSN 1530261X. doi: 10.1002/asl2.471. URL `http://doi.wiley.com/10.1002/asl2.471`.

I. M. Held and B. J. Soden. Water Vapor Feedback and Global Warming. *Annu. Rev. Energy Environ*, 25:441 – 475, 2000.

J. W. Hurrell and K. E. Trenberth. Spurious trends in satellite MSU temperatures from merging different satellite records, 1997. ISSN 0028-0836. URL http://www.nature.com/doifinder/10.1038/386164a0.

J. W. Hurrell and K. E. Trenberth. Difficulties in Obtaining Reliable Temperature Trends: Reconciling the Surface and Satellite Microwave Sounding Unit Records. *Journal of Climate*, 11(5):945–967, 1998. ISSN 0894-8755.

J. W. Hurrell, J. J. Hack, D. Shea, J. M. Caron, and J. Rosinski. A new sea surface temperature and sea ice boundary dataset for the community atmosphere model. *Journal of Climate*, 21(19):5145–5153, 2008. ISSN 08948755. doi: 10.1175/2008JCLI2292.1.

Y.-T. Hwang and D. M. W. Frierson. Increasing atmospheric poleward energy transport with global warming. *Geophysical Research Letters*, 37(24), 2010. ISSN 00948276. doi: 10.1029/2010GL045440. URL http://doi.wiley.com/10.1029/2010GL045440.

D. L. Jackson and B. J. Soden. Detection and Correction of Diurnal Sampling Bias in HIRS/2 Brightness Temperatures. *Journal of Atmospheric and Oceanic Technology*, 24(8):1425–1438, 2007. ISSN 0739-0572. doi: 10.1175/JTECH2062.1. URL http://journals.ametsoc.org/doi/abs/10.1175/JTECH2062.1.

C. M. Johanson and Q. Fu. Robustness of tropospheric temperature trends from MSU channels 2 and 4. *Journal of Climate*, 19(17):4234–4242, 2006. ISSN 08948755. doi: 10.1175/JCLI3866.1.

V. O. John and B. J. Soden. Temperature and humidity biases in global climate models and their impact on climate feedbacks. *Geophysical Research Letters*, 34(18):L18704, 2007. ISSN 0094-8276. doi: 10.1029/2007GL030429. URL http://doi.wiley.com/10.1029/2007GL030429.

N. C. Johnson and S.-P. Xie. Changes in the sea surface temperature threshold for tropical

convection. *Nature Geoscience*, 3(12):842–845, 2010. ISSN 1752-0894. doi: 10.1038/ngeo1008. URL http://dx.doi.org/10.1038/ngeo1008.

Y. Kamae, H. Shiogama, M. Watanabe, M. Ishii, H. Ueda, and M. Kimoto. Recent slowdown of tropical upper-tropospheric warming associated with Pacific climate variability. *Geophysical Research Letters*, pages n/a–n/a, 2015. ISSN 00948276. doi: 10.1002/2015GL063608. URL http://doi.wiley.com/10.1002/2015GL063608.

T. R. Karl, S. J. Hassol, C. D. Miller, and W. L. Murray, editors. *Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences*, volume 17. Climate Change Science Program and Subcommittee on Global Change Research, Washington, 2006. ISBN 1234358743. URL www.climatescience.gov/Library/sap/sap1-1/finalreport/sap1-1-final-all.pdf.

E. C. Kent, N. A. Rayner, D. I. Berry, M. Saunby, B. I. Moat, J. J. Kennedy, and D. E. Parker. Global analysis of night marine air temperature and its uncertainty since 1880: The HadNMAT2 data set. *Journal of Geophysical Research Atmospheres*, 118(3):1281–1298, 2013. ISSN 21698996. doi: 10.1002/jgrd.50152.

J. T. Kiehl, J. J. Hack, G. B. Bonan, B. A. Boville, B. P. Briegleb, and D. L. Williamson. Description of the NCAR Community Climate Model (CCSM3). NCAR Tech. Note NCAR/TN-4201+STR. Technical report, National Center for Atmospheric Research, Boulder, CO, 1996.

D. Kim, Y. S. Jang, D. H. Kim, Y. H. Kim, M. Watanabe, F. F. Jin, and J. S. Kug. El Niño-Southern Oscillation sensitivity to cumulus entrainment in a coupled general circulation model. *Journal of Geophysical Research: Atmospheres*, 116(22):1–9, 2011. ISSN 01480227. doi: 10.1029/2011JD016526.

D. Kim, A. H. Sobel, A. D. Del Genio, Y. Chen, S. J. Camargo, M.-S. Yao, M. Kelley, and L. Nazarenko. The Tropical Subseasonal Variability Simulated in the NASA GISS

General Circulation Model. *Journal of Climate*, 25(13):4641–4659, 2012. ISSN 0894-8755. doi: 10.1175/JCLI-D-11-00447.1. URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-11-00447.1.

A. Kottayil, V. O. John, and S. A. Buehler. Correcting diurnal cycle aliasing in satellite microwave humidity sounder measurements. *Journal of Geophysical Research: Atmospheres*, 118(1):101–113, 2013. ISSN 2169897X. doi: 10.1029/2012JD018545. URL http://doi.wiley.com/10.1029/2012JD018545.

J. R. Lanzante. A Cautionary Note on the Use of Error Bars. *Journal of Climate*, pages 3699–3703, 2005.

J. Le Marshall, J. Jung, J. Derber, M. Chahine, R. Treadon, S. J. Lord, M. Goldberg, W. Wolf, H. C. Liu, J. Joiner, J. Woollen, R. Todling, P. van Delst, and Y. Tahara. Improving Global Analysis and Forecasting with AIRS. *Bulletin of the American Meteorological Society*, 87(7):891–894, 2006. ISSN 0003-0007. doi: 10.1175/BAMS-87-7-891.

O. Lecomte, T. Fichefet, M. Vancoppenolle, F. Domine, F. Massonnet, P. Mathiot, S. Morin, and P. Y. Barriat. On the formulation of snow thermal conductivity in large-scale sea ice models. *Journal of Advances in Modeling Earth Systems*, 5:1–16, 2013. doi: 10.1002/jame.20039.

R. Lei, N. Li, P. Heil, B. Cheng, Z. Zhang, and B. Sun. Multiyear sea ice thermal regimes and oceanic heat flux derived from an ice mass balance buoy in the Arctic Ocean. *Journal of Geophysical Research : Oceans*, 119(January 2013):537–547, 2014. doi: 10.1002/2012JC008731.

P. Lin and Q. Fu. Changes in various branches of the Brewer-Dobson circulation from an ensemble of chemistry climate models. *Journal of Geophysical Research: Atmospheres*, 118(1):73–84, 2013. ISSN 21698996. doi: 10.1029/2012JD018813.

P. Lin, Q. Fu, S. Solomon, and J. M. Wallace. Temperature trend patterns in Southern Hemisphere high latitudes: Novel indicators of stratospheric change. *Journal of Climate*, 22(23):6325–6341, 2009. ISSN 08948755. doi: 10.1175/2009JCLI2971.1.

A. V. Lindfors, I. a. Mackenzie, S. F. B. Tett, and L. Shi. Climatological diurnal cycles in clear-sky brightness temperatures from the high-resolution infrared radiation sounder (HIRS). *Journal of Atmospheric and Oceanic Technology*, 28:1199–1205, 2011. ISSN 07390572. doi: 10.1175/JTECH-D-11-00093.1.

A. C. Lorenc and R. T. Marriott. Forecast sensitivity to observations in the Met Office Global numerical weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, 140(January):209–224, 2014. ISSN 00359009. doi: 10.1002/qj.2122.

Q. Lu and W. Bell. Characterizing Channel Center Frequencies in AMSU-A and MSU Microwave Sounding Instruments. *Journal of Atmospheric and Oceanic Technology*, 31 (8):1713–1732, 2014. ISSN 0739-0572. doi: 10.1175/JTECH-D-13-00136.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JTECH-D-13-00136.1`.

I. A. MacKenzie, S. F. B. Tett, and A. V. Lindfors. Climate ModelSimulated Diurnal Cycles in HIRS Clear-Sky Brightness Temperatures. *Journal of Climate*, 25(17):5845–5863, 2012. ISSN 0894-8755. doi: 10.1175/JCLI-D-11-00552.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-11-00552.1`.

S. Manabe and R. T. Wetherald. The Effects of Doubling the CO2 Concentration on the Climate of a General Circulation Model. *Journal of Atmospheric Sciences*, 32(1):3 – 15, 1975.

S. Manabe, K. Bryan, and M. J. Spelman. Transient Response of a Global Ocean-Atmosphere Model to a Doubling of Atmospheric Carbon Dioxide. *Journal of Physical Oceanography*, 20:722 – 749, 1990.

S. Manabe, R. J. Stouffer, M. J. Spelman, and K. Bryan. Transient Responses of a Coupled Ocean-Atmosphere Model to Gradual Changes of Atmospheric CO2. Part I: Annual Mean Response. *Journal of Climate*, 4:785 – 818, 1991.

T. Mauritsen and B. Stevens. Missing iris effect as a possible cause of muted hydrological change and high climate sensitivity in models. *Nature Geoscience*, (April):8–13, 2015. ISSN 1752-0894. doi: 10.1038/ngeo2414. URL `http://www.nature.com/doifinder/10.1038/ngeo2414`.

R. Mckitrick, S. Mcintyre, and C. Herman. Panel and multivariate methods for tests of trend equivalence in climate data series. *Atmospheric Science Letters*, 11(4):270–277, 2010. ISSN 1530261X. doi: 10.1002/asl.290.

C. a. Mears and F. J. Wentz. The effect of diurnal correction on satellite-derived lower tropospheric temperature. *Science (New York, N.Y.)*, 309(5740):1548–51, 2005. ISSN 1095-9203. doi: 10.1126/science.1114772. URL `http://www.ncbi.nlm.nih.gov/pubmed/16141071`.

C. a. Mears and F. J. Wentz. Construction of the remote sensing systems V3.2 atmospheric temperature records from the MSU and AMSU microwave sounders. *Journal of Atmospheric and Oceanic Technology*, 26(6):1040–1056, 2009a. ISSN 07390572. doi: 10.1175/2008JTECHA1176.1.

C. a. Mears and F. J. Wentz. Construction of the RSS V3.2 Lower-Tropospheric Temperature Dataset from the MSU and AMSU Microwave Sounders. *Journal of Atmospheric and Oceanic Technology*, 26:1493 – 1509, 2009b. ISSN 07390572. doi: 10.1175/2008JTECHA1176.1.

C. A. Mears and F. J. Wentz. Sensitivity of satellite-derived tropospheric temperature trends to the diurnal cycle adjustment. *Journal of Climate*, 5:3629–3646, 2016. ISSN 0894-8755. doi: 10.1175/JCLI-D-15-0744.1.

C. a. Mears, M. C. Schabel, and F. J. Wentz. A reanalysis of the MSU channel 2 tropospheric temperature record. *Journal of Climate*, 16(22):3650–3664, 2003. ISSN 08948755.

C. A. Mears, F. J. Wentz, P. Thorne, and D. Bernie. Assessing uncertainty in estimates of atmospheric temperature changes from MSU and AMSU using a Monte-Carlo estimation technique. *Journal of Geophysical Research — Atmospheres*, 116(8):1–16, 2011. ISSN 01480227. doi: 10.1029/2010JD014954. URL http://onlinelibrary.wiley.com/doi/10.1029/2010JD014954/abstract.

C. A. Mears, F. J. Wentz, and P. W. Thorne. Assessing the value of microwave sounding unit-radiosonde comparisons in ascertaining errors in climate data records of tropospheric temperatures. *Journal of Geophysical Research: Atmospheres*, 117:D19103, 2012. ISSN 01480227. doi: 10.1029/2012JD017710.

B. Medeiros, C. Deser, R. A. Tomas, and J. E. Kay. Arctic Inversion Strength in Climate Models. *Journal of Climate*, 24:4733–4740, 2011. doi: 10.1175/2011JCLI3968.1.

G. A. Meehl, J. M. Arblaster, C. M. Bitz, C. T. Y. Chung, and H. Teng. Antarctic sea-ice expansion between 2000 and 2014 driven by tropical Pacific decadal climate variability. *Nature Geoscience*, (July):1–7, 2016. ISSN 1752-0894. doi: 10.1038/ngeo2751. URL http://www.nature.com/doifinder/10.1038/ngeo2751.

W. Meier, F. Fetterer, M. Savoie, S. Mallory, R. Duerr, and J. Stroeve. NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 2, 2013. URL http://dx.doi.org/10.7265/N55M63M1.

D. M. Mitchell, P. W. Thorne, P. a. Stott, and L. J. Gray. Revisiting the controversial issue of tropical tropospheric temperature trends. *Geophysical Research Letters*, 40(11):2801–2806, 2013a. ISSN 00948276. doi: 10.1002/grl.50465. URL http://doi.wiley.com/10.1002/grl.50465.

D. M. Mitchell, P. W. Thorne, P. a. Stott, and L. J. Gray. Revisiting the controversial issue of tropical tropospheric temperature trends. *Geophysical Research Letters*, 40(11):2801–2806, 2013b. ISSN 00948276. doi: 10.1002/grl.50465. URL `http://doi.wiley.com/10.1002/grl.50465`.

T. Mo. A study of the NOAA-15 AMSU-A brightness temperatures from 1998 through 2007. *Journal of Geophysical Research: Atmospheres*, 114(11):1–10, 2009. ISSN 01480227. doi: 10.1029/2008JD011267.

C. P. Morice, J. J. Kennedy, N. a. Rayner, and P. D. Jones. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *Journal of Geophysical Research: Atmospheres*, 117(8):1–22, 2012. ISSN 01480227. doi: 10.1029/2011JD017187. URL `http://onlinelibrary.wiley.com/doi/10.1029/2011JD017187/abstract`.

T. M. Pavelsky, J. Boé, A. Hall, and E. J. Fetzer. Atmospheric inversion strength over polar oceans in winter regulated by sea ice. *Climate Dynamics*, 36:945–955, 2011. doi: 10.1007/s00382-010-0756-8.

A. E. Payne, M. F. Jansen, and T. W. Cronin. Conceptual model analysis of the influence of temperature feedbacks on polar amplification. *Geophysical Research Letters*, 42:9561–9570, 2015. doi: 10.1002/2015GL065889.

G. Peng, W. N. Meier, D. J. Scott, M. H. Savoie, and N. Snow. A long-term and reproducible passive microwave sea ice concentration data record for climate studies and monitoring. *Earth System Science Data*, 5:311–318, 2013. doi: 10.5194/essd-5-311-2013.

F. Pithan and T. Mauritsen. Arctic amplification dominated by temperature feedbacks in contemporary climate models. *Nature Geoscience*, 7(3):181–184, 2014. ISSN 1752-0894. doi: 10.1038/ngeo2071.

S. Po-Chedley. *Reconciling tropospheric temperature trends from the Microwave Sounding Unit*. M.s. thesis, University of Washington, 2012.

S. Po-Chedley and Q. Fu. Discrepancies in tropical upper tropospheric warming between atmospheric circulation models and satellites. *Environmental Research Letters*, 7:044018, 2012a. ISSN 1748-9326. doi: 10.1088/1748-9326/7/4/044018. URL `http://iopscience.iop.org/article/10.1088/1748-9326/7/4/044018/pdf`.

S. Po-Chedley and Q. Fu. A Bias in the Midtropospheric Channel Warm Target Factor on the NOAA-9 Microwave Sounding Unit. *Journal of Atmospheric and Oceanic Technology*, 29(5):646–652, 2012b. ISSN 0739-0572. doi: 10.1175/JTECH-D-11-00147.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JTECH-D-11-00147.1`.

S. Po-Chedley and Q. Fu. Reply to "Comments on 'A Bias in the Midtropospheric Channel Warm Target Factor on the NOAA-9 Microwave Sounding Unit'". *Journal of Atmospheric and Oceanic Technology*, 30(5):1014–1020, 2013. doi: 10.1175/JTECH-D-12-00131.1.

S. Po-Chedley, T. J. Thorsen, and Q. Fu. Removing Diurnal Cycle Contamination in Satellite-Derived Tropospheric Temperatures : Understanding Tropical Tropospheric Trend Discrepancies. *Journal of Climate*, 28:2274–2290, 2015. doi: 10.1175/JCLI-D-13-00767.1.

L. M. Polvani and S. Solomon. The signature of ozone depletion on tropical temperature trends, as revealed by their seasonal cycle in model integrations with single forcings. *Journal of Geophysical Research: Atmospheres*, 117(D17), 2012. ISSN 01480227. doi: 10.1029/2012JD017719. URL `http://doi.wiley.com/10.1029/2012JD017719`.

L. R. Pralungo and L. Haimberger. New estimates of tropical mean temperature trend profiles from zonal mean historical radiosonde and pilot balloon wind shear observations. *Journal of Geophysical Research — Atmospheres*, 120:3700–3713, 2015. ISSN 2169-8996.

doi: 10.1002/2014JD022664. URL http://onlinelibrary.wiley.com/doi/10.1002/2014JD022664/full.

V. Ramanathan. Interactions between Ice-Albedo, Lapse-Rate and Cloud-Top Feedbacks: An Analysis of the Nonlinear Response of a GCM Model. *Journal of Atmospheric Sciences*, 34:1885 – 1897, 1977.

R. M. Randall and B. M. Herman. Using limited time period trends as a means to determine attribution of discrepancies in microwave sounding unit-derived tropospheric temperature time series. *Journal of Geophysical Research: Atmospheres*, 113(D5), 2008. ISSN 0148-0227.

W. J. Randel and F. Wu. Biases in stratospheric and tropospheric temperature trends derived from historical radiosonde data. *Journal of Climate*, 19(10):2094–2104, 2006. ISSN 08948755. doi: 10.1175/JCLI3717.1.

W. J. Randel and F. Wu. A stratospheric ozone profile data set for 19792005: Variability, trends, and comparisons with column ozone data. *Journal of Geophysical Research*, 112 (D6):D06313, 2007. ISSN 0148-0227. doi: 10.1029/2006JD007339. URL http://doi.wiley.com/10.1029/2006JD007339.

N. A. Rayner, D. Parker, E. Horton, C. Folland, L. Alexander, D. Rowell, E. Kent, and A. Kaplan. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, 108(D14): 4407, 2003. ISSN 0148-0227. doi: 10.1029/2002JD002670.

T. Reichler, M. Dameris, and R. Sausen. Determining the tropopause height from gridded data. *Geophysical Research Letters*, 30(20):1–5, 2003. doi: 10.1029/2003GL018240.

G. H. Roe, N. Feldl, K. C. Armour, Y.-T. Hwang, and D. M. W. Frierson. The remote impacts of climate feedbacks on regional climate predictability. *Nature Geoscience*, 8(2):

135–139, 2015. ISSN 1752-0894. doi: 10.1038/ngeo2346. URL http://dx.doi.org/10.1038/ngeo2346.

B. D. Santer, T. M. L. Wigley, C. Mears, F. J. Wentz, S. a. Klein, D. J. Seidel, K. E. Taylor, P. W. Thorne, M. F. Wehner, P. J. Gleckler, J. S. Boyle, W. D. Collins, K. W. Dixon, C. Doutriaux, M. Free, Q. Fu, J. E. Hansen, G. S. Jones, R. Ruedy, T. R. Karl, J. R. Lanzante, G. a. Meehl, V. Ramaswamy, G. Russell, and G. a. Schmidt. Amplification of surface temperature trends and variability in the tropical atmosphere. *Science (New York, N.Y.)*, 309(5740):1551–1556, 2005. ISSN 0036-8075. doi: 10.1126/science.1114867. URL http://www.sciencemag.org/content/309/5740/1551.abstract.

B. D. Santer, P. W. Thorne, L. Haimberger, K. E. Taylor, T. M. L. Wigley, J. R. Lanzante, S. Solomon, M. Free, P. J. Gleckler, P. D. Jones, T. R. Karl, S. A. Klein, C. Mears, D. Nychka, G. A. Schmidt, S. C. Sherwood, and F. J. Wentz. Consistency of modelled and observed temperature trends in the tropical troposphere. *International Journal of Climatology*, pages 1703 – 1722, 2008. doi: 10.1002/joc.1756.

B. D. Santer, C. Bonfils, J. F. Painter, M. D. Zelinka, C. Mears, S. Solomon, G. a. Schmidt, J. C. Fyfe, J. N. S. Cole, L. Nazarenko, K. E. Taylor, and F. J. Wentz. Volcanic contribution to decadal changes in tropospheric temperature. *Nature Geoscience*, 7(3):185–189, 2014. ISSN 1752-0894. doi: 10.1038/NGEO2098. URL http://www.nature.com/ngeo/journal/v7/n3/full/ngeo2098.html.

B. D. Santer, S. Solomon, G. Pallota, C. Mears, S. Po-Chedley, Q. Fu, F. Wentz, C.-Z. Zou, J. Painter, I. Cvijanovic, and C. Bonfils. Comparing Tropospheric Warming In Climate Models and Satellite Data. *Journal of Climate*, in press, 2016.

M. Sato, J. E. Hansen, M. P. Mccormick, and J. B. Pollack. Stratospheric aerosol optical depths, 1850-1990. *Journal of Geophysical Research*, 98(D12):22987–22994, 1993. ISSN 0148-0227. doi: 10.1029/93JD02553.

G. A. Schmidt, M. Kelley, L. Nazarenko, R. Ruedy, G. L. Russell, I. Aleinov, M. Bauer, S. E. Bauer, M. K. Bhat, R. Bleck, V. Canuto, Y.-h. Chen, Y. Cheng, T. L. Clune, A. D. Genio, R. D. Fainchtein, G. Faluvegi, J. E. Hansen, R. J. Healy, N. Y. Kiang, D. Koch, A. A. Lacis, A. N. Legrande, J. Lerner, K. K. Lo, E. E. Matthews, S. Menon, R. L. Miller, V. Oinas, and A. O. Oloso. Journal of Advances in Modeling Earth Systems contributions to the CMIP5 archive. *Journal of Advances in Modeling Earth Systems*, 6:141–184, 2014. ISSN 19422466. doi: 10.1002/2013MS000265.Received.

J. A. Screen and I. Simmonds. The central role of diminishing sea ice in recent Arctic temperature amplification. *Nature*, 464(7293):1334–1337, 2010. ISSN 0028-0836. doi: 10.1038/nature09051. URL `http://dx.doi.org/10.1038/nature09051`.

J. A. Screen, C. Deser, and I. Simmonds. Local and remote controls on observed Arctic warming. *Geophysical Research Letters*, 39(May):1–5, 2012. doi: 10.1029/2012GL051598.

D. J. Seidel, M. Free, and J. Wang. Diurnal cycle of upper-air temperature estimated from radiosondes. *Journal of Geophysical Research D: Atmospheres*, 110(9):1–13, 2005. ISSN 01480227. doi: 10.1029/2004JD005526.

D. J. Seidel, M. Free, and J. S. Wang. Reexamining the warming in the tropical upper troposphere: Models versus radiosonde observations. *Geophysical Research Letters*, 39 (22), 2012. ISSN 00948276. doi: 10.1029/2012GL053850. URL `http://doi.wiley.com/10.1029/2012GL053850`.

K. M. Shell, J. T. Kiehl, and C. A. Shields. Using the Radiative Kernel Technique to Calculate Climate Feedbacks in NCAR ' s. *Journal of Climate*, 21:2269–2282, 2008. doi: 10.1175/2007JCLI2044.1.

S. C. Sherwood and N. Nishant. Atmospheric changes through 2012 as shown by iteratively homogenized radiosonde temperature and wind data (IUKv2). *Environmental Research*

Letters, 10(5):054007, 2015. ISSN 1748-9326. doi: 10.1088/1748-9326/10/5/054007. URL http://iopscience.iop.org/article/10.1088/1748-9326/10/5/054007.

S. C. Sherwood, J. R. Lanzante, and C. L. Meyer. Radiosonde daytime biases and late-20th century warming. *Science (New York, N.Y.)*, 309(5740):1556–1559, 2005. ISSN 0036-8075. doi: 10.1126/science.1115640.

S. C. Sherwood, C. L. Meyer, R. J. Allen, and H. a. Titchner. Robust tropospheric warming revealed by iteratively homogenized radiosonde data. *Journal of Climate*, 21(20):5336–5350, 2008. ISSN 08948755. doi: 10.1175/2008JCLI2320.1.

A. H. Sobel, I. M. Held, and C. S. Bretherton. The ENSO signal in tropical tropospheric temperature. *Journal of Climate*, 15(18):2702–2706, 2002. ISSN 08948755.

B. J. Soden and I. M. Held. An Assessment of Climate Feedbacks in Coupled Ocean–Atmosphere Models. *Journal of Climate*, 19(14):3354–3360, 2006. doi: 10.1175/JCLI3799.1.

B. J. Soden, I. M. Held, R. Colman, K. M. Shell, J. T. Kiehl, and C. A. Shields. Quantifying Climate Feedbacks Using Radiative Kernels. *Journal of Climate*, 21(14):3504–3520, 2008. ISSN 0894-8755. doi: 10.1175/2007JCLI2110.1.

S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, editors. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, 2007.

S. Solomon, J. S. Daniel, R. R. Neely, J.-P. Vernier, E. G. Dutton, and L. W. Thomason. The Persistently Variable "Background" Stratospheric Aerosol Layer and Global Climate Change. *Science*, 333(6044):866–870, 2011. ISSN 0036-8075. doi: 10.1126/science.1206027. URL http://www.sciencemag.org/cgi/doi/10.1126/science.1206027.

S. Solomon, P. J. Young, and B. Hassler. Uncertainties in the evolution of stratospheric ozone and implications for recent temperature changes in the tropical lower stratosphere. *Geophysical Research Letters*, 39(17):2–5, 2013. ISSN 00948276. doi: 10.1029/2012GL052723.

R. W. Spencer and J. R. Christy. Precise of Monitoring Trends from Global Satellites Temperature. *Science*, 247(4950):1558–1562, 1990. URL `http://www.sciencemag.org/content/247/4950/1558.short`.

R. W. Spencer and J. R. Christy. Precision and Radiosonde Validation of Satellite Gridpoint Temperature Anomalies. Part I: MSU Channel 2. *Journal of Climate*, 5:847–857, 1992.

A. K. Steiner, B. C. Lackner, F. Ladstdter, B. Scherllin-Pirscher, U. Foelsche, and G. Kirchengast. GPS radio occultation for climate monitoring and change detection. *Radio Science*, 46(6):1–17, 2011. ISSN 00486604. doi: 10.1029/2010RS004614. URL `http://onlinelibrary.wiley.com/doi/10.1029/2010RS004614/abstract`.

R. J. Stouffer, S. Manabe, and K. Bryan. Interhemispheric asymmetry in climate response to a gradual increase of atmospheric CO2. *Nature*, 342(6250):660–662, 1989. ISSN 0028-0836. doi: 10.1038/342660a0.

R. T. Sutton, B. Dong, and J. M. Gregory. Land/sea warming ratio in response to climate change: IPCC AR4 model results and comparison with observations. *Geophysical Research Letters*, 34(2):2–6, 2007. ISSN 00948276. doi: 10.1029/2006GL028164.

K. E. Taylor, R. J. Stouffer, and G. a. Meehl. An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4):485–498, 2012. ISSN 00030007. doi: 10.1175/BAMS-D-11-00094.1.

D. W. J. Thompson, D. J. Seidel, W. J. Randel, C.-Z. Zou, A. H. Butler, C. Mears, A. Osso, C. Long, and R. Lin. The mystery of recent stratospheric temperature trends. *Nature*, 491 (7426):692–7, 2012. ISSN 1476-4687. doi: 10.1038/nature11579. URL `http://www.ncbi.nlm.nih.gov/pubmed/23192146`.

P. W. Thorne, D. E. Parker, B. D. Santer, M. P. McCarthy, D. M. H. Sexton, M. J. Webb, J. M. Murphy, M. Collins, H. A. Titchner, and G. S. Jones. Tropical vertical temperature trends: A real discrepancy? *Geophysical Research Letters*, 34(16):1–5, 2007. ISSN 00948276. doi: 10.1029/2007GL029875.

P. W. Thorne, J. R. Lanzante, T. C. Peterson, D. J. Seidel, and K. P. Shine. Tropospheric temperature trends: History of an ongoing controversy. *Wiley Interdisciplinary Reviews: Climate Change*, 2(1):66–88, 2011. ISSN 17577799. doi: 10.1002/wcc.80.

B. Tian, B. J. Soden, and X. Wu. Diurnal cycle of convection, clouds, and water vapor in the tropical upper troposphere: Satellites versus a general circulation model. *Journal of Geophysical Research D: Atmospheres*, 109(10):1–16, 2004. ISSN 01480227. doi: 10.1029/2003JD004117.

H. A. Titchner, P. W. Thorne, M. P. McCarthy, S. F. B. Tett, L. Haimberger, and D. E. Parker. Critically reassessing tropospheric temperature trends from radiosondes using realistic validation experiments. *Journal of Climate*, 22(3):465–485, 2009. ISSN 08948755. doi: 10.1175/2008JCLI2419.1.

H. Tokinaga, S.-P. P. Xie, C. Deser, Y. Kosaka, and Y. M. Okumura. Slowdown of the Walker circulation driven by tropical Indo-Pacific warming. *Nature Research Letter*, 491(7424):439–443, 2012. ISSN 0028-0836. doi: 10.1038/nature11576. URL `http://www.nature.com/nature/journal/v491/n7424/full/nature11576.html?WT.ec{_}id=NATURE-20121115$\delimiter"026E30F$nhttp://dx.doi.org/10.1038/nature11576$\delimiter"026E30F$nhttp://usrexp-sandbox.nature.com/nature/journal/v491/n7424/full/nature11576.html`.

L. Tomassini, A. Voigt, and B. Stevens. On the connection between tropical circulation, convective mixing, and climate sensitivity. *Quarterly Journal of the Royal Meteorological Society*, 141(689):1404–1416, 2015. ISSN 1477870X. doi: 10.1002/qj.2450.

G. A. Vecchi, S. Fueglistaler, I. M. Held, T. R. Knutson, and M. Zhao. Impacts of Atmospheric Temperature Trends on Tropical Cyclone Activity. *Journal of Climate*, 26: 3877–3891, 2013. doi: 10.1175/JCLI-D-12-00503.1.

J. P. Vernier, L. W. Thomason, J. P. Pommereau, A. Bourassa, J. Pelon, A. Garnier, A. Hauchecorne, L. Blanot, C. Trepte, D. Degenstein, and F. Vargas. Major influence of tropical volcanic eruptions on the stratospheric aerosol layer during the last decade. *Geophysical Research Letters*, 38(12):1–8, 2011. ISSN 00948276. doi: 10.1029/2011GL047563.

J. M. Wallace. Effect of deep convection on the regulation of tropical sea surface temperatures. *Nature*, 356:230 – 231, 1992.

J. M. Wallace and Coauthors. *Reconciling observations of global temperature change*. National Academy Press, 2000. ISBN 0309594006. doi: 10.1029/2001GL014074.

W. Wang and C.-Z. Zou. AMSU-A-Only Atmospheric Temperature Data Records from the Lower Troposphere to the Top of the Stratosphere. *Journal of Atmospheric and Oceanic Technology*, 31:808–825, 2014. ISSN 0739-0572. doi: 10.1175/JTECH-D-13-00134.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JTECH-D-13-00134.1`.

S. Watanabe, T. Hajima, K. Sudo, T. Nagashima, T. Takemura, H. Okajima, T. Nozawa, H. Kawase, M. Abe, T. Yokohata, T. Ise, H. Sato, E. Kato, K. Takata, S. Emori, and M. Kawamiya. MIROC-ESM: model description and basic results of CMIP5-20c3m experiments. *Geoscientific Model Development Discussions*, 4(2):1063–1128, 2011. ISSN 1991-9603. doi: 10.5194/gmdd-4-1063-2011.

F. J. Wentz and T. Meissner. AMSR ocean algorithm, version 2. Remote Sensing Systems Tech. Rep. 121599A-1. Technical report, Remote Sensing Systems, 1999.

F. J. Wentz and M. Schabel. Effects of orbital decay on satellite-derived lower-tropospheric temperature trends. *Nature*, 394(6694):661–664, 1998. doi: 10.1038/29267.

M. Winton. Amplified Arctic climate change : What does surface albedo feedback have to do with it ? *Geophysical Research Letters*, 33:1–4, 2006. doi: 10.1029/2005GL025244.

G. Y. Yang and J. Slingo. The diurnal cycle in the tropics. *Monthly Weather Review*, 129 (1994):784–801, 2001. ISSN 0027-0644.

M. S. Yao and Y. Cheng. Cloud simulations in response to turbulence parameterizations in the GISS model E GCM. *Journal of Climate*, 25(14):4963–4974, 2012. ISSN 08948755. doi: 10.1175/JCLI-D-11-00399.1.

P. J. Young, S. M. Davis, B. Hassler, S. Solomon, and K. H. Rosenlof. Modeling the climate impact of Southern Hemisphere ozone depletion: The importance of the ozone data set. *Geophysical Research Letters*, pages 9033–9039, 2014. ISSN 00948276. doi: 10.1002/2014GL061738.Received.

E. Yulaeva and J. M. Wallace. The signature of ENSO in global temperature and precipitation fields derived from the microwave sounding unit. *Journal of Climate*, 7(11):1719–1736, 1994. ISSN 08948755. doi: 10.1175/1520-0442(1994)007⟨1719:TSOEIG⟩2.0.CO;2.

M. Zhao, J.-C. Golaz, I. M. Held, V. Ramaswamy, S.-J. Lin, Y. Ming, P. Ginoux, B. Wyman, L. J. Donner, D. Paynter, and H. Guo. Uncertainty in model climate sensitivity traced to representations of cumulus precipitation microphysics. *Journal of Climate*, 29:543 – 560, 2016. ISSN 0894-8755. doi: 10.1175/JCLI-D-15-0191.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-15-0191.1`.

C. Z. Zou and W. Wang. Intersatellite calibration of AMSU-A observations for weather and climate applications. *Journal of Geophysical Research: Atmospheres*, 116(May):D23113, 2011. ISSN 01480227. doi: 10.1029/2011JD016205. URL `http://onlinelibrary.wiley.com/doi/10.1029/2011JD016205/full`.

C.-Z. Zou and W. Wang. MSU/AMSU radiance fundamental climate data record calibrated using simultaneous nadir overpasses. Climate Data Record Program Tech.

Rep. CDRP-0015. Technical report, National Oceanic and Atmospheric Administration, 2013. URL http://www.star.nesdis.noaa.gov/smcd/emb/mscat/documents/MSU{_}AMSU{_}CATBD{_}V1.0.pdf.

C. Z. Zou and W. H. Wang. Stability of the MSU-Derived Atmospheric Temperature Trend. *Journal of Atmospheric and Oceanic Technology*, 27(11):1960–1971, 2010. ISSN 0739-0572.

C.-Z. Zou, M. D. Goldberg, Z. Cheng, N. C. Grody, J. T. Sullivan, C. Cao, and D. Tarpley. Recalibration of microwave sounding unit for climate studies using simultaneous nadir overpasses. *Journal of Geophysical Research*, 111(D19):D19114, 2006. ISSN 01480227. doi: 10.1029/2005JD006798. URL http://doi.wiley.com/10.1029/2005JD006798$\delimiter"026E30F$npapers3://publication/doi/10.1029/2005JD006798.

C.-Z. Zou, M. Gao, and M. D. Goldberg. Error Structure and Atmospheric Temperature Trends in Observations from the Microwave Sounding Unit. *Journal of Climate*, 22(7):1661–1681, 2009. ISSN 0894-8755. doi: 10.1175/2008JCLI2233.1. URL http://journals.ametsoc.org/doi/abs/10.1175/2008JCLI2233.1$\delimiter"026E30F$npapers3://publication/doi/10.1175/2008JCLI2233.1.