# Climate Response to Solar Variation: Cyclic and Secular

Jiansong Zhou

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

Doctor of Philosophy

University of Washington

2012

Reading Committee:

Ka-Kit Tung, Chair

Christopher S. Bretherton

Randall J. LeVeque

John M. Wallace

Program Authorized to Offer Degree:
Applied Mathematics

University of Washington

**Abstract**

Climate Response to Solar Variation: Cyclic and Secular

Jiansong Zhou

Chair of the Supervisory Committee:
Professor Ka-Kit Tung
Department of Applied Mathematics

The radiation emitted by the Sun varies both cyclically and secularly. We first study in this thesis the response of the Earth's temperature to the 11-year solar cycle at the surface and in the troposphere. Then we study multi-century long climate response and attribute it to various causes, including the Sun. The solar variation is one of the many factors that cause climate change on Earth. However, the solar cycle response in the atmosphere has not been consistently identified due to its small amplitude compared to other climate variability. A comprehensive study of extracting the solar cycle signal from various global temperature records is presented in this dissertation. Specifically, a clean decadal solar cycle response is proved statistically to exist in the 150-year-long sea surface temperature with the contamination by other climate phenomena quantified small. The existence and the statistical significance of the zonal-mean global temperature response to the 11-year solar cycle are also established throughout the troposphere and parts of the lower stratosphere, which reveals a spatial pattern consistent with a "bottom-up" mechanism that explains the effects of solar forcing on the atmosphere. Evidence is also found that the Sun is not to blame for global warming. These observational findings can be used for calibrating models to improve their ability to make more reliable predictions on Earth's climate change.

## TABLE OF CONTENTS

## LIST OF FIGURES

## ACKNOWLEDGMENTS

The author wishes to express sincere appreciation to his advisor Prof. Ka-Kit Tung for all the support he has provided. The author would also like to thank his committee members, Christopher Bretherton, Randall LeVeque, John Wallace and Peter Guttorp for help and advice. The author also acknowledges insightful communications about the work included in this thesis with Charles Camp, Patrick Fischer, King-Fai Li, Zhaohua Wu and Norden Huang.

In addition, the author wishes to acknowledge the financial support through five years of this work, funded by National Science Foundation (NSF), Climate Dynamics Program, under Grants ATM-0332364 and ATM-0808375, by Division of Mathematical Sciences under Grant DMS-0940342, and by National Aeronautical and Space Administration (NASA) under Grant NNX11AC75G.

## DEDICATION

to my dear wife, Chun-Hsuan

# Chapter 1

# INTRODUCTION

The total solar irradiance (TSI), also called the solar constant, is the amount of solar radiative energy incident on the top of Earth's atmosphere integrated over all wavelengths. Solar energy is the most important driving force for the atmospheric circulation and climate. The uneven distribution of incoming solar radiation in both space and time, for example, between the tropics and poles, between the land and oceans, and between seasons, etc., contributes significantly to the variations of the Earth's temperature and other climate variability. Despite its name, there is some small variation of the solar constant in time [81, 82, 159] both in cycles and secularly. Since 1978 orbiting satellites have measured that the TSI averages 1366 $\mathrm{Wm}^{-2}$ and fluctuates quasi-periodically on a roughly 11-year time scale with an amplitude of about 0.1% from peak to trough during the course of a sunspot cycle. In its active phase, called the solar maximum (max), the Sun has more dark sunspots and accompanying bright faculae. The magnitude and indeed even the sign of this variation were uncertain before that time, when direct measurement of the solar constant above the earth's atmosphere were not feasible. But using sunspot and other proxy indices, the variation of the TSI can be extended using solar models back to the seventeenth century.

The response of the Earth's climate to the variable solar forcing is controversial, especially with regards to the temperature at the surface. Since the terrestrial response is believed to be very small compared to other climate variability and noise, there have even been doubts that such a small signal can be extracted in data, especially in the currently available short-term meteorological observations. Historically there were debates as to whether the earth was warmer or colder during the solar max as compared to the solar minimum (min). Although some researchers have previously found using modern temperature records that the global-mean temperature is warmer during solar max (see, e.g., [11, 145]), there were controversial reports that perhaps in an earlier epoch the response was opposite. For

example, the literature reviewed by [61, chapter 5] suggests that the surface temperature is negatively correlated with the TSI during the period 1800-1920 and positively correlated from 1920 to the present, and a sign reversal was observed in the apparent dependence of water levels in Lake Victoria around 1920 [13]. Such a phase reversal, if true, is difficult to understand on physical grounds and makes the search for the mechanism of the solar cycle response more elusive. One possibility could be that our Sun is at the borderline between overcompensation and under-compensation of the dimming effect of the sunspots by the brightening effect of the faculae. However, modern reconstructions of the TSI (e.g., [83, 85, 84]) do not show this reversal between TSI maximum and sunspot number maximum. Many studies have attempted to correlate temperature measurements and solar activity during the past 100 years, and the Sun-climate relationships seem dependent upon the chosen geographical region, the time interval, and the analysis technique.

In this thesis a persistent and unambiguous solar cycle response will be established using an objective technique called *composite-mean difference projection* (CMD-P) [11]. We will show that the solar max always warms Earth's surface in the global mean. This result is also relevant with respect to the recent finding (unconfirmed) that because the ultra-violet part of the solar radiation change in the recent solar cycle is so unexpectedly large that the part of the radiation reaching the surface may be opposite in sense so that the surface may cool during solar max.

Furthermore, there is also a secular trend in the solar radiation, which is rather controversial (see, e.g., [125]). This is because it has often been used by global warming skeptics as a cause for the observed global warming. The relative significance of solar variability and other forcings of climate change (e.g., the greenhouse gases which have increased greatly during the industrial era due to human activities) is still under debate. Scafetta and West [124] suggested that the Sun might have contributed up to approximately 50% to the observed global warming during the industrial era by studying the lower tropospheric temperature. On the other hand, Hansen et al. [54] suggested, by running a global climate model, that the secular changes in the solar radiation are too small to explain the substantial increase in the Earth's temperature in recent decades. It still remains an open question as to the role of solar variations (both cyclic and secular) in climate change. In order to establish solid

evidence on the effects of the 11-year solar cycle and longer term solar variability on the climate and better understand the associated mechanisms, it is necessary and important to decipher the observations. The observed response to the solar cycle variations can also be used to calibrate models on their ability of climate prediction.

Isolating the solar cycle signal from a mixture of other signals and noise is a challenging undertaking due to the short length of the currently available observed data record. Many researchers have explored this issue using various techniques. However, even the reality of the solar response has remained controversial until very recently. In 2002, Labitzke et al. [76] asserted the existence of a global solar cycle signal in the stratosphere by calculating the correlations based on the NCEP/NCAR re-analyses [69] (Figure 1.1), but their argument is unconvincing. The correlations are very small in most areas and no connection between the temperature and the solar forcing has been shown statistically. Such small correlations can be achieved even if the temperature data contain no solar cycle signal. And it would also be difficult for them to pass statistical tests. Shortly afterwards Gleisner and Thejll [39] tried to filter out the unrelated variability to obtain higher correlations so that they are statistically significant at the 95% level in most areas. However, Coughlin and Tung [20] have pointed out that their results were problematic since removing variability that is uncorrelated with the remainder will always increase the correlation even if the original time series does not involve the signal of interest. Some other researchers have applied the classical technique of multiple linear regression (MLR) to disentangle responses to multiple phenomena including the solar response. For example, Haigh [45] deduced solar cycle signals in zonal mean temperature from stratosphere to surface (Figure 1.2); Roy and Haigh [120] studied the solar signals in sea level pressure (SLP) and sea surface temperature (SST); Lean and Rind [86] compared the influences of the four major climate drivers (solar variation, El Niño-Southern Oscillation (ENSO), anthropogenic warming, and volcano cooling) on observed surface temperatures. One major drawback of their work is that their geographical response patterns were not examined statistically or statistical significance was established in only a very few small regions based on the classical Student's $t$-test. Therefore the question of whether the claimed "solar" signal is really solar-related or whether there exists a coherent global temperature signal for the 11-year solar cycle was not clearly answered.

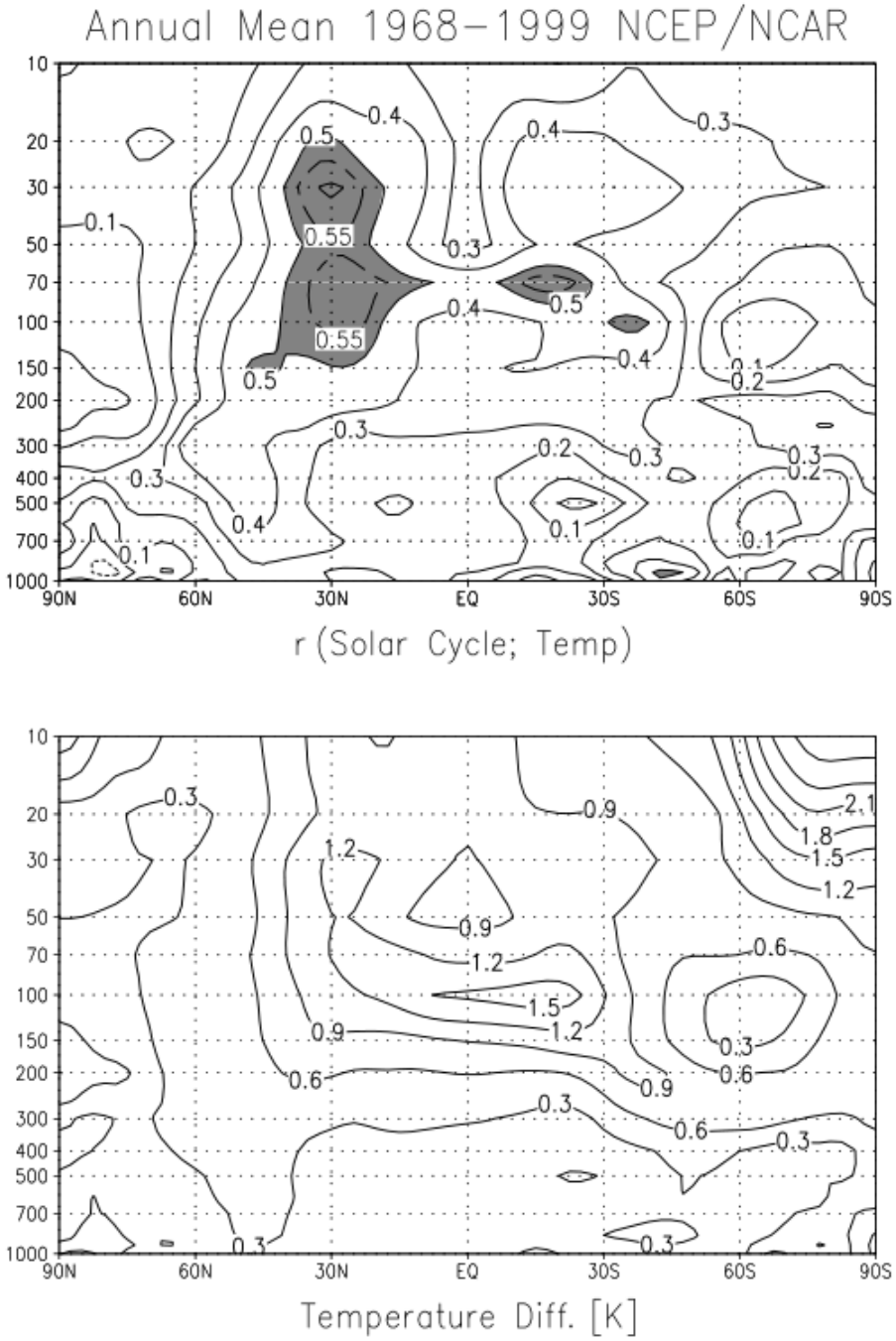


Figure 1.1: Copy of Figure 3(b) of Labitzke et al. [76]. Meridional cross-sections for the period 1968-1999 of (top): the correlations between the zonally averaged annual mean temperatures and the 11-year solar cycle (shaded for emphasis where the correlations are above 0.5); (bottom): the zonally averaged temperature differences (K) of solar maxima minus solar minima for the annual means.

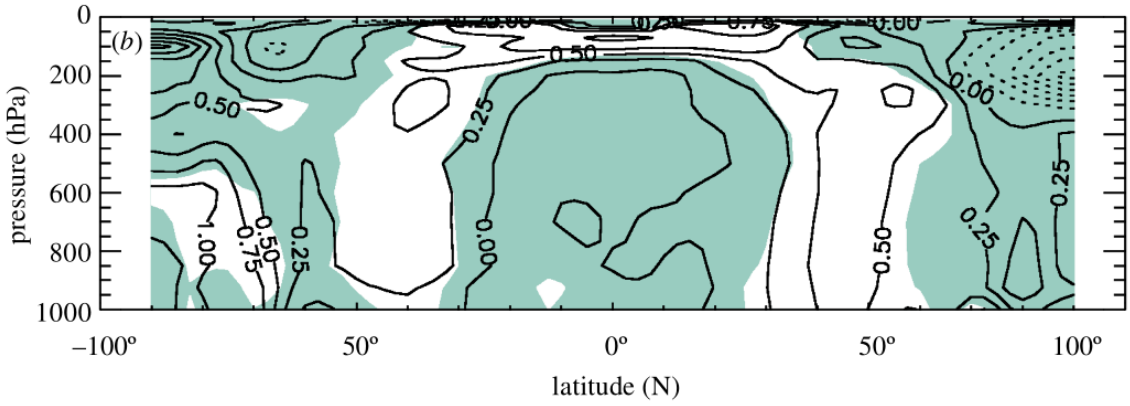


Figure 1.2: Copy of Figure 3(b) of Haigh [45]. Amplitudes of the components of variability in zonal mean temperature due to solar forcing. The unit is maximum variation (K) over the data period. Shaded areas are not statistically significant at the 95% level using a Student's $t$-test.

Recently Camp and Tung [11] made a major breakthrough in finding evidence, by applying the CMD-P method to the surface temperatures over the past 50 years, which covers four-and-a-half solar cycles, that the variable solar radiation has a planet-wide influence on the Earth's atmosphere and the global average temperatures oscillated by 0.12-0.18°C between high and low points in the 11-year sunspot cycle. The CMD technique has already been widely used in the fields of climate research ([76, 148, 60, 78], see also Figure 1.1), but the additional *projection* procedure and the following statistical test are new. Chapter 2 of this thesis describes an advanced version of this method. An alternative technique that functions similarly is also introduced.

In Chapter 3, the CMD-P method will be employed to investigate the solar cycle response in global sea surface temperature (SST) over the last 150 years. The discovery resulting from analysing the longest instrumental record of surface temperature with global ocean coverage adds to our confidence in a global temperature signal attributable to the 11-year solar cycle. It is also confirmed that the earth was warmer during the solar max as compared to the solar min, dispelling previous claims that the solar cycle response before 1920 is opposite to that of the modern era. The magnitude of the solar cycle response averaged over the oceans between 60°S and 60°N is about 0.1°C of warming for each $\mathrm{Wm}^{-2}$ variation of the solar constant (but

is slightly lower, at ∼0.085°C, when periods of suspected bad data are averaged in, which is consistent with previous work). In the second part of Chapter 3, we derive a constraint for the transient climate response (TCR) using observational results on the response to the 11-year solar variation. The derivation is completely based on statistically robust solar cycle signals, which are extracted from multiple surface temperature datasets, including reanalysis and blended *in situ* land-ocean data. The coupled atmosphere-ocean models participating in the 4th Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC) span a large range in their TCR. It is seen, however, that most models assessed by IPCC AR4 have too low a TCR compared with our derived constraint.

The existence of the 11-year solar cycle signal in temperature will also be established statistically throughout the troposphere in Chapter 4. A reasonable mechanism for producing the temperature response in the lower atmosphere of the Earth to the solar cycle forcing will be revealed afterwards by examining the latitude-height patterns of the response. Previously Haigh [45] obtained a "horse-shoe" pattern of warming using the multiple linear regression analysis: a heating center at the tropical tropopause and two vertical strips of warming outside the edge of the tropics in the troposphere. The author brought forward a "top-down" mechanism for this shape of the response. In her view, that pattern in temperature is a result of weakening and poleward shifts of the Hadley circulation, driven by enhanced ozone heating in the stratosphere in the solar max relative to the solar min [44, 47]. However, alternative mechanisms are available as well. As suggested by models, deep convection due to surface radiative heating and evaporative feedback can produce similar amplitude and spatial pattern of a temperature response in the troposphere [95, 94, 50, 59]. Unlike the "top-down" mechanism, the enhanced UV change does not play a role in this physical process which is termed the "bottom-up" mechanism [8]. It will be demonstrated that our observational results favor the latter mechanism.

The focus of Chapter 5 will be given to the Pacific's response to surface heating in the two longest historical instrumental sea surface temperature (SST) datasets [the Extended Reconstructed SST dataset (ERSST) and the Met Office Hadley Centre Sea Ice and SST dataset (HadISST)]. Using a modified method of multiple linear regression, it is found that the response to increased greenhouse forcing is a warm SST in the mid- to eastern Pacific

Ocean in the equatorial region in the annual or seasonal mean. The warming is statistically significant at the 95% confidence level. Consistent with this, the smaller radiative heating from solar forcing produces a weak warming also in this region, and the spatial pattern of the response is neither La Niña-like nor El Niño-like. It is noted that previous reports of a cold-tongue (La Niña-like) response to increased greenhouse or to solar-cycle heating [148, 147, 99, 100] were likely caused by contamination due to the dominant mode of natural response in the equatorial Pacific. Our result has implications on whether the Walker circulation is weakened or strengthened in a warmer climate and on coupled atmosphere-ocean climate model validation.

So far, our work has been concentrated on the solar response. In the final chapter of the thesis, we will make an effort to address a question that many may have been concerned with: is the current global warming speeding up? The answer seemed to be positive in IPCC AR4 [134], where various linear trends in global mean temperature were examined: the recent 25-year trend appears to be larger than the 50-year trend, which is in turn larger than the 100-year and 150-year trends (Figure 1.3). However, apparent fluctuations can also be seen throughout the entire period. The globe experienced cooling in 1970s, for example, and the period of accelerating warming seen after that time seems to have stalled since 2005. Many explanations have been given for these changes in global temperature [132, 133, 72], including increases in stratospheric water vapor and background aerosol. Even the Sun was thought to blame for historical episodes of warming by some researchers [1, 123]. We will demonstrate, by comparing the global-mean surface temperature available after 1850 with the 352-year-long Central England Temperature (CET) record, that the current anthropogenic global warming rate is almost constant ($\sim 0.08^{\circ}$C per decade) over the past 75 years. An oceanic internal oscillation with an average period of 70 years is verified to be real with 5 cycles, which accentuates the severe cold in the Little Ice Age (LIA) and the cooling in 1970s, explains a few episodes of warming in the preindustrial era and similarly the controversial Early Twentieth Century Warming, and adds to theories for the recent slowing of global warming. The solar forcing is found not to be responsible for rapid warming either at the end of Maunder Minimum or in the industrial era.

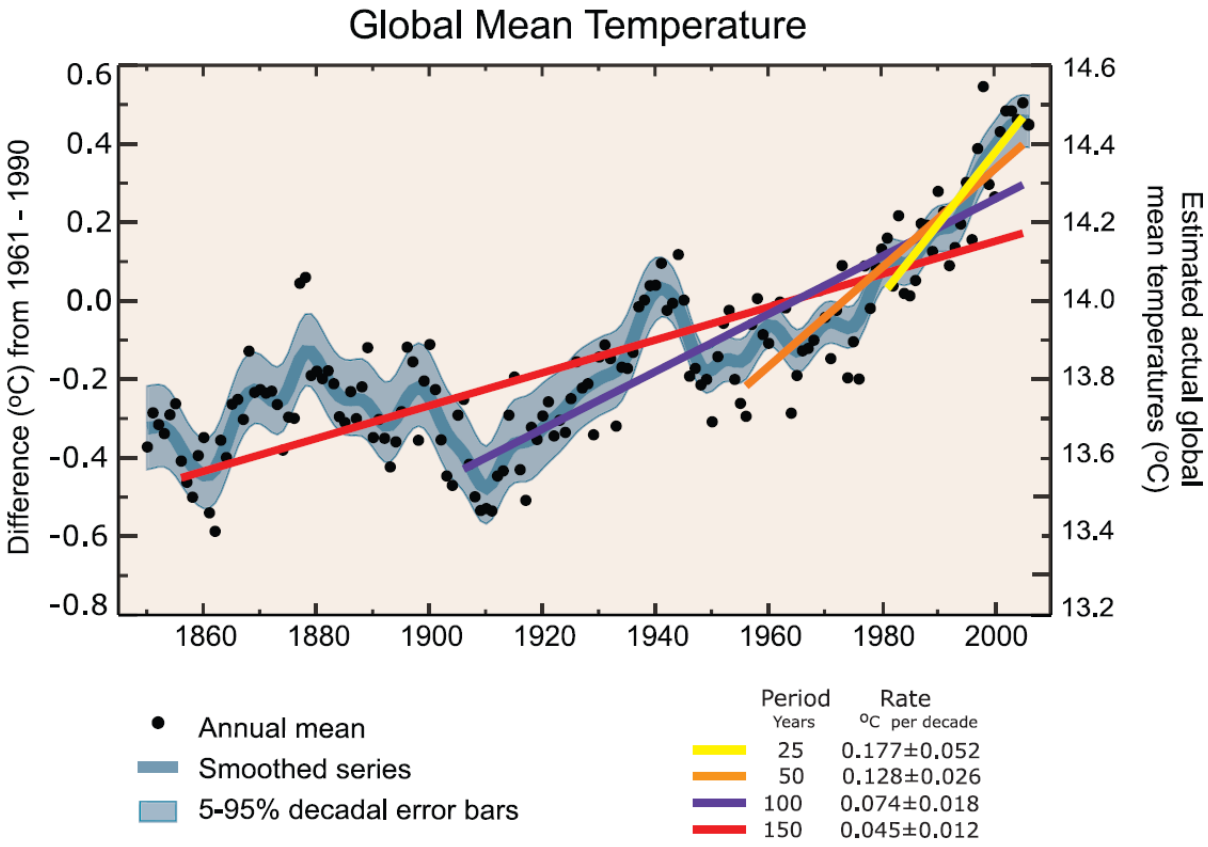


Figure 1.3: Annual global mean observed temperatures (black dots) along with simple fits to the data. The left hand axis shows anomalies relative to the 1961 to 1990 average and the right hand axis shows the estimated actual temperature (°C). Linear trend fits to the last 25 (yellow), 50 (orange), 100 (purple) and 150 years (red) are shown, and correspond to 1981 to 2005, 1956 to 2005, 1906 to 2005, and 1856 to 2005, respectively. Figure appeared in IPCC AR4 [134].

# Chapter 2

## THE CMD PROJECTION METHOD

Composite Mean Difference (CMD) is a simple and common technique for finding the spatial pattern responsible for the difference between two states. The observation, a spatial-time variable, is first divided into two groups in time, for example, solar maximum years and solar minimum years, and then the spatial pattern of the corresponding climate response is constructed by differencing the averages of the two composites. Spatial CMD patterns have been derived for the solar-cycle response in atmospheric temperature by Labitzke et al. [76], in the Pacific Ocean by van Loon et al. [148], and for the Quasi-Biennial Oscillation (QBO) phenomenon in the stratosphere by Holton and Tan [60] and Labitzke and van Loon [78], among others. Later Camp and Tung [11] extended the CMD method to verify the statistical significance of the spatial pattern generated by this technique. The extended CMD procedure, called the *CMD Projection method* in their article (hereinafter referred to as "CMD-P"), is essentially an empirical orthogonal mode expansion of the original data using the spatial CMD pattern as a basis. It is similar to the Empirical Orthogonal Function (EOF) expansion commonly used in climate data analysis except that its first mode is empirically defined to be phenomenon-specific instead of being the mode that explains the most variance in the data as in the latter. The EOF, being a mathematically defined statistical quantity, may contain contributions from several phenomena, making its physical interpretation difficult. Furthermore, because the modes in the EOF expansion are ordered according to the magnitudes of their variances, the first few modes are often made up of internal variability, with smaller forced responses relegated to the higher modes, which are seldom studied. In CMD, a known (forcing) index of a quasi-periodic phenomenon under consideration is used to define two groups (e.g., high and low index years) by composite mean difference of the data segregated in time into these two groups, yielding a spatial pattern of the "response" to this forcing. An associated time series which represents the

temporal behaviour of this "response" is then obtained by projecting the original spatial-time data onto the spatial CMD pattern. A statistical test on the first mode of the expansion, conducted using bootstrap Monte Carlo simulations, has been devised to test whether the association of the "response" with the "forcing" index is distinguishable from that obtained by random chance. This Monte Carlo test, based on the CMD-P method, notwithstanding its simplicity, turns out to be surprisingly effective in detecting a small quasi-periodic signal embedded in relatively large noise and other irrelevant atmospheric variability. As an overall test, it makes the target signal (if it exists) stand out by taking advantage of the spatial CMD pattern to filter out the noise and thus increases the probability of rejecting the null hypothesis of the absence of the signal.

In this chapter the CMD-P procedure and its associated botstrap Monte Carlo test are briefly reviewed. A Student's $t$-test is also proposed to establish the statistical significance of the local spatial characteristics of the climate signal being sought. At the end of this chapter Linear Discriminant Analysis (LDA) is described as an alternative method to find features that distinguish the two states of a quasi-periodic oscillation.

### *2.1 CMD Projection and Pairwise CMD with Shift*

The algorithm for the classical CMD method is straightforward. Take as an example the extraction of the 11-year solar cycle signal from temperature record. Given an observed data sequence $\{D_i(x)\}_{i=1}^{n}$ where $x$ denotes the space variable and $n$ is the data length in time. Note that the observations need not necessarily be continuous in time. Suppose $n_1$ ($n_2$) years of observations are classified as solar-max (solar-min) years ($n_1 + n_2 = n$). This classification can be done according to proper objective information about the solar variation, such as the total solar irradiance (TSI) time series [87, 153, 74]. The spatial pattern of the solar response is then calculated as follows:

$$P_1(x) = \sum_{i=1}^{n} w_i T_i(x), \tag{2.1}$$

where

$$w_i = \begin{cases} \dfrac{1}{n_1}, \text{ if the } i^{\text{th}} \text{ year belongs to the solar-max group} \\ -\dfrac{1}{n_2}, \text{ if the } i^{\text{th}} \text{ year belongs to the solar-min group} \end{cases}. \tag{2.2}$$

The accuracy of this approach relies on the degree to which the contaminations caused by the stochastic noise and other deterministic signals such as El Niño-Southern Oscillation (ENSO) can be eliminated by the CMD procedure. It is usually expected that the CMD value of a random sequence is small as long as the data record is long enough, but this property does not apply to deterministic noise. The residual from the deterministic components can be roughly evaluated by composite differencing their proxy indices, for instance, the Cold Tongue Index (CTI) for ENSO, which is the averaged SST over 6°N-6°S,180°-90°W minus the global mean SST (data and references offered online by the University of Washington at `http://jisao.washington.edu/data/cti/`). Luckily this part is also small for most of the proxy data, since the variations of different climate forcings are usually not in phase. An important exception is the anthropogenic forcing. A monotonic positive trend can be seen in the air and sea surface temperature during the most recent five to six decades, which may be due to forcing agents other than the solar input. The increase in greenhouse gases could be a candidate since it agrees with the concurrent change in temperature. Therefore it is necessary to remove a trend from the temperature data before applying the CMD analysis, as was done by Camp and Tung [11]. The data that they analyzed (1959-2004) has an apparent linear trend; thus the detrending process is straightforward. However, in some longer-term record such as the 150-year SST data analyzed by Zhou and Tung [169], the presence of a nonuniform trend adds complexity to the task of detrending. In such circumstances, taking out trends piecewisely may introduce undesired artificial jumps into the data. However, performing the CMD method piecewisely can sometimes deal with a complicated trend. This idea leads to an improved CMD procedure called *pairwise differencing with shift*, which was developed and employed for the first time by Zhou and Tung [169].

The modified method involves implementing the CMD procedure on the data (temperature) twice on every subperiod consisting of one and a half solar cycles: The first CMD pattern is obtained by differencing solar max and solar min years of the first complete cycle. Then by shifting half a period, another CMD pattern (solar max minus solar min) is obtained. The half cycle (either solar max or solar min) in the middle of the 1.5-cycle subset is used twice. If there is a decadal linear trend present it would be canceled out

when we average the two CMD patterns obtained in this 1.5-cycle subset, without affecting the long-term trend from the original data. Finally, the CMD pattern for the oscillatory part of the signal for the overall period is the average of all the subpatterns derived from every subset of one and a half cycles. Let $m$ be the number of half solar cycles (solar maxes or solar mins), and $n_k$ ($k = 1, \cdots, m$) be the number of data points in the $k^{\text{th}}$ half-cycle solar-max or solar-min subgroup. A total of $m-2$ subsets can be exploited within the entire data sequence. The spatial pattern obtained by the method of pairwise CMD with shift can still be formulated as (2.1), but with the coefficients $w_i$ updated as follows:

$$w_i = \begin{cases} \dfrac{(-1)^s}{2(m-2)n_k}, & \text{if } \displaystyle\sum_{j=1}^{k-1} n_j < i \le \sum_{j=1}^{k} n_j \quad \text{for } k = 1, m \\ \dfrac{(-1)^s}{n_k}, & \text{if } \displaystyle\sum_{j=1}^{k-1} n_j < i \le \sum_{j=1}^{k} n_j \quad \text{for } k = 2 \text{ and } m = 3 \\ \dfrac{3(-1)^s}{2(m-2)n_k}, & \text{if } \displaystyle\sum_{j=1}^{k-1} n_j < i \le \sum_{j=1}^{k} n_j \quad \text{for } k = 2, m-1 \text{ and } m \ge 4 \\ \dfrac{2(-1)^s}{(m-2)n_k}, & \text{if } \displaystyle\sum_{j=1}^{k-1} n_j < i \le \sum_{j=1}^{k} n_j \quad \text{for } k = 3, \cdots, m-2 \text{ and } m \ge 5 \end{cases}, \tag{2.3}$$

where $s = 0$ ($s = 1$) when the $i^{\text{th}}$ year belongs to the solar-max (solar-min) group.

In certain cases one may need to divide the whole data record into a few smaller segments and apply the method of pairwise CMD with shift separately to each of the segments because of the possible discontinuities residing in data to be analyzed. This was done in [169] because the years of 1942-1950 were removed from their analysis of the 150-year SST data for bad data quality [139]. As a result the two half solar cycles right before and after the deleted years are not naturally connected. Some other factors such as the inclusion of advanced measurements from satellites during the past few decades can also introduce troublesome discontinuities into data. Thus deriving CMD patterns for two or more segments of the data record is necessary to avoid introducing large errors caused by the artificial gap. The final spatial pattern associated with the entire period can then be simply defined as the average of all the segment patterns.

The CMD-P method takes an additional step to project the original data onto the spatial CMD pattern $P_1(x)$, resulting in a time series of the projection coefficients $C_1(t)$. The

temperature data $D(t,x)$ can be decomposed in an empirical orthogonal mode expansion as

$$D(t,x) = \sum C_i(t)P_i(x), \tag{2.4}$$

where

$$C_i(t) = \frac{\int D(t,x)P_i(x)dx}{\int P_i^2(x)dx}. \tag{2.5}$$

The integration is over the space of $x$. For a spherical earth, the integration over the surface is given by $a\cos\theta d\theta d\varphi$, where $\theta$ is latitude and $\varphi$ is longitude ($a$ is the radius of the earth). What is of interest is the first mode. All the higher modes combined together, called the remainder, need not be decomposed. The way in which $C_1(t)$ is constructed guarantees that the "solar cycle" mode $P_1(x)$ is orthogonal to the remainder.

### *2.2 Bootstrap Monte Carlo Test*

The correlation coefficient $\rho$ between $C_1(t)$ and the TSI is then calculated and used as the test statistic in what we shall call the $\rho$-test, for the purpose of answering the question of whether the "solar cycle signal" extracted from the temperature data by the CMD method is just a chance outcome: the null hypothesis. We address this question by looking into the likelihood that the statistic $\rho$ could randomly achieve a value as large as the observed correlation in magnitude. Consider the original temperature data as a sample from a statistical population. It is impossible to infer or mimic the theoretical sampling distribution of $\rho$ since almost no attribute of the population is known and no other sample beyond the obeserved temperature is obtainable. In such a case *bootstrapping* is a favored practice of estimating the unknown population distribution and therefore approaching the sampling distribution of a given statistic based solely on one sample.

In bootstrap Monte Carlo simulations, numerous resampled datasets (10,000 surrogates are usually adopted in our studies) are constructed by random resampling with replacement from the original dataset. Each resampled dataset must have $n$ years of data, the same as the original dataset, and its construction is completed in $n$ steps. At each step one randomly drawn observation from the real data record is assigned to a year of the resampled dataset which is empty prior to this assignment and then returned to the pool for later draws. (The year chosen in a previous step must be returned to the pool before another year is

chosen; otherwise later draws would not be independent of the early ones. For example, if the years were not returned to the pool and $k-1$ yr were chosen, the $k^{\text{th}}$ year would be dependent on the previous $k-1$ yr.) The original association of the temperature with the solar variation is destroyed in the synthetic dataset created by this means, even if it is present in the real data. We produce many such resampled datasets in a similar fashion. In the meanwhile the relationship between the TSI and years is not randomized but held fixed as the real values, so the grouping of years into the solar groups remains the same as when we analyse the original data. The CMD-P method is then repeatedly applied to each of these resampled datasets to yield a large number of values for the statistic $\rho$. Thus the bootstrap distribution of $\rho$ is formed for the original sample and can be used to carry out a statistical test for the observed $\rho$. The probability of rejecting the null hypothesis (i.e., no solar cycle signal exists) can then be established by seeing where the observed value of the correlation coefficient sits in the bootstrap distribution of $\rho$. For example, if more than 95% of the randomly resampled datasets yield a value of $\rho$ smaller than the observed value in magnitude, then the null hypothesis can be rejected at 5% significance level.

A useful technique for the bootstrap resampling to account for the inherent autocorrelation of the climate data is to draw blocks of $L$ consecutive observations with replacement instead of individual observations. This modified bootstrap method is called the *moving-block bootstrap*, which captures the structure of dependence of neighboring observations [29, 79, 89, 157]. The value of $L$ should be chosen so that data units of $L$ distance apart from each other are essentially independent. Generally the larger the $L$, the smaller the risk of rejecting a null hypothesis when it does hold in the population. Under the assumption that the original time series is modeled as a first-order autoregressive process (AR(1)), Wilks [158, chapter5] has suggested that a good choice of the block length $L$ is given by

$$L = (n - L + 1)^{2/3(1-n'/n)}, \tag{2.6}$$

where $n$ is the sample size, $n' = n(1-\rho_1)/(1+\rho_1)$ is the approximated effective sample size, and $\rho_1$ is the lag-1 autocorrelation coefficient. However, the estimated value for $L$ given by (2.6) may not be a suitable choice since an AR(1) process is usually not a good fit for climate data. It is in general difficult to theoretically determine an appropriate block

length [89]. For certain practical problems we can just adopt a very simple way: repeat the computation for a variety of $L$ values (e.g., $L = 1, 2, 3, \cdots, 10, 11, 12$, etc.) and report the most conservative result obtained. By this means the test result should be underestimated, but the same conclusion can be drawn if it is still statistically significant at some preselected probability threshold (e.g., 1% or 5%).

Although Camp and Tung [11] have illustrated the usefulness of the bootstrap $\rho$-test associated with CMD-P, no analytical theory on the reliability and validity of this novel statistical test is available due to the complexity of both the *test statistic* (the correlation coefficient $\rho$) and the composition of the climate data. As a substitution, Zhou and Tung [168] recently empirically confirmed the trustworthiness of this test, even in the presence of other deterministic variability in the "noise", by conducting a large number of numerical simulations. In other words, the probability that the authors have reported a solar cycle signal that does not actually exist is very low if the test result turns out to be at 95% confidence level.

### *2.3 Student's $t$-Test*

The bootstrap Monte-Carlo $\rho$-test mentioned above technically justifies the existence of the target signal at a given atmospheric data record. In addition to this, we are also interested in knowing whether the CMD-derived spatial characteristics of the solar cycle response are distinguishable from random noise at various spatial locations. Many authors have applied the conventional two-sample Student's $t$-test (also known as Welch's $t$-test if the variances of the two populations are assumed to be unequal) to answer this question (see, e.g., [148]). However, this classical student's $t$-test merely tests the validity of the null hypothesis that the means of two populations from which the data units of the two composites are sampled are equal. For this reason there is a possibility that the $t$-test passes due simply to the effects of some other deterministic phenomena, even in the case of the signal of interest being absent. In addition, the presence of serial correlations in climate data [170, 171] violates the fundamental assumption of the two-sample $t$-test. Despite the complications mentioned above, we can still conduct Student's $t$-tests on our CMD results with autocorrelations being taken into account by assuming a regression-like model for the

time series to be analysed. The model takes the form of $y(t) = \beta \cdot \mathrm{TSI}(t) + \epsilon(t)$, where $y(t)$ represents the temperature data and the error term $\epsilon(t)$ is assumed to be a random noise sequence. The possible solar cycle signal in the temperatures is assumed to be proportional to the TSI index time series. More generally, the model can be extended to account for more than one deterministic signal, as does the multiple linear regression (MLR):

$$y(t) = \sum_{j=1}^{k} x_j(t)\beta_j + \epsilon(t), \tag{2.7}$$

where $k$ is the number of deterministic signals under study and each $\beta_j$ represents the (unknown) true response to corresponding climate influence whose time behavior is described by the observable index $x_j(t)$. The predictor variables $x_j$ are simply assumed to be error-free and uncorrelated with the error term $\epsilon$. In the case of univariate CMD for deriving the solar response, model (2.7) carries only one forcing index $x_1(t)$ which is the TSI.

There are a number of ways to estimate the unknown parameters $\beta_j$ in linear regression. The simplest and most commonly used technique is the Ordinary Least Squares (OLS), which minimizes the *residual sum of squares*. Given $n$ observations $\{y_i, x_{i1}, \cdots, x_{ik}\}_{i=1}^{n}$, the OLS estimator of $\boldsymbol{\beta} = [\beta_1, \cdots, \beta_k]^{\mathrm{T}}$, denoted by $\hat{\boldsymbol{\beta}} = [\hat{\beta}_1, \cdots, \hat{\beta}_k]^{\mathrm{T}}$, is given in vector form as follows:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y}, \quad \text{where} \quad \mathbf{Y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \text{and} \quad \mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nk} \end{bmatrix}. \tag{2.8}$$

Note that the design matrix $\mathbf{X}$ is required to have full column rank $k$ to identify the unknown parameter vector $\boldsymbol{\beta}$. This condition is generally satisfied in climate research. The **Gauss-Markov theorem** states that the OLS estimator in (2.8) is the best linear unbiased estimator (*BLUE*) of $\boldsymbol{\beta}$ under the assumption that the errors have expectation zero and are mutually uncorrelated and have equal finite variances (this notion is called *homoscedasticity*).

The CMD method yields an alternative estimator of $\boldsymbol{\beta}$ by assuming that the composite mean difference of the noise itself is small and thus ignored. One minor difference between the MLR and CMD models is that the former usually contains an *intercept* whose

corresponding explanatory variable is identical across all the observations while the latter does not need such a constant term because it is always eliminated by the composite mean difference. When multiple phenomena are involved in the model to allow a correction for possible cross-contamination, $\boldsymbol{\beta}$ can be estimated by applying the following procedure based on CMD. Given $n$ observations as in MLR, we first perform the CMD procedure $k$ times on the data, each with the two groups selected according to different forcing agent (thus the observations contributing to calculate the CMD may vary for each climate signal), to establish the following linear equations:

$$\mathbf{W}^{\mathrm{T}}\mathbf{Y} = \mathbf{W}^{\mathrm{T}}\mathbf{X}\boldsymbol{\beta} + \mathbf{W}^{\mathrm{T}}\boldsymbol{\epsilon}, \tag{2.9}$$

where $\boldsymbol{\beta}$, $\mathbf{Y}$ and $\mathbf{X}$ are the same as those in MLR. The unknown sample values of the random noise embedded in the $n$ observations are written in the vector $\boldsymbol{\epsilon} = [\epsilon_1, \cdots, \epsilon_n]^{\mathrm{T}}$. The $n$-by-$k$ matrix $\mathbf{W}$ stores in columns the weights for implementing the $k$ CMD procedures. For example, the weights for performing the pairwise CMD with shift according to the solar forcing is given in (2.3). By ignoring the noise CMD, $\mathbf{W}^{\mathrm{T}}\boldsymbol{\epsilon}$, which is assumed to be small, the CMD estimator of $\boldsymbol{\beta}$ is obtained as:

$$\hat{\boldsymbol{\beta}} = (\mathbf{W}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{W}^{\mathrm{T}}\mathbf{Y}. \tag{2.10}$$

Here $\mathbf{X}$ again must have full column rank to invert $\mathbf{W}^{\mathrm{T}}\mathbf{X}$. Besides, since the variations of different climate forcings are usually not in phase, the matrix $\mathbf{W}^{\mathrm{T}}\mathbf{X}$ are expected to have large main diagonal elements because of the selection of the two CMD groups while its non-diagonal elements are all expected to be small, as they turn out to be in our practical problems. Therefore the nonsigularity of $\mathbf{W}^{\mathrm{T}}\mathbf{X}$ is generally guaranteed and consequently $\boldsymbol{\beta}$ is identifiable by this means. The linear estimator in (2.10) is also unbiased if the errors have expectation zero. However, its variance is larger than the variance of the OLS estimator if the conditions provided in the Gauss-Markov theorem all hold true. The extension of the univariate CMD described above was first proposed by Zhou and Tung [169] and called the method of *multiple CMD* (MCMD).

In climate data the errors are usually autocorrelated. Therefore the autocorrelations must be taken into account when we perform a statistical analysis of $\boldsymbol{\beta}$. Next the Student's

$t$-test on CMD results will be introduced by assuming that the errors in (2.7) follow an autoregressive Gaussian process of order $p$ (AR($p$)):

$$\epsilon(t) = \sum_{i=1}^{p} \phi_i \epsilon(t-i) + \tilde{\omega}(t), \tag{2.11}$$

where $\{\tilde{\omega}(t)\}$ is identically distributed Gaussian white noise and $\phi_1, \cdots, \phi_p$ are the autoregressive coefficients. Statistical theory asserts that an AR($p$) red noise time series of length $n$ can be rewritten in the form of $\boldsymbol{\epsilon} = \boldsymbol{\Phi}\boldsymbol{\omega}$, where the elements of $\boldsymbol{\omega}$ are independent and identically distributed (i.i.d.) Gaussian random variables and $\boldsymbol{\Phi}$ is the inverse of the $n$-by-$n$ lower triangular matrix

$$\boldsymbol{\Phi}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{v_0}} & 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ -\frac{\phi_{1,1}}{\sqrt{v_1}} & \frac{1}{\sqrt{v_1}} & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ -\frac{\phi_{2,2}}{\sqrt{v_2}} & -\frac{\phi_{2,1}}{\sqrt{v_2}} & \frac{1}{\sqrt{v_2}} & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ -\frac{\phi_{p-1,p-1}}{\sqrt{v_{p-1}}} & -\frac{\phi_{p-1,p-2}}{\sqrt{v_{p-1}}} & \cdots & -\frac{\phi_{p-1,1}}{\sqrt{v_{p-1}}} & \frac{1}{\sqrt{v_{p-1}}} & 0 & \cdots & \cdots & 0 & 0 \\ -\phi_p & -\phi_{p-1} & \cdots & \cdots & -\phi_1 & 1 & \cdots & \cdots & 0 & 0 \\ 0 & -\phi_p & -\phi_{p-1} & \cdots & \cdots & -\phi_1 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & -\phi_p & \cdots & \cdots & \cdots & -\phi_1 & 1 \end{bmatrix}. \tag{2.12}$$

The reverse Levinson-Durbin recursion can be used to compute the elements of $\boldsymbol{\Phi}^{-1}$. Let $\phi_{p,i} = \phi_i$ for $i = 1, \cdots, p$. Given $\phi_{h,1}, \cdots, \phi_{h,h}$ and $v_h$ (with $v_p \equiv 1$), the procedure then recursively yields

$$\phi_{h-1,i} = \frac{\phi_{h,i} + \phi_{h,h}\phi_{h,h-i}}{1 - \phi_{h,h}^2} \quad \text{for} \quad 1 \le i \le h-1, \tag{2.13}$$

and

$$v_{h-1} = \frac{v_h}{1 - \phi_{h,h}^2}. \tag{2.14}$$

The model (2.7) then turns out in the matrix-vector form to be

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\Phi}\boldsymbol{\omega}. \tag{2.15}$$

The general form of a linear estimator of $\boldsymbol{\beta}$ is $\hat{\boldsymbol{\beta}} = \mathbf{MY}$, where the matrix $\mathbf{M} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$ for OLS and $\mathbf{M} = (\mathbf{W}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{W}^{\mathrm{T}}$ for MCMD. In the case of known autoregressive coefficients $\phi_1, \cdots, \phi_p$, the *BLUE* of $\boldsymbol{\beta}$ is given by the Generalized Least Squares (GLS) estimator:

$$\hat{\boldsymbol{\beta}} = \mathbf{MY}, \quad \text{where} \quad \mathbf{M} = \left(\mathbf{X}^{\mathrm{T}}\boldsymbol{\Phi}^{-\mathrm{T}}\boldsymbol{\Phi}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\Phi}^{-\mathrm{T}}\boldsymbol{\Phi}^{-1}. \tag{2.16}$$

This is simply the OLS estimator applied to the transformed data, $\boldsymbol{\Phi}^{-1}\mathbf{Y} = \boldsymbol{\Phi}^{-1}\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\omega}$, for which the assumptions of the Gauss-Markov theorem are met. It is usually unrealistic to know the true values of the autoregressive coefficients. So we will need to exploit appropriate approximations to handle a practical problem. Among many the Yule-Walker estimation is one most commonly used technique to approach the autoregressive coefficients. The order $p$ can be determined by minimizing the AICC (Akaike Information Criterion with a Correction) statistic [5, chapter 5]. Once an AR($p$) model is fitted to the residue $\hat{\boldsymbol{\epsilon}} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$ resulting in estimated autoregressive coefficients $\hat{\phi}_1, \cdots, \hat{\phi}_p$, statistical inference for the parameters in $\boldsymbol{\beta}$ can be approximately done by taking $\hat{\phi}_1, \cdots, \hat{\phi}_p$ as the known values of $\phi_1, \cdots, \phi_p$.

For an estimator designed subject to $\mathbf{MX} = \mathbf{I}$ (i.e., the identity matrix), we have

$$\hat{\boldsymbol{\beta}} = \mathbf{MY} = \boldsymbol{\beta} + \mathbf{M}\boldsymbol{\Phi}\boldsymbol{\omega}. \tag{2.17}$$

Since $\omega_1, \cdots, \omega_n$ are i.i.d. normal random variables with expectation zero and (unknown) variance $\sigma^2$, the random vector $\hat{\boldsymbol{\beta}}$ has a multivariate normal distribution:

$$\hat{\boldsymbol{\beta}} \sim \mathcal{N}\left(\boldsymbol{\beta}, \sigma^2\boldsymbol{\Sigma}\right), \quad \text{where the covariance matrix } \boldsymbol{\Sigma} = \mathbf{M}\boldsymbol{\Phi}(\mathbf{M}\boldsymbol{\Phi})^{\mathrm{T}}. \tag{2.18}$$

Furthermore, we can define another random vector

$$\boldsymbol{\eta} = \mathbf{V}^{\mathrm{T}}\mathbf{Y} = \mathbf{V}^{\mathrm{T}}\mathbf{X}\boldsymbol{\beta} + \mathbf{V}^{\mathrm{T}}\boldsymbol{\Phi}\boldsymbol{\omega}, \tag{2.19}$$

which satisfies the following conditions:

(i) All the elements of $\boldsymbol{\eta}$ are i.i.d. normal random variables with expectation zero and variance $\sigma^2$;

(ii) The random vectors $\hat{\boldsymbol{\beta}}$ and $\boldsymbol{\eta}$ are independent.

In order to find such an $\boldsymbol{\eta}$, we first find a matrix $\tilde{\mathbf{V}}$ whose columns form an orthonormal

basis for the null space of the matrix $[(\mathbf{M\Phi})^{\mathrm{T}},\ \mathbf{\Phi}^{-1}\mathbf{X}]^{\mathrm{T}}$, and then let $\mathbf{V} = \mathbf{\Phi}^{-\mathrm{T}}\tilde{\mathbf{V}}$. The column rank of $\mathbf{V}$ and $\tilde{\mathbf{V}}$ must be equal and full, which we denote by $r$ ($r$ is at most $n-k$ since the rank of $\mathbf{\Phi}^{-1}\mathbf{X}$ is $k$). Note that the matrix $\mathbf{V}$ defined in this way has the following properties: $\mathbf{V}^{\mathrm{T}}\mathbf{X} = \mathbf{0}$, $\mathbf{V}^{\mathrm{T}}\mathbf{\Phi}\mathbf{\Phi}^{\mathrm{T}}\mathbf{V} = \mathbf{I}$ and $\mathbf{M\Phi}\mathbf{\Phi}^{\mathrm{T}}\mathbf{V} = \mathbf{0}$. Therefore $\boldsymbol{\eta}$ satisfies the above conditions (i) and (ii) according to *Lemma 5.3.3* of [12]. In addition, *Theorem 4.6.12* of [12] also states that the random variable

$$S^2 \triangleq \frac{\boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{\eta}}{r} = \frac{\mathbf{Y}^{\mathrm{T}}\mathbf{V}\mathbf{V}^{\mathrm{T}}\mathbf{Y}}{r} \tag{2.20}$$

is independent of $\hat{\boldsymbol{\beta}}$. Actually $S^2$ is an unbiased estimator of $\sigma^2$ since the expected value of $S^2$ is

$$\mathbb{E}(S^2) = \frac{1}{r}\sum_{j=1}^{r}\mathbb{E}(\eta_j^2) = \frac{1}{r}\sum_{j=1}^{r}\Big(\mathrm{Var}(\eta_j) + (\mathbb{E}(\eta_j))^2\Big) = \sigma^2. \tag{2.21}$$

Consequently, it is straightforward to see the chi-square distribution with $r$ degrees of freedom

$$\frac{rS^2}{\sigma^2} \sim \chi_r^2 \tag{2.22}$$

and a family of student's $t$-distribution with $r$ degrees of freedom

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{S^2\Sigma_{jj}}} \sim t_r, \quad j = 1, \cdots, k, \tag{2.23}$$

where $\Sigma_{jj}$ is the $j^{\mathrm{th}}$ main diagonal element of $\mathbf{\Sigma}$.

The student's $t$ statistic in (2.23) can be used to form a student's $t$-test of the null hypothesis $\boldsymbol{H_0}$ : $\beta_j = 0$ versus the alternative hypothesis $\boldsymbol{H_1}$ : $\beta_j \neq 0$ (two-sided) or $\boldsymbol{H_1}$ : $\beta_j > 0$ (one-sided; $\boldsymbol{H_1}$ : $\beta_j < 0$ is tested in a similar way). We reject $\boldsymbol{H_0}$ at significance level $\alpha$ in favor of $\boldsymbol{H_1} : \beta_j \neq 0$ if

$$\left|\frac{\hat{\beta}_j - 0}{\sqrt{S^2\Sigma_{jj}}}\right| > t_{r,\frac{\alpha}{2}}; \tag{2.24}$$

or in favor of $\boldsymbol{H_1} : \beta_j > 0$ if

$$\frac{\hat{\beta}_j - 0}{\sqrt{S^2\Sigma_{jj}}} > t_{r,\alpha}. \tag{2.25}$$

A $100(1-\alpha)\%$ confidence interval for $\beta_j$ can also be constructed using expression (2.23):

$$\hat{\beta}_j - t_{r,\frac{\alpha}{2}}\sqrt{S^2\Sigma_{jj}} < \beta_j < \hat{\beta}_j + t_{r,\frac{\alpha}{2}}\sqrt{S^2\Sigma_{jj}}. \tag{2.26}$$

## 2.4 The LDA Method

### 2.4.1 Discriminant analysis

Linear discriminant analysis (LDA) is a longstanding statistical technique used for the classification of multivariate data into predefined groups [158, 119]. Since an oscillatory event can always be categorized into two groups (e.g., high and low index states), LDA can be used to find features which characterize or separate the two states. The LDA method was originally applied in climate data analysis by Schneider and Held [127] to deduce the temperature trends, and then Camp and Tung [9] and Tung and Camp [143] studied the Quasi-Biennial Oscillation (QBO) and solar cycle influences on various temperature data records by this means. The LDA procedure takes a more sophisticated form than does the CMD-P method. It in principle is an optimization process which is able to deal with multiple phenomena simultaneously and it was thought by Tung and Camp [143] of being more powerful for improving the purity of the extracted signal.

Given a set of bivariate observations (e.g., in time and space), $\mathbf{X}$, and some pre-defined groups, the LDA method seeks spatial weights $\boldsymbol{w}$ in a specific vector space that make the weighted linear combination of the original data, $\mathbf{X}\boldsymbol{w}$, have the greatest separation among pre-defined groups. The separation can be measured by the ratio of between-group variance to either within-group or total variance. Suppose that $\mathbf{X}(n \times p)$ is a centered data set, consisting of $n$ observations of $p$ variables, $\boldsymbol{x}_i(1 \times p)$. Centering is performed by removing the mean observations from the original data. Furthermore, suppose that we can partition the observations into $g$ groups, $G_j(j = 1, \cdots, g)$, with $n_j$ observations in each group. Let $[i]$ denote the group of observation $\boldsymbol{x}_i$. We can define the group matrix $\mathbf{G}(n \times g) = [g_{ij}]$, where $g_{ij} = 1$ if $j = [i]$ and equals 0 otherwise. Note that $\mathbf{G}^{\mathrm{T}}\mathbf{G} = \mathrm{diag}(n_j)$.

The between-group covariance matrix is defined as

$$\begin{aligned} \mathbf{\Sigma}_{\mathrm{b}} &= \frac{g}{n(g-1)}\mathbf{X}^{\mathrm{T}}\mathbf{G}(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{X} \\ &= \frac{g}{n(g-1)}(\mathbf{G}\mathbf{M})^{\mathrm{T}}(\mathbf{G}\mathbf{M}), \end{aligned} \tag{2.27}$$

where the $g$-by-$p$ matrix $\mathbf{M} = (\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{X}$ is the matrix of group means, whose $j^{\mathrm{th}}$ row denotes the mean of the observations in the $j^{\mathrm{th}}$ group. Incidently, the within-group

covariance matrix is defined as

$$\mathbf{\Sigma}_{\mathrm{w}} = \frac{1}{n-g}(\mathbf{X} - \mathbf{GM})^{\mathrm{T}}(\mathbf{X} - \mathbf{GM}). \tag{2.28}$$

The total covariance matrix $\mathbf{\Sigma}_{\mathrm{t}}$ is a linear combination of $\mathbf{\Sigma}_{\mathrm{b}}$ and $\mathbf{\Sigma}_{\mathrm{w}}$:

$$\begin{aligned} \mathbf{\Sigma}_{\mathrm{t}} &= \frac{\mathbf{X}^{\mathrm{T}}\mathbf{X}}{n-1} \\ &= \frac{1}{n-1}\left(\frac{n(g-1)}{g}\mathbf{\Sigma}_{\mathrm{b}} + (n-g)\mathbf{\Sigma}_{\mathrm{w}}\right). \end{aligned} \tag{2.29}$$

A separability measure for any linear combination $\mathbf{X}\boldsymbol{w}$ can then be defined by

$$\gamma = (\boldsymbol{w}^{\mathrm{T}}\mathbf{\Sigma}_{\mathrm{t}}\boldsymbol{w})^{-1}(\boldsymbol{w}^{\mathrm{T}}\mathbf{\Sigma}_{\mathrm{b}}\boldsymbol{w}). \tag{2.30}$$

The canonical variates, $\boldsymbol{c}_k = \mathbf{X}\boldsymbol{w}_k$ $(k = 1, \cdots, g-1)$, are the linear combinations which optimize $\gamma$. Therefore $\boldsymbol{w}_k$ must satisfy

$$\frac{\partial\gamma}{\partial\boldsymbol{w}} = 2(\boldsymbol{w}^{\mathrm{T}}\mathbf{\Sigma}_{\mathrm{t}}\boldsymbol{w})^{-1}\mathbf{\Sigma}_{\mathrm{b}}\boldsymbol{w} - 2(\boldsymbol{w}^{\mathrm{T}}\mathbf{\Sigma}_{\mathrm{t}}\boldsymbol{w})^{-2}(\boldsymbol{w}^{\mathrm{T}}\mathbf{\Sigma}_{\mathrm{b}}\boldsymbol{w})\mathbf{\Sigma}_{\mathrm{t}}\boldsymbol{w} = 0. \tag{2.31}$$

Simplifying this gives an equation in a neat form:

$$\mathbf{\Sigma}_{\mathrm{t}}^{-1}\mathbf{\Sigma}_{\mathrm{b}}\boldsymbol{w}_k = \gamma_k\boldsymbol{w}_k, \tag{2.32}$$

which implies that the $\gamma_k$ are the (ordered) eigenvalues of $\mathbf{\Sigma}_{\mathrm{t}}^{-1}\mathbf{\Sigma}_{\mathrm{b}}$ and the $\boldsymbol{w}_k$ are the associated right eigenvectors. This optimization is equivalent to maximizing the ratio

$$R = (\boldsymbol{w}^{\mathrm{T}}\mathbf{\Sigma}_{\mathrm{w}}\boldsymbol{w})^{-1}(\boldsymbol{w}^{\mathrm{T}}\mathbf{\Sigma}_{\mathrm{b}}\boldsymbol{w}). \tag{2.33}$$

The canonical variates are orthogonal time series indices. The first canonical variate $\boldsymbol{c}_1$ represents the largest separation of the observations into two groups recognized by low and high values of $\boldsymbol{c}_1$. The second index $\boldsymbol{c}_2$ captures the largest separability subject to the constraint that it is uncorrelated with the previous index $\boldsymbol{c}_1$. So does for each of the rest canonical variates. The discriminating patterns, $\boldsymbol{p}_k$ $(k = 1, \cdots, g-1)$, which distinguish the observations between groups as categorized according to the associated canonical variates, are defined as the regression coefficients of the centered data $\mathbf{X}$ onto $\boldsymbol{c}_k$, i.e.,

$$\boldsymbol{p}_k = (\boldsymbol{w}_k^{\mathrm{T}}\mathbf{\Sigma}_{\mathrm{t}}\boldsymbol{w}_k)^{-1}\mathbf{\Sigma}_{\mathrm{t}}\boldsymbol{w}_k. \tag{2.34}$$

Note that the discriminant patterns are not necessarily to be orthogonal to each other.

### *2.4.2 Regularization and the truncation number*

Since the number of variables $p$ in climate data exceeds the number of observations $n$, the system to be analyzed is overdetermined and the LDA based on temporal groups is ill-posed. Note that in such a case $\mathbf{\Sigma}_\mathrm{t}$ is a singular (noninvertible) matrix. Therefore, the weight vector $\boldsymbol{w}_k$ for maximizing the variance ratio $\gamma$ cannot be obtained before some regularization step is taken [55]. The LDA is then performed using a truncated principal component representation of the original dataset $\mathbf{X}$; retaining only a pre-selected number (say, $r$) of the leading modes. The weight vectors are computed according to (2.32) but using the Moore-Penrose pseudoinverse of the regularized toal covariance matrix (i.e., $\mathbf{\Sigma}_\mathrm{r} = \mathbf{X}_\mathrm{r}^\mathrm{T}\mathbf{X}_\mathrm{r}/(n-1)$, where $\mathbf{X}_\mathrm{r}$ is the regularized dataset) in place of the inverse $\mathbf{\Sigma}_\mathrm{t}^{-1}$. The regularization restricts the weight vectors $\boldsymbol{w}$ to the range of $\mathbf{\Sigma}_\mathrm{r}^{\dagger}$. The choice of the truncation number $r$ is crucial to the performance of the LDA method. If $r$ is too large, the model is overfitting the data. Since the number of observations in each group is small, choosing a large $r$ gives the model enough degrees of freedom to find a pattern nearly orthogonal to the within-group variability and so artificially high separation will be seen. If $r$ is chosen too small, desired variability may be missing in the truncated dataset which can only yield a poor and meaningless separation. This is particularly true for smaller amplitude signals such as the solar cycle response.

Notwithstanding the importance of the truncation level, there is still no known way of determining a reasonable number for use by LDA. Camp and Tung [9] selected their truncation numbers by requiring both the between-group variance $v_b$ and the within-group variance $v_w$ to contribute significantly to the variance ratio $R$. Generally both $R$ and $v_b$ increase while $v_w$ decreases as $r$ increases. When $r$ is above some critical value, we can see that $v_w$ decreases rapidly as $r$ grows, meanwhile the change in $v_b$ becomes relatively tiny. Thus the large separability $R$ is due mainly to the abnormally small $v_w$, which is most likely an artifact of the small sample size and the large truncation number. In practice $r$ is constrained to a certain level so that the relative importance of $v_b$ and $v_w$ to $R$ is substantial. However, this kind of relative importance is hard to be evaluated for many problems and as a result the selection of $r$ through such a process is still quite subjective. From this point

of view, the CMD-P method is preferable since it can always be applied objectively. There are no selectable parameters in this method.

### *2.4.3 Statistical tests associated with LDA*

For an overdetermined system, a "good" separation between pre-defined groups is always obtainable (i.e., the variance ratio can be arbitrarily large) using a large enough truncation number. The spurious results then need to be pruned by statistical tests. A bootstrap Monte Carlo test was developed by Camp and Tung [9] to verify the statistical significance of the separation. The statistical test associated with the LDA is quite similar to the one accompanying the CMD-P method except for the use of the variance ratio $R$ instead of the correlation coefficient $\rho$ as the test statistic, and it is thus named an $R$-test. The LDA method is performed on numerous synthetic datasets constructed by resampling with replacement the observations while preserving the group structure and truncation parameter of the original analysis to create a distribution of variance ratios. The percentile of the observed $R$ within the distribution of synthesized variance ratios denotes the statistical significance of that observed ratio. The use of a uniform truncation number for all synthetic datasets adds complexity to the interpretation of the test results because it is hard to tell if the statistical significance is contributed by the physical signal in the original dataset or by the selection of the truncation number.

It is worth mentioning that in theory the ratio $R$ of the between-group variance to the within-group variance calculated for the CMD-P-derived temporal behaviour $C_1(t)$ can be also used as the test statistic for conducting a statistical test. Recall that the input information for constructing the spatial CMD pattern is rather minimal and less controversial: Only the classification of whether a year belongs to the solar-max or solar-min group needs to be known. Likewise, the LDA procedure merely requires the same minimal information about the groups. Detailed index values like the TSI time series is however necessary for calculating the correlation coefficient $\rho$. If such detailed information is not available, the variance ratio $R$ should be a good alternative. Generally speaking, the $\rho$-test is more powerful than the $R$-test if the solar response follows the TSI variations in time because the

rather large within-group variability in TSI itself adversely affects the value of $R$. In certain cases when the within-group variability is expected to be small (e.g., the QBO phenomenon is suggested by physical considerations to be sensitive only to the sign but not to the magnitude of the zonal wind at the equatorial lower stratosphere), the variance ratio $R$ may also work as a good test statistic.

Chapter 3

# THE 11-YEAR SOLAR CYCLE SIGNAL IN GLOBAL SURFACE TEMPERATURE RECORD



The purpose of this chapter is to demonstrate that a solar cycle response exists in surface temperature using the longest global dataset available, which is in the form of 1854-2007 sea surface temperature (SST), with an emphasis on methods and procedures, statistical tests, data quality and the removal of deterministic signals, such as volcano aerosol forcing and greenhouse gas warming. Using the method of composite-mean difference projection (CMD-P), signals of warming during solar max and cooling during solar min years are found in the global SST over the 14 cycles. The existence of a solar cycle response at the earth's surface is proved at high statistical confidence by Monte Carlo tests. The long-term solar cycle signal is robust provided that the years near the Second World War are excluded, during which transitions from British ships to U.S. ships introduced warm bias in the SST, as recently pointed out by D. Thompson and his colleagues. Contamination of the signal by volcano aerosols is estimated using the multiple CMD (MCMD) method and found to be small over this long record, although ENSO contamination varies depending on the period chosen but is also small.

The multidecadal trend of response to solar forcing is found to account for no more than a quarter of the observed warming in SST during the past 150 yr, under a reasonable but

unproven assumption that the climate response to secular solar forcing and to solar cycle forcing has the same spatial pattern.

### *3.1 A Solar Cycle Response in the 150-Year-Long Global SST*

When dealing with historical data, a major problem is that of data quality, especially during periods of world wars. Camp and Tung [11] and Tung and Camp [143] found a statistically significant global temperature warming at the surface (land plus ocean) during solar max in two reanalysis datasets since the late 1950s, by which time some of the data problems likely had been corrected. Tung et al. [145] additionally found a similar response in the two *in situ* data records during the same period. Questions remain concerning the existence of the solar cycle response at the surface in earlier decades and in century-long records. A simple extension of the previous work, which was done for the period from 1950s on, to earlier periods immediately runs into the period of World War II (WWII), when the data were problematic, as pointed out recently by Thompson et al. [139].

Although "global" surface temperature datasets are available that start from 1880, large continental areas have missing coverage, with the exception of parts of North America, Europe, and Japan. Some datasets fill in the missing data using various methods, as reviewed in [145]. Generally, the solar cycle signal obtained by CMD is smaller in areas where the missing data were filled in, as various interpolation schemes tend to reduce the anomaly to varying degrees. Recent satellite data (used in reanalysis) show that larger responses tend to occur over continents relative to the oceans and that they are larger over the Arctic and Antarctic relative to the tropics. Since these higher response regions are the ones more likely to have experienced severe missing data in the long-term record, it is expected that the global mean signal in the long-term historical record with missing data is smaller than what could have been found in a geographically complete data record.

A related issue on the existence of the solar cycle response is the fact that there were major volcano eruptions that happened to be spaced on a decadal scale during the recent period: Agung in 1963, El Chichón in 1982, and Pinatubo in 1991. Previously, using 100 yr of surface temperature data and optimal filters constructed using a two-dimensional energy balance model, North and Stevens [104] found that the volcano signal contributed signifi-

cantly to the decadal peak in the climate signal spectrum. Such a contamination prevented the authors from detecting the solar signal with confidence, in contrast to their earlier work [135], where a "fairly robust solar signal" was found when other deterministic climate signals (such as volcano eruptions and anthropogenic warming) were ignored. Lean and Rind [86] recently also pointed out that such volcano contamination could affect methods such as Fourier analysis, which is global in time. It should in principle not affect as much the local-in-time methods such as those used by us [11, 143, 145]. In our previous work, we additionally removed two years after major eruptions, when significant aerosol-induced cooling was observed. Nevertheless, it would be reassuring if the solar signal could still be found during periods when the stratosphere was clear of volcano aerosols, or when the period studied is long enough that the time of occurrence of major volcanoes can be taken as random and averaged out when we take the composite means of solar max and solar min and then difference them. A long data record affords us both possibilities. In addition, we will present an analysis using the method of MCMD to show that volcano and ENSO contaminations are small in our solar results.

Previously there have been a number of important papers in the oceanographic literature dealing with the upper oceans' response to the radiative forcing from the sun. Of these, the work of White et al. [155] stands out. They pointed out that since almost 90% of the change in TSI on decadal and interdecadal time scales is at wavelengths that penetrate to the troposphere, it is plausible that direct radiative forcing by the changing solar insolation of the upper ocean can give rise to a solar signal in the SST. Using 92 yr of the Global Ice and Sea Surface Temperature (GISST) data from 1900-91, White et al. [155] obtained a band-passed decadal signal with an amplitude of $0.08 \pm 0.01$°C per $\mathrm{Wm}^{-2}$ of the TSI in the globally averaged (from 40°S to 60°N) SST. The methods used were cross-spectrum and singular spectrum analyses. The peaks of the SST appear to approximately align with the peaks of the TSI except during the beginning of the century and during 1940s and early 1950s; they suspected that the latter discrepancy occurs because of the disruption in the collection of marine data during WWII, which turned out to be the case. Allen [2] applied the multitaper frequency-domain singular value decomposition method to the Hadley Center global surface temperature record from 1871 to 1994 and found a strong spectral peak in the

10-13-yr period, which he called the quasi-decadal oscillation (QDO). A visual inspection of the time series of this QDO now shows coherence with the 11-yr solar TSI variation, although no correlative study was done by the author. Nevertheless it appears that an 11-yr solar signal in global surface temperature exists in Allen's filtered data. White and Tourre [156] similarly found a statistically significant QDO peak in the 93-yr (1900-92) SST spectrum, as did Tourre et al. [141] earlier in the 92-yr (1900-91) SST spectrum, and commented that the time series of the QDO appear to align with the solar irradiance variation. These methods are all of the Fourier type and may be subject to the volcano contamination mentioned above. We hope our work will be able to directly address the contamination due to volcano and other deterministic signals, such as greenhouse gas warming and ENSO. A new method is introduced to separate out these various other contributions.

### *3.1.1 Data*

Currently the longest homogeneous instrumented record of surface temperature is that of sea surface temperature, which spans 150 yr from 1854 to 2007, in the form of extended reconstructed sea surface temperature (ERSST), as described in [129, 130, 131]. [NOAA_ERSST_V3 data are provided by the National Oceanic and Atmospheric Administration (NOAA) Office of Oceanic and Atmospheric Research (OAR), Earth System Research Laboratory (ESRL), Physical Sciences Division (PSD), Boulder, Colorado, and are available at `http://www.esrl.noaa.gov/psd/`.] The dataset is based on the Comprehensive Ocean Atmosphere Dataset (COADS; available online at `http://icoads.noaa.gov/Release_1/coads.html`; [161]). Since 1982, SST is measured directly by satellite with global coverage, in contrast to marine surface air temperature. The global data were separated into "low frequency" (interdecadal) and "high frequency" (decadal) parts, and missing data were filled in using different methods. Of relevance here is the procedure for the high-frequency interpolation. The global data were expanded in empirical orthogonal teleconnections (EOTs), which are similar to empirical orthogonal functions (EOFs) with the exception noted below. The available ship and buoy data were projected onto these to help calibrate the satellite data. Prior to the availability of satellite data, there were

large ocean areas without ship or buoy measurements. The available data were projected onto the leading EOTs deduced from satellite measurements after 1982. The influence of any measurement point is truncated beyond 8000 km and damped beyond 5000 km. These ranges of influence are larger than can be justified but are necessitated by the sparse coverage. The Goddard Institute for Space Studies (GISS) data set, for example, allows a single measurement to influence other grid points only up to 1200 km based on a correlation analysis of the data points [51]. Analyzed in this manner, ocean SST data between 60°S and 60°N appear to be more geographically complete than land surface temperature data, and they greatly influence the global mean temperature used in the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4; [134]). Recently, Thompson et al. [139] found that the global temperature data used in the IPCC AR4 report are problematic during the Second World War, when British ships were replaced by U.S. ships. The U.S. ships measured SST using engine-water intake, which tended to be warmer than the British method of measuring SST on deck from water drawn up using buckets. The authors argued that this might account for the anomalous warming seen in the global temperature displayed in the AR4 report in the 1940s and the subsequent cooling as British ships resumed measurement in the mid-1940s. This warming and cooling were suggested from time to time by some, perhaps erroneously, to be of solar origin, arguing that they were not expected from greenhouse trends. In our current study, the years 1942-50 are excluded from our record as problematic years not yet adjusted in the data record, according to [139]. Their removal resolves much of the sensitivity we were encountering with the historical data with respect to the length of record analyzed.

Additionally, because of sparse data, the ERSST data were heavily damped before 1880, but it is claimed that after 1880 the signal strength was more consistent over time. We originally performed our calculation only for the period 1880-2007. Later when we repeated the calculation for the whole period 1854-2007, the results show very little difference in the overlapping period. Hence the full record, encompassing 14 solar cycles, is shown in Figure 3.1, although we do not have confidence in the spatial patterns of the response before 1880.

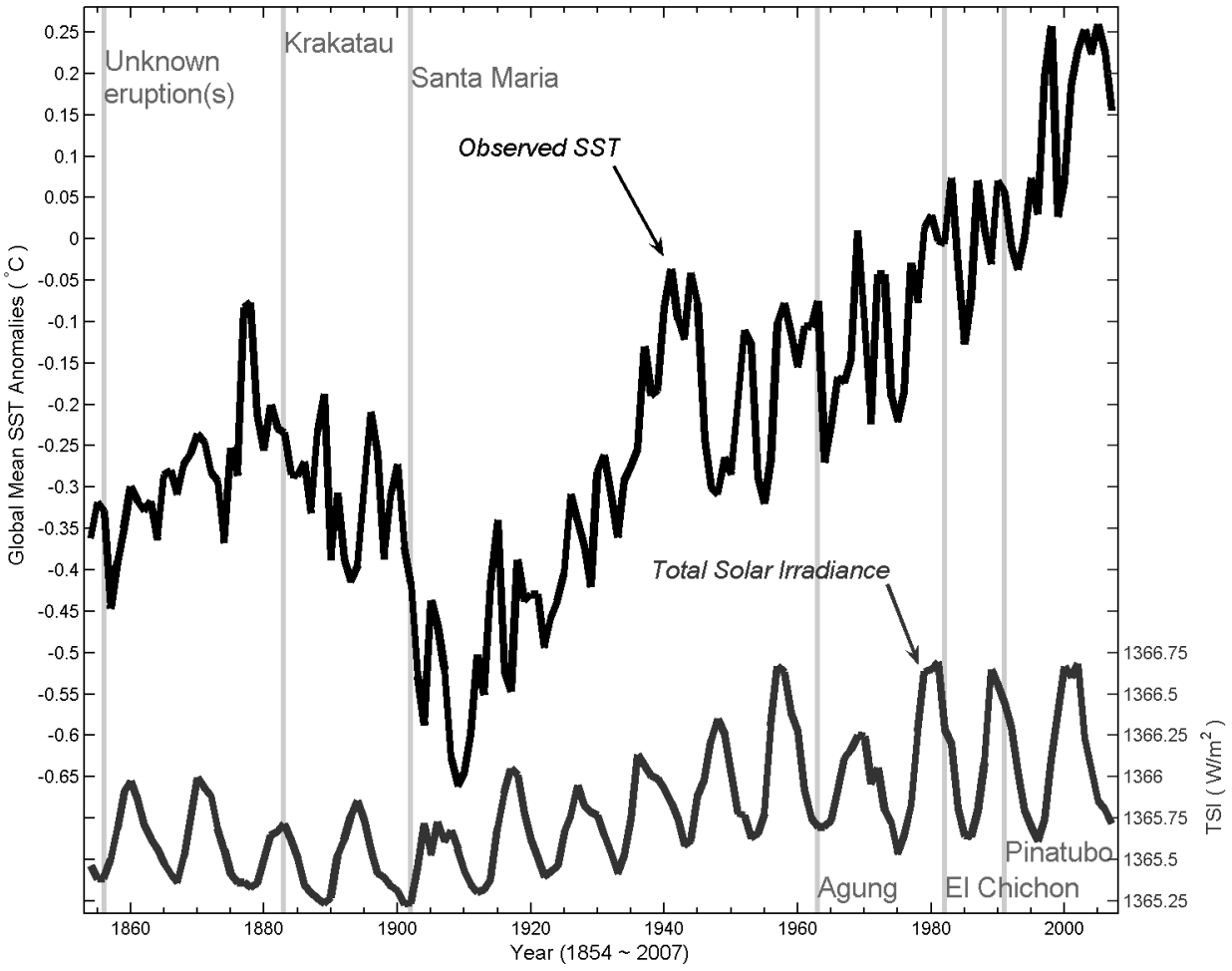


Figure 3.1: (top): SST, averaged from 60°S to 60°N over the ocean areas, from NOAA's ERSST dataset from 1854 to 2007. (bottom): Solar constant (TSI) from Lean et al. [87] and Wang et al. [153].

### *3.1.2 Methodology*

We use the method of composite-mean difference projection (CMD-P). A similar but more sophisticated method is available in the form of linear discriminant analysis (LDA) [127, 9, 143]. The present method was chosen because it is more intuitive and much easier to implement. Briefly, this method separates temperature data into two groups, the solar max group and the solar min group. The separation is done objectively according to the TSI, as discussed below. A global spatial pattern is obtained by composite-mean difference. The original data are then projected onto this CMD spatial pattern, yielding a time series that may or may not vary in phase with the solar (TSI) time series. The method is successful when the correlation is high. The correlation of these two time series is tested using a Monte Carlo simulation. The unknown atmospheric population distribution is estimated by bootstrap resampling with replacement of the original temperature data, by assigning a year to either a solar max or solar min group randomly while preserving the number of years in each group. The same CMD projection method is applied to this synthetic data to produce a time series. The percentage of the time when this randomly generated time series has a correlation coefficient with the TSI equal to or higher in magnitude than the observed is noted, and this number is often less than 0.02%. To take into account the inherent autocorrelation of the climate data, the resampling is repeated using a block of $L$ consecutive years, where the length of $L$ is supposed to be determined by the autocorrelation time of the time series. As mentioned in Chapter 2, it is not an easy task to choose an appropriated value for $L$. For our problem, the block length calculated using the formula (2.6) varies from 1 to about 20 yr depending on the spatial location, even if the temperature data are assumed to follow AR(1) processes. To circumvent this complication, we simply repeat the calculation for $L = 1, 2, 3, ...., 10, 11, 12$, etc., and report the lowest confidence level obtained, which occurs at $L = 10$. There is, however, not much change (less than 0.02%) between $L = 1$ and $L = 10$ for the decadal signal under study. The same holds true when we repeat the previously published results in [11] and [145], using $L = 10$ instead of $L = 1$. That is, all results previously deemed to be statistically significant at above 95% confidence level remain so using the moving-block resampling method.

Figure 3.1 shows the global mean (from 60°S to 60°N) and annual mean SST from ERSST described above for the period 1854-2007, along with the annual mean TSI from Lean et al. [87] and Wang et al. [153] extended to 2007 using data kindly provided to us by J. Lean. It is visually apparent that there exist non-uniform trends in both the SST and the TSI. There was a pronounced cooling of over 0.6 K in the globally averaged SST in a short period of time from 1895 to 1910. Then the SST warmed by an even larger increment of 0.8 K from 1910 to 1945. In 1945 there was the sudden anomalous drop in SST studied by Thompson et al. [139], followed by the modern global warming of 0.5 K until 2007. This latest warming is usually attributed to the increase in greenhouse gases. The warming from 1910 to 1945 has been attributed to the solar forcing in some studies, because the TSI coincidentally also increased during this period (see Figure 3.1). The solar max of 1910 was abnormally weak and the solar max of 1955-60 was abnormally strong, and there was a general increasing trend in between. We will show, however, that this trend in TSI during the period 1910-45 was too weak to account for the "observed" warming, which was likely due to the data problem described above.

### *3.1.3 To detrend or not to detrend*

In the period 1959-2004 previously analyzed by Camp and Tung [11], the TSI from Lean et al. [87] has little or no trend. In the longer-term record we are analyzing here, the presence of the nonuniform trend, also evident in Lean et al. [87], makes some solar max TSI values in the earlier part of the record lower than even the solar min TSI values in the more recent part of the record. Since on physical grounds it is the absolute irradiance that matters, with the higher TSI warming the earth more than the lower TSI does, it is not clear that a trend should be removed to center the TSI data. To compound the problem, the magnitude of the trend in TSI is uncertain and is currently under debate; see the IPCC AR4 report ([33], p.132). Since the controversy concerning the TSI reconstruction is related to the secular trend of the TSI and generally is not about the classification of whether a year belongs to solar max or solar min (see, e.g., [125]), we have decided not to detrend but instead to implement a *pairwise differencing* procedure which requires only the latter

minimal information. We divide the TSI time series into subperiods each containing just one whole solar cycle (with one solar max and one solar min). Since there is very little TSI trend within a decadal period, the solar max (min) years are defined as the years in which the TSI is 0.06 $Wm^{-2}$ above (below) the mean TSI for that particular short subperiod. [The 0.06 $Wm^{-2}$ threshold was introduced by Camp and Tung [11] so as not to count years as either solar max or solar min when their TSI variations are within ∼10% of the mean peak variation of ±0.6 $Wm^{-2}$.] This grouping/identification is objectively done for each solar cycle period. The CMD is performed on the SST data one solar cycle period at a time by taking the difference between the temperature at solar max years and at solar min years. This difference for each solar cycle period is then averaged over all solar cycles in the longer data record. This method works well even with undetrended data when the secular trends are small. During the last three decades, however, somewhat different results are obtained when a subperiod is defined as solar max following a solar min versus a solar min following a solar max. This problem is remedied by the procedure of *pairwise differencing with shift*, as described in Chapter 2 and below.

The monotonic positive trend in surface temperature in the recent decades may be due to forcing agents other than the TSI. An obvious candidate is the increase in greenhouse gases. To remove this contamination, we perform the above-described pairwise differencing with the following modification. A whole solar cycle subperiod is first defined as solar min following solar max. Then we repeat the procedure but by defining a wholee po solar cycle subperiod as solar max following solar min. This is done by shifting the years comprising a solar cycle forward by half a cycle. The CMD spatial pattern that we will use is obtained by averaging the patterns obtained with these two definitions over this one and a half period. If there is a positive 5-10-yr trend that exists within a solar cycle, it would manifest itself by yielding a higher CMD warming if the solar max follows the solar min than if the solar min follows the solar max. The averaging then eliminates the short-term trend that might be present within a solar cycle, because the positive and negative contributions of the trend to the CMD cancel each other locally (within that one and a half solar cycles). Interdecadal variations are not removed. Previously, in [11], the linear trend that exists in the temperature record of 1959-2004 was removed by linear detrending. This is not feasible

in the 150-yr data because no single linear trend exists. Piecewise linear trend removal introduces artificial jumps in temperature, which is undesirable. Our method of pairwise differencing with shift works very well and greatly reduces the sensitivity we have had in our previous trials with trend removal.

### *3.1.4 CMD projection and statistical tests*

Figure 3.2 shows the longitude-latitude distribution of warming and cooling obtained by CMD (*pairwise differencing with shift*), as described in the previous section, for the period 1854-2007. The spatial distribution prior to 1880 is probably not reliable. Therefore we repeated the calculation using the period 1880-2007, shown in Figure 3.3. The differences between the two relate mostly to the fact that the amplitudes of the warming and cooling centers are slightly larger in Figure 3.3, probably because the data prior to 1880 are heavily damped in the dataset. This spatial CMD pattern is denoted $P_1(x)$. In the CMD projection method, the original SST data are projected onto $P_1(x)$ to obtain its associated temporal coefficients $C_1(t)$, as described in Chapter 2. By the definition of $C_1(t)$, the "solar cycle" mode $P_1(x)$ is orthogonal to sum of all the remaining modes, which theoretically include all other variability and noise. The lower panel in Figure 3.2 shows the projected time series $C_1(t)$ in blue. It is the time variation of the solar response in the SST data corresponding to the spatial pattern shown in the top panel. For convenience of presentation, $C_1(t)$ is additionally normalized by the global mean of $P_1(x)$, so the magnitude of $C_1(t)$ is interpretable as the magnitude of the globally averaged SST variation in response to the solar TSI variation.

Looking at the time series of solar cycle response in Figure 3.3, we see that the solar max warms relative to the solar min in globally averaged SST in the 13 solar cycles examined. There was not a phase reversal in 1920 or during any other period. The amplitude of the global SST response is about 0.1°C per each $\mathrm{Wm}^{-2}$ [the scale of TSI and $C_1(t)$ are scaled 1 $\mathrm{Wm}^{-2}$ to 0.1°C to facilitate this comparison]. There are, however, a few cycles in which the amplitude is smaller, and this can usually be attributed to questionable data. When regressed over all cycles (excluding, however, the period 1942-50 mentioned earlier),

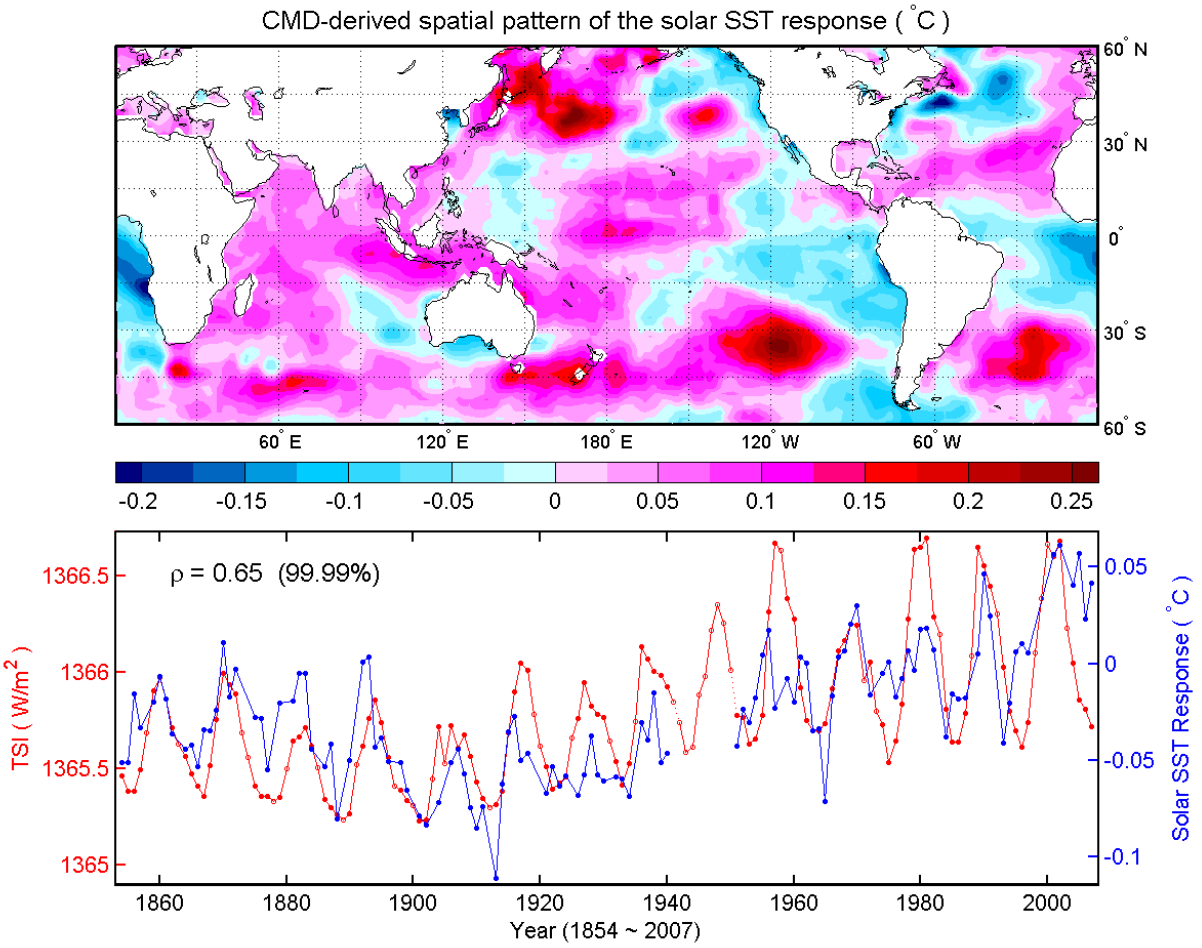


Figure 3.2: (top): For the period 1854-2007, the spatial pattern of SST obtained from CMD. (bottom): The time series obtained by projecting the original data onto this pattern (blue) and the TSI index (red). The correlation coefficient of the two curves is given by $\rho = 0.65$, which is statistically significant at above the 95% confidence level (99.99%) using bootstrap resampling with 10-yr blocks of data.

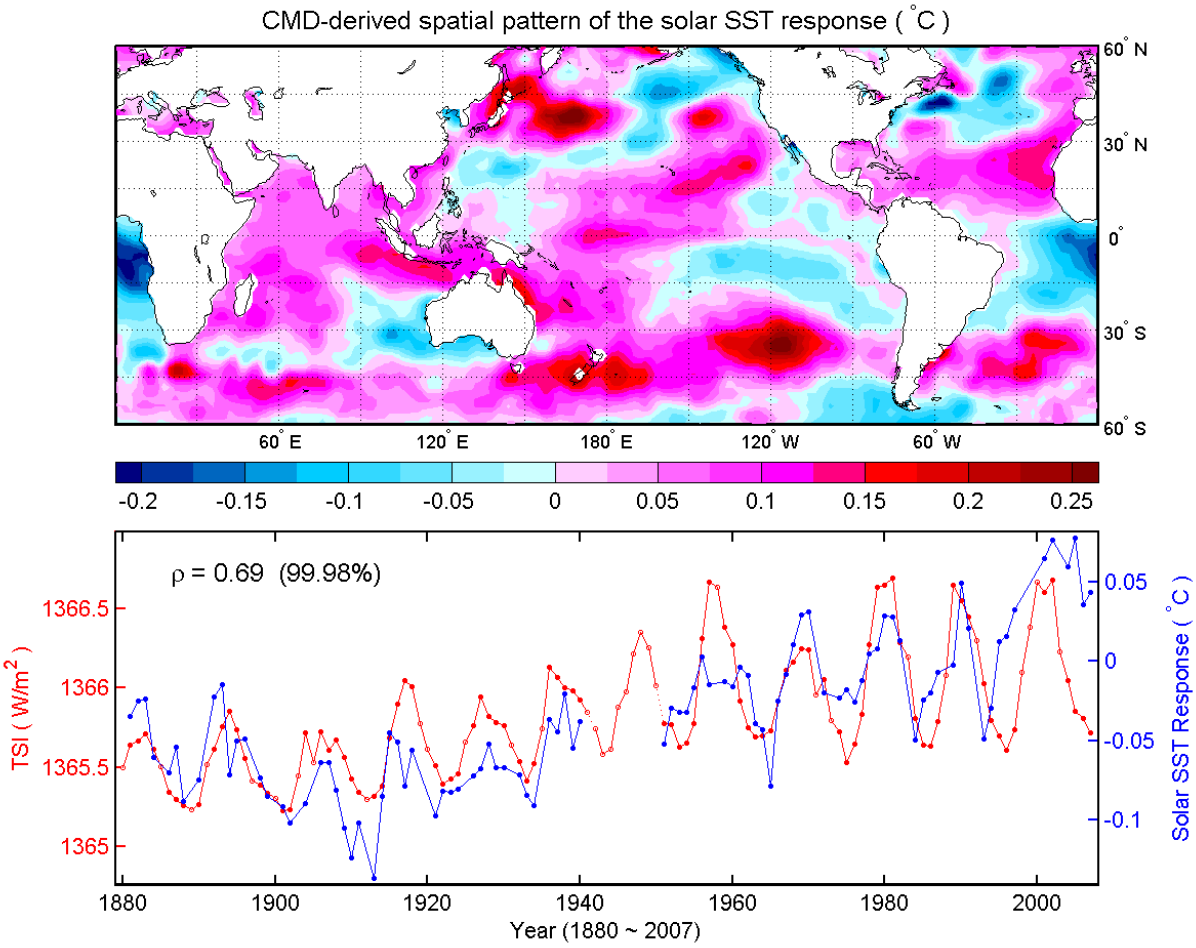


Figure 3.3: As in Figure 3.2, but for 1880-2007.

including periods of remaining bad data, the warming in globally averaged SST (between 60°S and 60°N) is $\kappa \sim 0.085$°C per Wm$^{-2}$ for the period 1880-2007. This amplitude is about the same that found by White et al. [155] for the period 1900-91. The solar cycle response amplitude found here for the SST is about 60%-70% of that found in the land-ocean average found by Tung et al. [145] for the *in situ* data of GISS and HadCRUT3. This finding is consistent with the value of $\kappa = 0.12$°C per Wm$^{-2}$ found for the land-ocean average in those two *in situ* datasets because the warming is usually stronger over the continents. The ratio of this ocean average to the global average of land and oceans is even smaller in the recent reanalysis data, also shown in Tung et al. [145], which includes areas poleward of 60°N and 60°S with amplified warming, not included in the *in situ* data.

There is a pronounced cooling trend after the eruption of Santa Maria in October 1902 that lasted more than a decade, longer than can be expected from volcano aerosol cooling. Interestingly, this cooling does not project onto the solar response pattern, indicating that this severe cooling may be due to noise or more likely bad data, and is effectively filtered out by our projection method. During the decade after WWII the solar max response was smaller than expected from the TSI. The global SST was actually very warm during that solar max (see the black line in Figure 3.1), but it does not project onto the solar response spatial pattern. This is an indication that the spatial pattern of the SST anomalies during that decade was atypical, because the mix of British and U.S. ships was changing [139]. The WWII years likewise do not project significantly onto the solar response spatial pattern (not shown) if the latter is obtained for the period 1854-2007 with the WWII years excluded. Hence the warming and cooling during that period do not appear to be solar related. Nevertheless the erroneous temperature discontinuity is so large that if the WWII years had been included in our calculation of the spatial patterns, they would have contaminated that pattern.

The correlation coefficient $\rho$ between the temperature response $C_1(t)$ and the TSI is about 0.69 for the period 1880-2007 and 0.65 for the period 1854-2008, both quite high for such a long data record and extremely unlikely to be producible by chance if there were no solar cycle signal in the SST (the null hypothesis). Figure 3.4 shows the distribution of $\rho$ in 10,000 synthetic SST time series generated using the method of moving-block bootstrap with

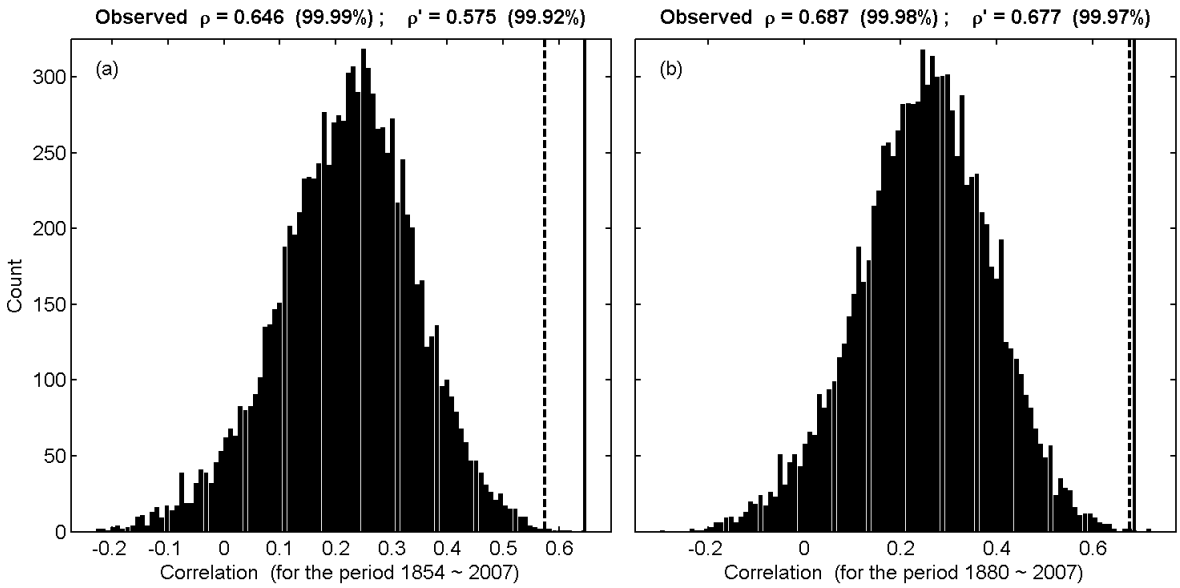


Figure 3.4: Distribution of the correlation coefficients between the projected SST time series $C_1(t)$ and the solar TSI index, obtained using synthetic data generated by the bootstrap resampling Monte Carlo method for the periods (a) 1854-2007 and (b) 1880-2007. The solid vertical line is the observed $\rho$, and the dashed vertical line is the observed $\rho$ obtained when the data was first detrended, denoted by $\rho'$. Note that here $C_1(t)$ is not normalized by the global mean of $P_1(x)$, which explains why the distribution is not bimodal.

replacement, as described in Chapter 2 and the previous section of *Methodology*. We find that very few (less than 0.1%) of the 10,000 synthetic SST time series achieve a correlation equal to or higher than the observed value. Thus, we have effectively ruled out the null hypothesis that our method can by random chance generate an apparent "signal" that is highly correlated with the TSI when no real solar signal exists in the data.

One may be suspicious of this high confidence level and question whether the correlation is enhanced by the presence of the long-term trend in the observed time series, with the temperature in earlier decades before WWII lower than in the more recent decades after the war, while in the synthetic data there is no consistent trend because of the scrambling of the years. It turns out that unlike the regression coefficient $\kappa$, the correlation coefficient is not sensitive to the presence of trend in $C_1(t)$. When we remove the trend in $C_1(t)$ before correlating it with TSI, $\rho$ is changed only slightly. The results of such a calculation are indicated in Figure 3.4.

### *3.1.5 Spatial features in ocean basins*

We will next discuss the features in Figure 3.3, obtained using the better data since 1880. We will show that the response over oceans exhibits both warming and cooling distributed in some characteristic patterns, more so than the warming over continents found in previous work. The ocean area-averaged temperature is therefore smaller than the local SST anomaly, which ranges from $-0.2^\circ$ to $+0.2^\circ$C. In the Atlantic Ocean, the tropics are cold south of the equator but warm a little north of it. The western north Atlantic is cold. The Indian Ocean is warm. These features are robust. The robust basinwide warming in the Indian Ocean – a small ocean basin – is indicative of a radiative response to solar forcing, in contrast to the situation in the larger ocean basin of the Pacific, which is capable of fast dynamical responses involving coupled atmosphere and oceans [99, 141, 156].

In the Pacific Ocean, there is a robust warming center located in the northwestern Pacific and cooling off the west coast of the United States. There is generally cooling in tropical eastern Pacific, with the exception of a thin warming strip located in the equatorial Pacific, where the ENSO variance is large. The warming center in the northwestern Pacific is robust,

but the warming strip in the eastern Pacific is not (cf. Figures 2 and 3, and see later figures).

Recently, van Loon et al. [148] and van Loon and Meehl [147] studied specifically the spatial pattern in the Pacific during northern winter using the same ERSST data since 1854. They calculated their composite mean difference by taking the difference of the mean of the "solar peak years" (one year per solar cycle) and the climatology, in effect using only 14 degrees of freedom. The climatology was calculated over a different period than that from which the solar peak years were chosen [The period used in the climatology calculation was based on only 30 yr, 1950-79, in [148]. A different 29 yr, 1968-96, was used in [147].] Over the equatorial eastern Pacific, they found a cold event (La Niña)-like condition, which was deemed statistically significant by the Student's $t$-test. Our Monte Carlo test of bootstrap resampling cannot be applied to their methodology because there is only one data point in each solar cycle. There is no time series information on the response for us to test the similarity between the response and the forcing when only one year is used for each solar cycle. The Student's $t$-test they used does not actually test whether the signal is solar related; it merely tests if the mean of the solar peak years is significantly different from the mean of the years used in defining the "climatology". It is in this regard that the subjective choice of the years used in the calculation of climatology affects the result of the Student's $t$-test. Because the period 1968-96, chosen by van Loon and Meehl [147] for the climatology, is warmer, it yields a larger-amplitude equatorial Pacific SST cold tongue when it is subtracted from the solar peak mean, and therefore it passes the Student's $t$-test. This is their best result, reproduced here in the top panel of Figure 3.5. The yellow contour encloses regions of statistical significance at the 95% confidence level, and we see that the cold tongue at the eastern Pacific and a warm pool over the northwestern Pacific are both statistically significant, as discussed in detail by van Loon and Meehl [147]. This result, however, is not robust to the choice of either the so-called solar peak years or of the base period for the calculation of climatology. The composite difference SST field shown in the middle panel in Figure 3.5 was constructed in the same way as in [147] except that the peak solar years are chosen objectively according to the peaks in TSI. The spatial pattern is rather different – the La Niña pattern is disrupted – but nevertheless the eastern Pacific is cold and still statistically significant. This changes again when the full 1854-2007 climatology

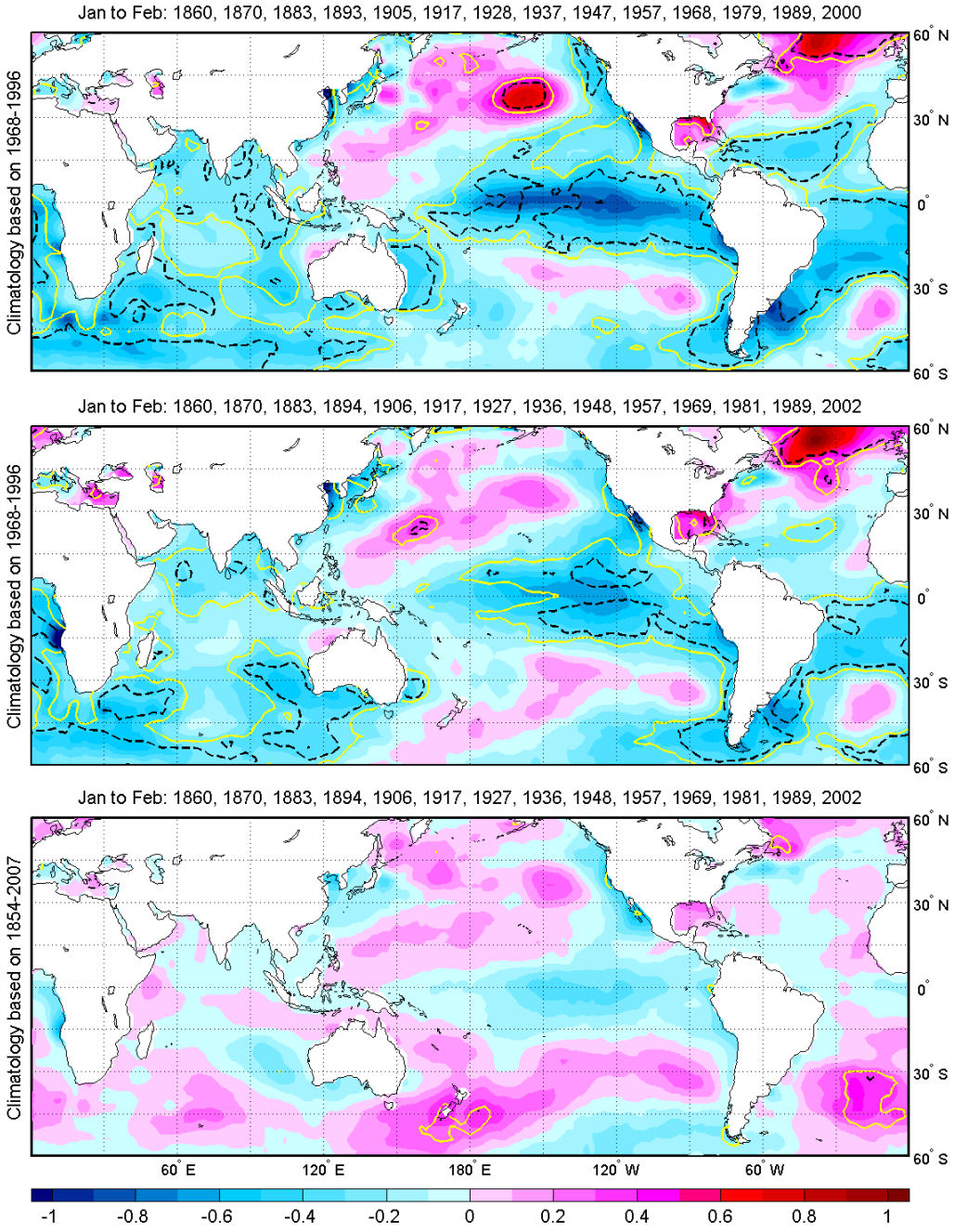


Figure 3.5: Difference in SST between the mean of the "solar peak years" (indicated at the top of the figure) during January to February over the period 1854-2007 and the "climatology" (computed for periods indicated along the left edge of the figure). Yellow contours enclose regions of 95% confidence level; dashed black contours enclose regions of 99% confidence level. (top): Solar peak years determined by peaks in sunspot number; climatology determined by the period 1968-96 chosen by van Loon and Meehl [147]. (middle), (bottom): Solar peak years determined by peaks in TSI; climatology determined by (middle) the period 1968-96 and (bottom) the whole period of the data record, 1854-2007.

is used. This is the most objective way of constructing the composite difference and the result is shown in the bottom panel of Figure 3.5. None of the features in the Pacific are statistically significant by the Student's $t$-test.

The hypothesis that it is the solar peak years that causes the La Niña-like response in the equatorial Pacific [147, 148] and that one or two years later the response switches to an El Niño-like pattern [99] may still be correct, and it appears to be supported by modeling results as reported in [100]. But it may be a reflection of ENSO-related lag correlations in SST that bear no relation to the sunspot cycle. In any case, observational support for this hypothesis is not yet available. It is likely that 150 yr of data is not long enough for us to separate out different behaviours in the first versus second year of a solar max.

The question of whether the equatorial Pacific responds to a warmer climate in a La Niña- or an El Niño-like pattern is under debate in the context of global warming. Vecchi et al. [149] showed that the ERSST data we are using exhibit a long-term trend in the form of an El Niño- like pattern while a different SST dataset, HadISST, exhibits a La Niña-like pattern. They attributed the difference to the difference between the two datasets in two periods: the 1930s and the 1980s, which corresponded to the periods of greatest change in the "buckets-to-intake" correction of SST measurements previously implemented (i.e., prior to [139]) and the beginning of SST retrievals using satellites.

### *3.2 Multidecadal Trends*

Since our method does not involve detrending of temperature or TSI, there is a secular SST response seen in Figure 3.2 to the secular trend in the solar forcing. Generally, the level of SST solar response is consistent with the level of TSI forcing, with periods of high positive values of $C_1(t)$ associated with periods of high TSI. By regressing $C_1(t)$, using just the solar min years, or just the solar max years, or the entire time series, onto the years to determine the slope of the time series, the amplitude of the global SST trend attributable to the solar influence is found to be about $0.004 \pm 0.0012$°C per decade for the period of 1854-2007 or $0.009 \pm 0.0017$°C per decade for the period 1880-2007. These bracket the solar trend over the last century, $0.007 \pm 0.001$°C per decade, reported by Lean and Rind [85]. However, this slope of $C_1(t)$ is not a robust quantity since the actual trend is nonlinear.

To set an upper bound on the solar forcing contribution to the warming trend, we give two maximum values, one based on only peak solar max years and one for peak solar min years: $C_1(t)$ warmed by 0.18°C from the solar max of 1909 to the solar max of 2002. During this period the global mean SST warmed by 0.89°C, and so no more than 20% of that may be attributed to solar forcing during this period. A larger warming of 0.21°C is found from the solar min of 1913 to the solar min of 2005. This last number, 0.21°C, is deemed the upper bound in the secular change in SST that can be attributed to solar forcing, first because that is the difference between the lowest and the highest temperatures in the solar min in the entire record, and second because some greenhouse warming residue may arguably remain in $C_1(t)$ during the most recent solar cycle (possibly since the solar min in 2007 is the last half cycle analyzed) despite our best efforts to remove it. During this same period of time, the global mean SST warmed by 0.81°C, and so no more than 26% of it can be attributed to solar forcing. These are upper bounds; the true solar-induced trend is probably smaller. These changes in SST associated with the interdecadal changes in solar forcing are quite modest and in no way can account for the observed warming trend in SST during the last century (see Figure 3.1). The latter must have been caused by other forcing agents, including anthropogenic ones.

### *3.3 Volcano, ENSO, and Greenhouse Warming Contamination*

When there is a long enough data record, within-group variances caused by volcanoes and ENSO, which are not consistently correlated with the solar cycle, are hopefully greatly reduced by the composite means and by the differencing of the two groups. Nevertheless, how well these variances are removed is a continuing concern. In the analysis shown in Figures 3.2 and 3.3, no years with volcanic eruptions were removed before processing, unlike the procedure in [11]. The result is not so different from that obtained (not shown) by excluding the volcanic years from the analysis. The time series $C_1(t)$ is highly correlated with the solar index ($\rho = 0.69$) and not correlated with the volcanic aerosol index [121]; the latter correlation coefficient, $\rho_{\mathrm{AI}} = -0.08$, is practically zero. We will show directly below that volcanic contamination is indeed very small. Global warming due to increases in greenhouse gases is another important source of contamination in the solar signal. Nonetheless,

the method that we introduced in the section entitled "*To detrend or not to detrend*" to obtain the solar signal reduces this contamination greatly, as we will quantify below. ENSO is a prominent source of variability in the Pacific Ocean and can affect significantly the SST patterns studied here, more so than the land-ocean patterns studied previously. In the present study, extreme ENSO years, defined as years in which the winter [December-February (DJF)] mean cold tongue index (CTI) exceeds 1.2°C in absolute magnitude, are excluded in the analysis presented in previous sections. The resulting $C_1(t)$ has a correlation coefficient with the annual mean CTI index of $-0.13$ for the period 1880-2007, which is small enough for the ENSO contribution to the derived solar signal to be deemed negligible. To verify that these contaminations are already small in what we have produced, we shall now try to separate out these four deterministic signals and show that our results relating to the solar cycle response are not changed.

In a typical error analysis, one assumes that the data consist of the signal under study and a remainder, called "noise." A noise model needs to be constructed; usually either a random white noise model or a red noise model is assumed. As pointed out by North and Stevens [104], neither of these noise models is appropriate for analyzing climate data because the data contain prominent deterministic signals such as ENSO and volcanic aerosols, that need to be explicitly taken into account. We shall assume that our data consist of the following four deterministic signals: solar, ENSO, volcanic, and anthropogenic greenhouse gas increases; and the residual noise is assumed to be random.

Here we employ the MCMD method (see Section 3 of Chapter 2) on the data to estimate the spatial patterns of the four signals for the period 1880-2007, and the results are shown from top to bottom in Figure 3.6. Each column of the weight matrix for performing the MCMD procedure is formed using one corresponding climate forcing. We pick the years and group them to compute the CMD as follows (the years 1942-50 are always removed beforehand): The solar max (min) years are defined according to the TSI index as having a TSI 0.06 $\mathrm{Wm}^{-2}$ greater (smaller) than the local mean of a complete solar cycle. The warm (cold) ENSO years are defined as the years when the annual mean CTI is greater (less) than 0.25°C(-0.25°C). The years with volcanic eruptions are 1883-85, 1902-04, 1963-65, 1982-84, and 1991-93, including three years after each major eruption indicated in Figure 3.1. The

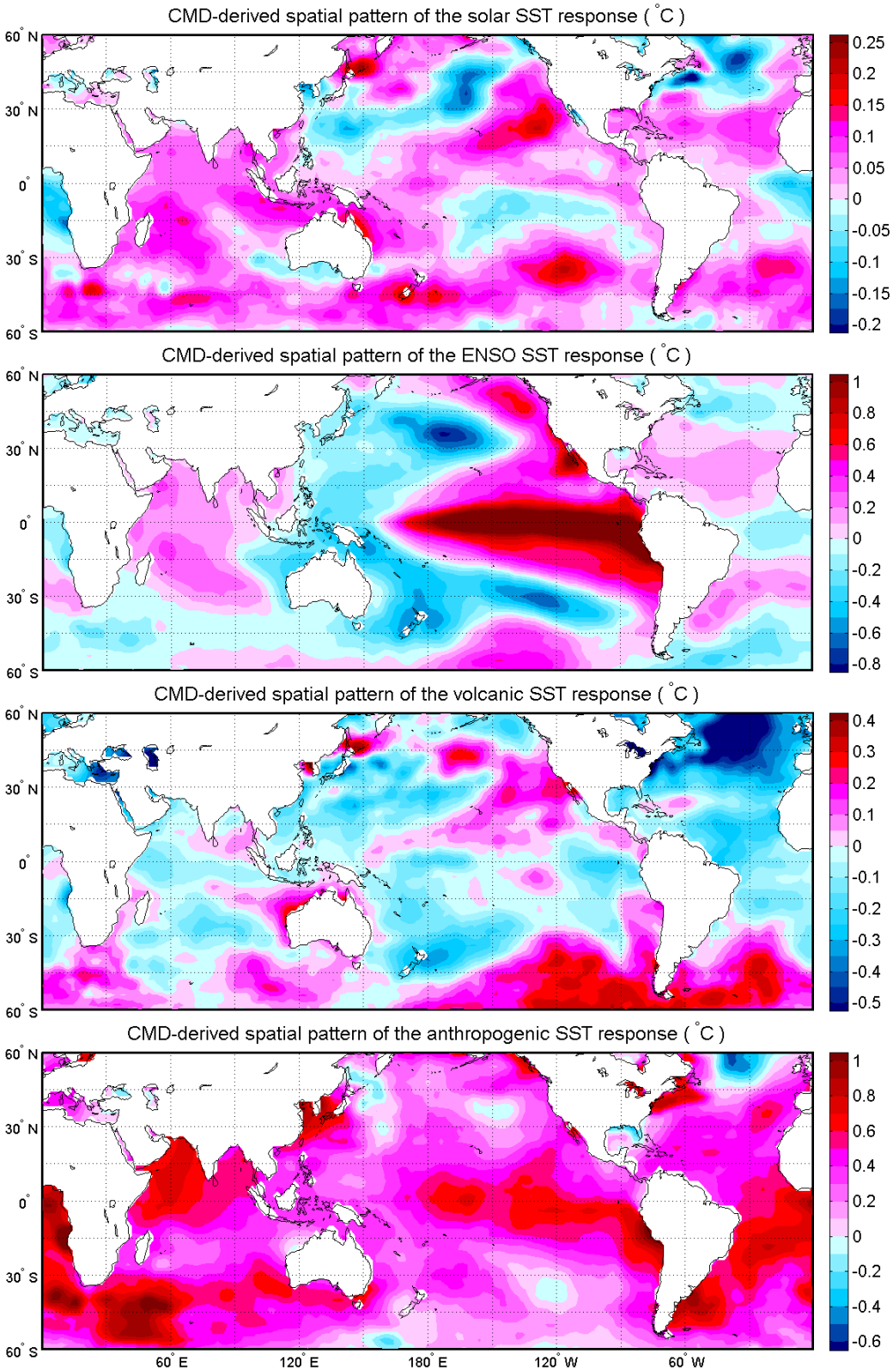


Figure 3.6: Spatial patterns in SST response derived based on the MCMD method for (from top to bottom) solar cycle, ENSO, volcano aerosols, and greenhouse gas increases.

nonvolcanic group contains years when the annual mean aerosol index is no larger than 0.005 optical depth. The two anthropogenic groups comprise the years in which the global mean $CO_2$ mixing ratios [52] are 10 ppm above or below the mean of the entire period of the data record, 1880-2007.

It is worthwhile pointing out that the column elements of the weight matrix corresponding to the solar forcing are derived by applying the simplest composite mean differencing instead of the pairwise differencing with shift. For example, let $P_1^S(x)$ be the simple difference between the mean temperature of all the solar max years and that of all the solar min years during the whole period. This simple differencing alone does not give a good estimate of the true solar cycle spatial pattern (not shown) because it contains other deterministic signals such as volcanic aerosols and greenhouse gas warming. However, it is only used as an intermediate step in the MCMD procedure, which is supposed to remove these deterministic contaminations. The surprising finding is that this estimate obtained via MCMD (in the top panel of Figure 3.6) is very close to that obtained using the method of single pairwise differencing with shift (shown in Figure 3.3) which does not explicitly take into account the other deterministic signals. The latter method is effective in removing the secular trend, which is presumably due to the anthropogenic greenhouse gases also removed in MCMD. The length of the record serves to average out the volcanic and ENSO contaminations, yielding very small differences in the spatial patterns calculated by these two very different methods.

### *3.4 Estimating the Transient Climate Response*



The Transient Climate Response (TCR) is defined in IPCC 4th Assessment Report (AR4) as the global mean warming in response to 1% per year compound increase in $CO_2$ at the time of its doubling. TCR is deemed more relevant in calibrating models on their ability to predict the warming resulting from transient increases in $CO_2$ than the Equilibrium Climate Sensitivity (ECS), which is defined as the equilibrium global mean surface change

[†]An edited version of this section was published by AGU. © Copyright (September 12, 2008) American Geophysical Union.

at doubled $CO_2$. The coupled atmosphere-ocean models participating in AR4 produce a range of TCR from 1.2 to 2.6 K [114]. This rather large range is difficult to constrain with observations, since the transient response does not easily discriminate between models with different climate feedback processes [49]. Previously Stott et al. [137] used the observed 20th century temperature change to constrain three models (HadCM3, GFDL-R30 and PCM) and then applied these models to the calculation of TCR for the future. The estimated TCR was around 2.1 K with the 5-95% probability range is 1.5 to 2.8 K.

In this section, we propose that the temperature response at the earth's surface to the 11-year solar cycle variation in total solar irradiance (TSI) can yield a useful constraint on the transient climate response. Multiple datasets, including reanalysis (NCEP/NCAR and ERA-40) and blended *in situ* land-ocean data (GISS and HadCRUT3) will be used in our analysis.

### *3.4.1 Datasets*

The solar-cycle temperature signal near the earth's surface is discernible in the presence of the larger unforced variability in our climate because its globally coherent spatial structure is mostly of one sign (warming) meridionally. The coupled atmosphere-ocean system naturally produces decadal variability of larger amplitude, but the unforced variability often takes the latitude-compensating form of annular modes of warming and cooling [98] and so can be filtered out using a spatial filter or a simple global average. ENSO, although an internal mode of oscillation in the atmosphere-equatorial ocean system, appears to the atmosphere as an "externally forced" response, in the sense that it contributes to changes in global-mean temperature. Nevertheless, the ENSO spatial pattern is different from the pattern of the solar response, with warming in the tropics and cooling in the mid-latitudes [128]. This and the removal of volcanic-aerosol- induced cooling and the secular trend of global warming, in addition, has been discussed previously [11, 143]. Briefly, volcano aerosol cools the surface, with most of the cooling occurring during the first two years after the volcano eruptions. Two years are excluded from our analysis after El Chichón and after Pinatubo eruptions. There may be some residual delayed cooling not removed by this method, but

this effect is believed to be small. A linear trend is removed before the analysis. Our analysis method emphasizes the difference between interleaved intervals of about 5 years in the temperature records, thereby minimizing the impact on the analysis of any residual secular anthropogenic effect not removed by a linear trend.

To extract the solar-cycle response signal by taking advantage of its spatial signature, it is preferable to use dataset that is globally complete. This was the reason that in previous work the geographically complete reanalyzed datasets of NCEP/NCAR and ERA-40 were used [11, 143]. Both reanalysis data use available station measurements, plus satellite, buoy and other forms of data. These are assimilated by a model, which dynamically supplies the missing information for one variable from constraints provided by other variables. In NCEP/NCAR [69], the surface air temperature is derived from observations of upper air variables and surface pressure. In ERA-40 [146], the surface temperature is called the 2-m temperature. It is not obtained directly as part of the three-dimensional variational analysis of atmospheric fields, but is an interpolation from the lowest model level (at $\sim$10-m) and the background forecast of the skin temperature. Without supplementation by satellite or other data, datasets based on *in situ* station measurements of surface temperature have large areas with missing or sparse coverage; these include the Antarctic, the Arctic, and the central African, South American, and the northern Asian continents. Interpolation in time and in space tends to reduce the amplitude of the response, which depends on the amplitude of the anomalies. These smaller amplitudes serve here as a lower bound for the solar-cycle response.

The land component of the Goddard Institute for Space Studies (GISS) global surface temperature dataset [51] consists of monthly mean station data from the Global Historical Climatology Network (GHCN) version 2 of [108] and the Scientific Committee on Antarctic Research (SCAR) data from Antarctic stations. All station records within 1200 km of a grid point are averaged. In data-sparse regions, a single station is used to fill in the estimated temperature up to 1200 km. The ocean component uses the sea-surface temperature (SST) [118] rather than the marine air temperature (MAT) because of historical measurement non-uniformity (with respect to ship height and speeds) associated with the latter. From 1982 on, satellite measurements of SST are used, calibrated with the help of thousands of

ship and buoy measurements. The same satellite-derived empirical orthogonal functions (EOF) were applied to the period prior to satellite observation [93]. Ship measurements were fitted into these predefined EOFs, which were then used to extend to regions without ship measurements. The Reynolds and Smith SST data are not defined south of 45°S.

HadCRUT3 [7] is the latest version of the historical blended air surface temperature over land and SST over ocean. The SST in HadSST2 [116] consists of a gridded dataset from *in situ* ship and buoy observations from the new International Comprehensive Ocean-Atmosphere dataset (ICOADS). Over 4000 land stations are used, with additional monthly data obtained from stations in Antarctica. Infilling of missing grid box values using data from surrounding grid boxes, used in the previous versions, is no longer done. Consequently coverage is sparsest over the interior of the continents of Africa and South America, and over Antarctica.

The period considered is from 1959-2004 for NCEP, GISS and HadCRUT3. ERA-40 is available only up to 2002. Figure 3.7 shows the 2D composite mean difference of the surface temperature between the solar max years and the solar min years. The solar max (min) years are defined as the years when the TSI is above (below) the record mean value, with a few years near the mean excluded (see [11]). Missing data areas are left blank, and serve to show that an *in situ* dataset such as HadCRUT3 is missing data over large areas in the interior of the continents. This situation has not improved in recent decades. Figure 3.7 also shows the effect of interpolation schemes used in filling in the missing data in GISS. It is seen that the spatial features revealed by all four datasets are very similar. Not surprisingly, the *in situ* datasets with their many regions of missing data requiring interpolation show smaller anomalies than the reanalysis data. The results for the two geographically complete datasets, NCEP and ERA-40, are strikingly similar in the latitude and longitude locations of warming and cooling, except for the rather larger cooling in Siberia seen in ERA-40 than in NCEP. GISS data are more similar to NCEP than to ERA-40 in the Arctic region, with no zonally averaged cooling near 70°N.

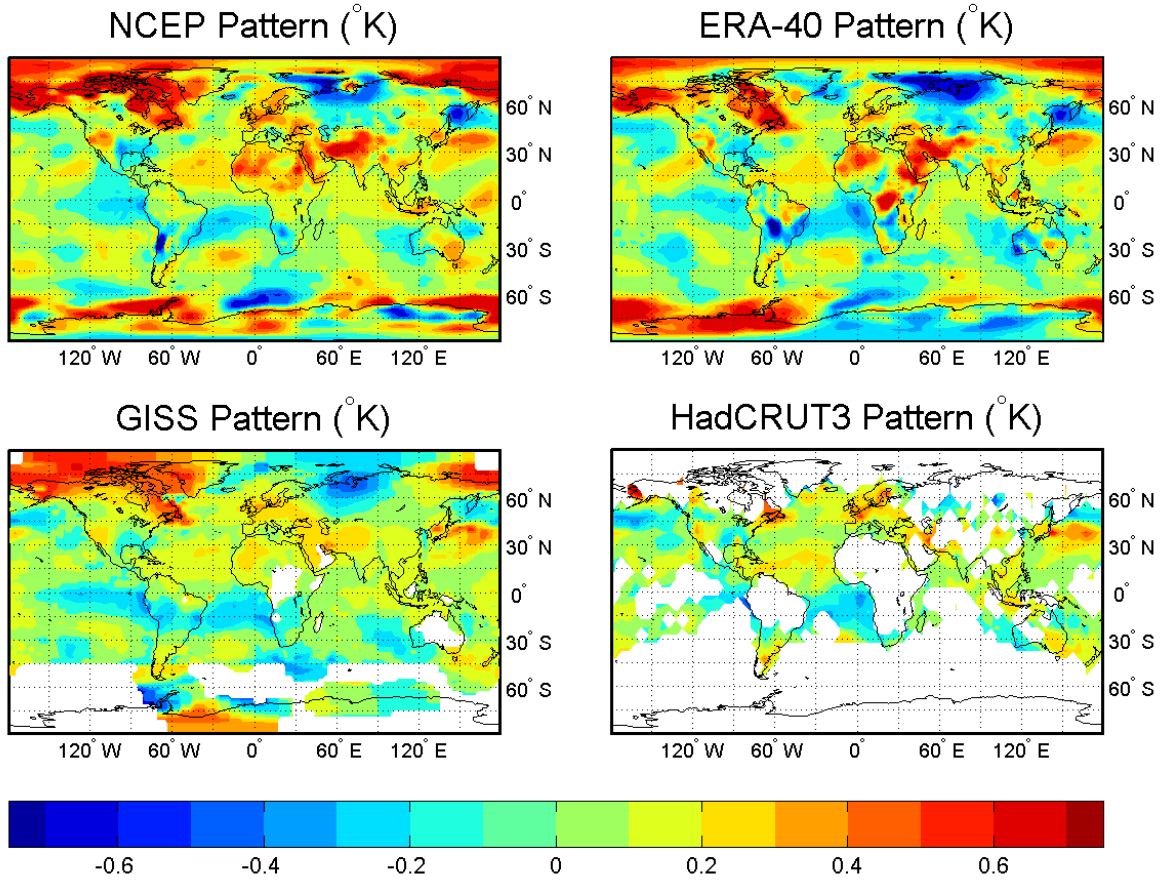


Figure 3.7: Composite mean difference between solar max years and solar min years in surface temperature in K; missing data areas are left blank. The annual average is the average of all four seasons, provided that at least three seasons are available and the missing season is not winter or summer. A seasonal average is the average of three months in the season, provided that at least two months are available.

### *3.4.2 Linear Discriminant Analysis (LDA)*

The method of LDA ([127, 143], see also Chapter 2) finds the spatial weights that best distinguish the solar max group of years from the solar min group of years in surface temperature after detrending and removal of the volcano years as described above. Here we use zonal mean patterns. The zonal mean is taken provided that data are available for at least 35/36 of the longitudes. That criterion needs to be relaxed for HadCRUT3 (to 6/7), because only a few latitudes satisfy the more stringent criterion. There is no useful zonal mean information south of 45°S in any of the *in situ* datasets. Projection of the original temperature data onto these spatial weights yields the time series ($C(+)$) shown in Figure 3.8. The regression of the original temperature data onto these time series yields the zonal mean spatial patterns in Figure 3.9. The zonal mean and 2D (not shown) spatial patterns obtained this way are very close to those obtained by composite mean difference (shown in Figure 3.7), and both are very close to that obtained by regressing temperature data upon the TSI index [84], meaning that these are the robust spatial patterns associated with the TSI variations. All LDA time series are statistically significant at above 95% confidence level as determined by a bootstrap Monte-Carlo test of the separation ratio $R$, which measures the ratio of variances between the solar max and solar min groups to the variances within each group [143]. The correlation coefficient $\rho$ between the signal time series and the TSI index is also statistically significant.

A conservative measure of the amplitude of the response is given by $\kappa$, which is the regression coefficient of the projected time series shown in Figure 3.8 against the TSI time series, also shown. We see that *in situ* data yield a solar cycle signal of $\kappa \sim$0.12 K per 1 $\mathrm{Wm}^{-2}$ variation of solar constant. The amplitude of the solar cycle signal is larger in NCEP (0.17 K) as expected. In subsequent sections we will adopt the range

$$\kappa = \frac{\partial T}{\partial S} \approx 0.12 \sim 0.17 \mathrm{K/Wm}^{-2}. \tag{3.1}$$

The $2\sigma$ regression errors, indicated in the range of $\kappa$ in Figure 3.8, are related to the goodness of fit of temperature response with TSI, and are affected by trend removal and method of analysis; they will not be discussed further here.

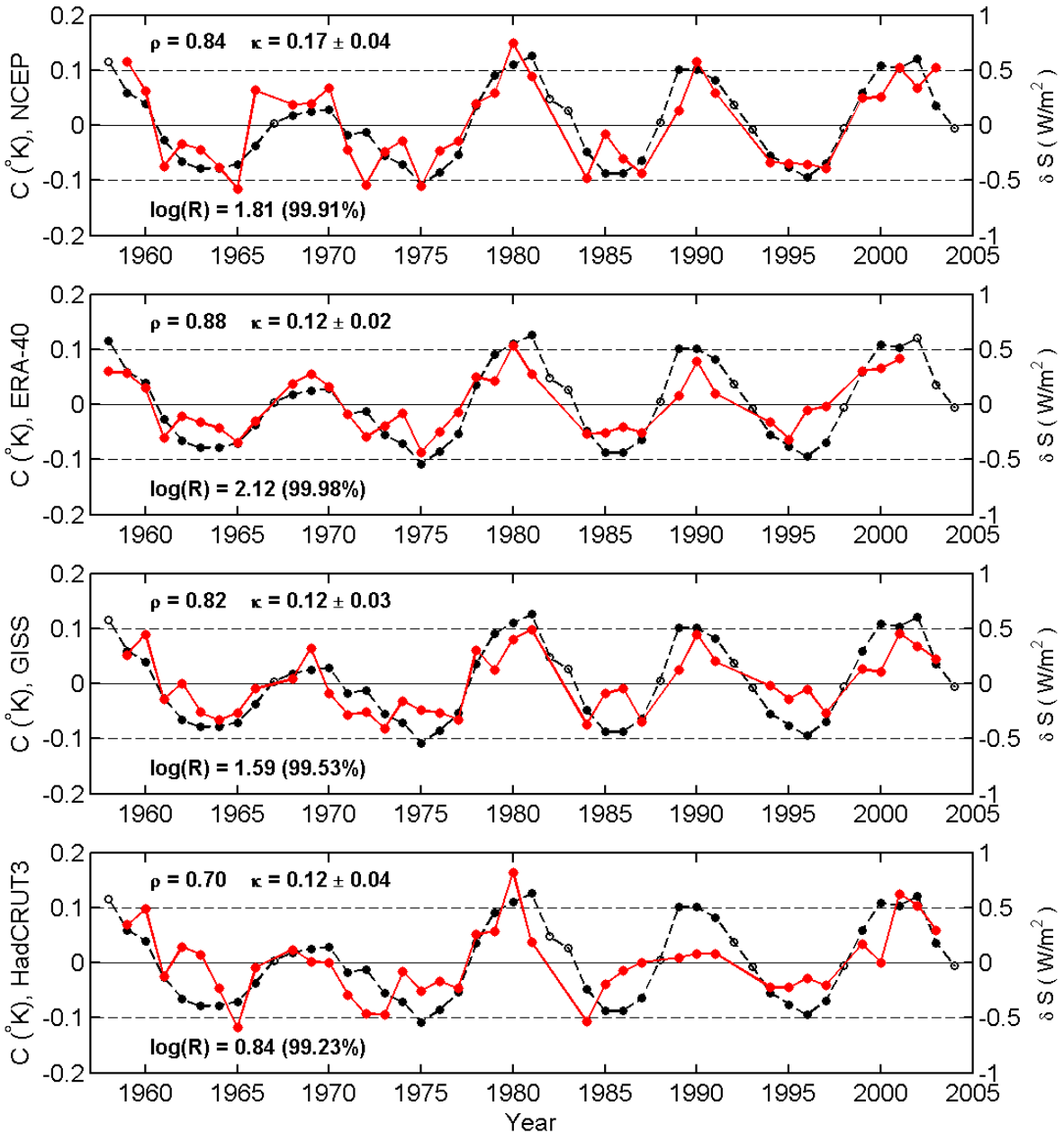


Figure 3.8: Surface temperature time series projected onto the zonal mean spatial weights obtained by LDA, normalized in such a way that the left axis indicates the globally averaged value (using the available data). The dotted line is the TSI index, whose scale is shown on the right axis.

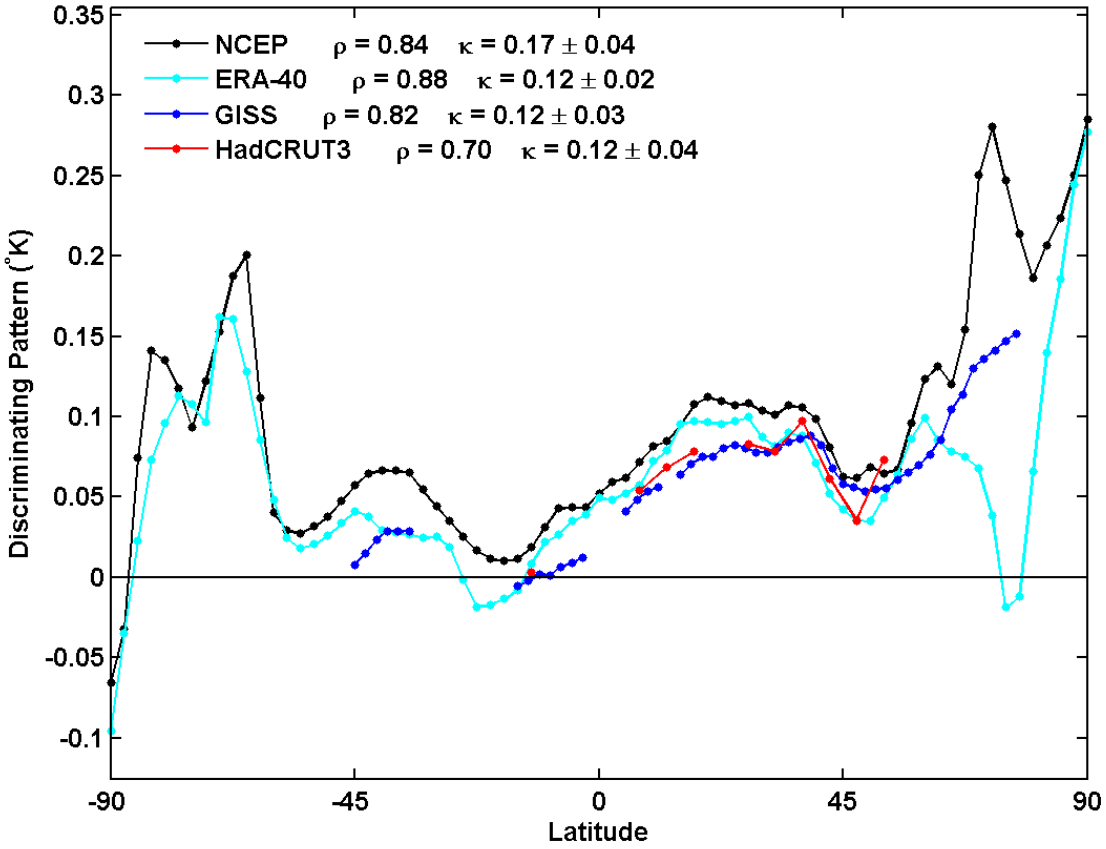


Figure 3.9: Zonal mean meridional profile that best distinguishes solar max years from the solar min years, obtained by LDA. Northern latitudes are given positive values and southern latitudes are given negative values.

### *3.4.3 Climate Sensitivity Parameter*

A measure of climate sensitivity can be defined as the ratio of the global-temperature response to the radiative forcing change,

$$\lambda = \frac{\delta T}{\varepsilon \delta F}, \tag{3.2}$$

where $\delta F$ is the radiative forcing (RF) change for the troposphere, evaluated above the top of the troposphere. This quantity $\lambda$, called the *climate sensitivity parameter*, is expected to be different for different time scales. In order that the definition of the climate sensitivity parameter be more general, and applicable to the greenhouse forcing as well as solar-cycle forcing, the RF change in equation (3.2) is multiplied by the *efficacy* factor $\varepsilon$, which measures the ratio of a unit of RF of, say, the solar-cycle phenomenon, to a unit of RF of $CO_2$ in terms of their effect in causing global warming. In particular, it is meant to take into account the slight difference in spectral distribution of the radiation between the two phenomena (with more bias towards the infrared in the greenhouse RF). The models in AR4 have estimated the efficacy for solar forcing to be close to 1, meaning that it is close to that of GHG forcing once reaching the troposphere, and all models in AR4 fall within the range of 0.7 to 1.0. Thus for solar-cycle forcing and response, we have

$$\lambda_{\text{solar cycle}} = \frac{\delta T}{\varepsilon \delta F} \geq \frac{\delta T}{\delta F} = \frac{4}{1-\alpha}\frac{\delta T}{\delta S}, \tag{3.3}$$

using $\delta F_{\text{solar cycle}} = \delta S(1-\alpha)/4$, where the factor of 4 accounts for the geometry of the circular disk on which the solar constant is measured and the spherical area on which the RF is expressed, and $\alpha \approx 0.3$ is the albedo, the fraction of the radiation reflected back to space by the surface and the clouds. Substituting $\delta T/(\delta S)$ from equation (3.1), equation (3.3) becomes

$$\lambda_{\text{solar cycle}} \geq 0.69 \text{ to } 0.97 \text{ K}/(\text{Wm}^{-2}). \tag{3.4}$$

The definition of RF used by IPCC differs from the usual top of atmosphere value in that the former is evaluated at the tropopause after the stratosphere has adjusted. Absorption of UV radiation by stratospheric ozone reduces the RF reaching the tropopause from the top of the atmosphere. Since 12-15% of the solar variability lies in the UV range (below 295

nm) [88], this reduction can potentially be as large as 12-15%. The stratospheric adjustment involves both the warmer temperature by the enhanced UV heating, which increases the longwave radiation reaching the troposphere, and the enhanced production of stratospheric ozone. There is some uncertainty in the net change of RF caused by the different predicted vertical distributions of enhanced ozone, as reviewed in Table 4.1 of [42]. We take the result from [80], RF$\sim$ 0.18 Wm$^{-2}$, which happens to be the same as our top of atmosphere estimate. Using a smaller RF however does not affect our inequality.

This solar RF turns out to be almost 1/20 that for the total change in RF due to a doubling of $CO_2$ (RF $\approx$ 3.7 Wm$^{-2}$). Therefore the annual rate of increase in radiative forcing of the lower atmosphere during the five years from solar min to solar max happens to be equivalent to that from an average 1% per year increase in greenhouse gases, close to that used in TCR calculations. The global pattern of warming and cooling for the solar cycle signal shown in Figure 3.7 is also quite similar to the IPCC AR4 global warming runs, for example, as given by Leroy et al. [90].

A climate sensitivity parameter for model TCR can be defined as

$$\lambda_{\rm TCR} = \frac{\delta T}{\delta F} = \frac{\rm TCR}{3.7\ {\rm Wm}^{-2}}. \tag{3.5}$$

Since TCR is defined as the $\delta T$ at the time of doubled $CO_2$ after it has been increasing at a compounded rate of 1% per year, we have

$$\lambda_{\rm TCR} > \lambda_{\rm solar\ cycle} \geq 0.69 \text{ to } 0.97\ {\rm K/(Wm}^{-2}). \tag{3.6}$$

Multiplying equation (3.6) by $\delta F = 3.7$ Wm$^{-2}$ we obtain the desired constraint:

$$\rm TCR > 2.5 \text{ to } 3.6\ K. \tag{3.7}$$

The equilibrium climate sensitivity (ECS) should be greater than TCR, by approximately a factor of 3/2 (see Appendix of [143] and [137]). So ECS should be greater than 3.8 to 5.4 K.

The difference in the time scales between an oscillatory forcing and a secular forcing works in the direction of the inequality in equation (3.7). For the TCR, at the time of evaluation, there have been 70 years of compound 1% increase in RF, and the delayed

Table 3.1: Equilibrium Climate Sensitivity and Transient Climate Response for Various Atmosphere-Ocean GCMs, From IPCC AR4

| AOGCM | Equilibrium Climate Sensitivity (°C) | Transient Climate Response (°C) |
|---|---|---|
| 1:BCC-CM1 | n.a. | n.a. |
| 2:BCCR-BCM2.0 | n.a. | n.a. |
| 3:CCSM3 | 2.7 | 1.5 |
| 4:CGCM3.1(T47) | 3.4 | 1.9 |
| 5:CGCM3.1(T63) | 3.4 | n.a. |
| 6:CNRM-CM3 | n.a. | 1.6 |
| 7:CSIRO-MK3.0 | 3.1 | 1.4 |
| 8:ECHAM5/MPI-OM | 3.4 | 2.2 |
| 9:ECHO-G | 3.2 | 1.7 |
| 10:FGOALS-g1.0 | 2.3 | 1.2 |
| 11:GFDL-CM2.0 | 2.9 | 1.6 |
| 12:GFDL-CM2.1 | 3.4 | 1.5 |
| 13:GISS-AOM | n.a. | n.a. |
| 14:GISS-EH | 2.7 | 1.6 |
| 15:GISS-ER | 2.7 | 1.5 |
| 16:INM-CM3.0 | 2.1 | 1.6 |
| 17:IPSL-CM4 | 4.4 | 2.1 |
| 18:MIROC(hires) | 4.3 | 2.6 |
| 19:MIROC(medres) | 4.0 | 2.1 |
| 20:MRI-CGCM2.3.2 | 3.2 | 2.2 |
| 21:PCM | 2.1 | 1.3 |
| 22:ULMO-HadCM3 | 3.3 | 2.0 |
| 23:UKMO-HadGEM1 | 4.4 | 1.9 |

heating due to ocean inertia adds to the instantaneous heating, while for the solar-cycle response at solar max, there have been only five heating years.

The TCRs of 19 coupled atmosphere-ocean GCMs in IPCC AR4 listed in Table (3.1) fall within the rather low range of 1.2-2.2 K with the exception of one, and thus *fail* the lower constraint of 2.5 K determined by ERA-40, GISS and HadCRUT3. The only exception is the Japanese MIROC (hi-res), with a TCR of 2.6 K. All models fail the higher constraint of 3.6 K determined by the NCEP data.

### *3.5 Chapter Summary*

It is often thought that the response to solar cycle is too weak at the earth's surface to be detectable, and that even if a signal is claimed to have been found its statistical significance cannot be established. Using 150 yr of sea surface temperature data from 1854 to 2007 and an objective method, we found a robust signal of warming over solar max and cooling over solar min, with high statistical significance in the time domain. The amplitude of the signal in the SST averaged over the ocean areas between 60°N and 60°S is ∼0.085°C of warming for each $\mathrm{Wm}^{-2}$ of the change in TSI, which is about 70% that found in land-ocean averages (∼0.12°C per $\mathrm{Wm}^{-2}$) found in the recent *in situ* data by Tung et al. [145], as was expected because the response over continents and over the Arctic is known to be larger.

Volcanic eruptions tend to have a significant contribution to the decadal period peak in any spectral analysis; therefore, contamination of the solar cycle signal by volcanic signal has been a longstanding concern. Using 150 yr of data we have now shown that the volcanic contamination is negligible using our method of *pairwise differencing with shift*. This is further confirmed using a new method of *multiple CMD*, similar to the multiple regression method, in which the deterministic volcanic signals are separated out.

Our method of projecting the observed data onto a consistent spatial pattern determined by composite-mean difference of the whole period appears to be effective in reducing contamination by short periods of erroneous data, which do not exhibit consistent spatial patterns. This effect is in contrast to that of the method of (multiple) regression using least squares fit of the time series, which is affected by outliers (which may likely be caused by bad data).

In the method of multiple regression as applied to solar variation by previous authors, an index of solar forcing as a function of time, often in the form of TSI or sunspot number, needs to be prescribed, and the resulting response is assumed to vary in time in exactly the same way as the imposed index, albeit with the possibility of a lag. Our method of CMD projection depends only on the classification of years into the solar max or solar min group and does not require that we know the detailed variation of the total solar irradiation or its long-term trend. In this way we circumvent the controversy concerning the magnitude of the solar forcing trend in these 150 yr. Assuming that multidecadal SST response has the same spatial pattern as the decadal response, we additionally obtain a century-long secular trend; the latter is consistent with Lean's reconstruction of solar forcing. Our result shows that less than a quarter of the observed temperature trend can be attributed to solar forcing.

We have also examined four datasets on global surface temperature, two reanalyses and two *in situ*. The existence of a solar-cycle signal has been established in all four datasets at a confidence level above 95%. The measured solar response is then used to provide a constraint on the transient climate response of models. Since missing coverage in the historical *in situ* station data reduces the amplitude of the response compared to the geographically complete reanalysis data, the smaller amplitude obtained from the *in situ* data is adopted to obtain a lower bound of TCR.

It is seen that most of the current generation of general circulation models assessed by IPCC AR4 have a lower transient climate response than the observed range of sensitivities inferred from the analysis of the solar cycle. This is consistent with the independent finding by Forest et al. [32] that these models simulate too large an ocean heat uptake as compared to observations of ocean temperature changes during the period 1961-2003. See [115] and [101] for different views on how ocean heat uptake affects TCR. This excessive transfer of heat into the oceans tends to reduce the transient climate response for the atmosphere, but does not affect the modeled equilibrium climate sensitivity, which was calculated with a slab ocean in thermal equilibrium with the atmosphere.

Chapter 4

# MECHANISMS FOR PRODUCING THE 11-YEAR SOLAR CYCLE RESPONSE AS REVEALED BY THE TROPOSPHERIC OBSERVATIONS

There is as yet no accepted mechanism for producing the temperature response in the lower atmosphere of the Earth to the 11-year solar cycle forcing, although candidates are numerous (see the reviews by Gray et al. [42], Gray et al. [41] and Haigh [46]). In this chapter we seek tell-tale signs in the observed spatial patterns of the response that may reveal possible mechanisms. Since the response is expected to be small given the small radiative forcing of 0.1% from solar min to solar max, which is embedded in stronger climate noise, many of the features "observed" may not be real (i.e., statistically significant). We therefore pay particular attention not only to establishing the existence of the 11-year solar signal in the response, but also to the statistically significance of the features that we will point to.

## *4.1 Two Candidate Mechanisms*

So far, we have established the existence of the solar cycle signal in temperature at Earth's surface using Reanalysis data [11, 143]) and in *in situ* data [145]) and in 150 years of sea-surface temperature data [169, 144]). To get some hints on the mechanism, we need to examine the latitude-height patterns of the response. The availability of upper air data restricts the earliest global data record we can examine to 1958, the International Geophysical Year, when a coordinated effort in upper air measurements began [65]. (It is understood that truly global coverage did not start until the satellite era in 1979.) The barely 5 cycles of data make the statistical evidence difficult to produce, especially when we take the autocorrelation of the data into account. It appears that it is almost impossible to establish statistical significance in seasonal-mean data over an appreciable region, but we have made some progress with annual-mean data. Although we will be using both forms of data to illustrate the mechanisms involved, this limitation needs to be kept in mind.

Using 32 years of NCEP Reanalysis [69] (1968-1999), Labitzke et al. [76] obtained the annual-mean latitude-height pattern of zonal mean temperature from 1000 to 10 hPa, and from South Pole to North Pole, in a composite mean difference between solar max years and solar min years. An important result, though not emphasized by the authors, is that it is positive everywhere, implying that during solar max, when the solar radiation as measured by the total solar irradiance (TSI) is higher, the atmosphere warms everywhere in the annual mean. (Later Coughlin and Tung [19] used this result to argue that in the annual mean the mechanism must be radiative, and not due to an overturning circulation, which would have created negative and positive regions.) No statistical test was attempted by Labitzke [75]. Haigh [45] used multiple regression analysis and obtained a latitude-height pattern as the regression coefficient of the TSI. Statistical test was performed using the Student's $t$-test after pre-whitening the data. Most of the signal in the latitude-height plane was found not to be statistically significant with the exception of the following areas: a horizontal region of warming at the tropical tropopause, and two vertical strips of warming at midlatitudes from the surface to the tropopause, thus forming a "horse shoe" pattern of warming. It should be noted that this "horse shoe" is only a pattern where statistical significance was established, and does not necessarily represent the physical shape of the response. The cause for this pattern has been attributed to a shift in the Hadley circulation, which weakens and broadens during the solar max [44]. Haigh et al. [47] further suggested that such shifts in the tropical circulation can be brought about by an increase in tropical temperature in the lower stratosphere caused by enhanced ozone absorption of solar ultraviolet radiation in the solar max relative to the solar min. This is often termed the "top-down" mechanism.

On the other hand, it is now known, from numerous model studies of the greenhouse gas warming problem, that a robust response to radiative heating near the surface also takes the form of an expanded and weakened Hadley circulation and a poleward shift of the tropospheric jets and the accompanying storm tracks (see [59] for a summary). It is also known that such patterns can also be produced with a uniform change in the total solar irradiance (TSI) without needing the enhanced UV change in the "top-down" mechanism [95, 94, 50]. In fact, a 2% change in TSI (20 times the observed variation from solar min to solar max) produces similar amplitude and spatial pattern in the troposphere as a $2\times CO_2$

radiative forcing.

Recently, Cai and Tung [8] explained how a "bottom-up" mechanism, involving only the troposphere, can lead to the response mentioned above: The bulk of the solar forcing, in absolute terms, is in the visible range, which penetrates to the surface of the Earth, and has the largest amplitude in the tropics. Here most of the absorbed energy goes into evaporating water. The "evaporative feedback" reduces the surface warming. Vertical convection in the tropics deposits this energy in the form of latent heat in the upper troposphere, near the tropopause, which defines the limit of vertical convection. It is here tha the warming occurs, rather than in the stratosphere (as would be in the case of ozone heating). The meridional temperature gradient in the upper troposphere (with temperature decreasing poleward) is further strengthened by water vapor feedback (with water vapor warming more strongly in the tropics than in the extratropics) and leads to enhanced large-scale poleward heat transport. The latter produces a poleward shift of the tropospheric jet. A warmed pole aloft then radiatively warms the polar surface via thermal downwelling; this is permitted (i.e. without causing vertical convection to redistribute the heat) by the very large static stability of the cold polar surface. The resulting pattern at the surface is one with polar amplification of warming. The surface meridional temperature gradient is such that the largest warming is found at high latitudes, where the solar forcing is the smallest, and the minimum warming is over the ocean in the tropics (because of evaporative feedback and poleward heat transport), just the opposite of what one would expect from the meridional shape of solar forcing.

It therefore appears that either a "top-down" or a "bottom-up" mechanism can produce the observed "horse shoe" pattern of solar cycle warming in the latitude-height plane. The mechanism of [8] yields additional predicted patterns that can be tested using observation. It is hoped that, with the longer observational record that is now available, we will be able to find more features that are statistically significant. This may then lead us to deduce which, or both, of the mechanisms proposed is dominant in producing the observed response.

## *4.2 Evidence of A "Bottom-Up" Mechanism*

### *4.2.1 Data and methodology*

We analyse the monthly mean air temperature data from the National Centers for Environmental Prediction-National Center for Atmospheric Research (NCEP-NCAR) reanalysis [69] over the time period 1958-2011. The air temperatures are available with global coverage at 17 atmospheric pressure levels from 1000 hPa to 10 hPa. We could not reproduce the spatial patterns and statistical confidence levels using a second reanalysis dataset, ERA-40 [146]. The problem was traced to differences in the data which was found to be in error that was subsequently fixed with ERA-Interim [23]. However, the new data record is too short (only since 1979) for statistical analysis, but has been used by us to verify the spatial patterns obtained using NCEP in the overlapping period.

We will be performing two different kinds of statistical analysis. First, we establish the existence of an 11-year solar cycle signal in the temperature at each pressure level using the CMD-P method. First the composite-mean difference (CMD) is calculated on the zonally averaged annual mean (for the months of December-November) and/or seasonal mean (for the months defining the season) temperature data after the trend is removed. The detrended data are then projected onto the CMD spatial pattern to yield a time series. This time series is correlated with the TSI index to yield a correlation coefficient $\rho$, and a bootstrap Monte-Carlo test is employed to estimate the likelihood that the observed correlation could be obtained by chance if there were no solar cycle signal in the temperatures ([11], see alsp Chapter 2). Due to the fact that the temperature data may be serially correlated, the *moving-block bootstrap* method [29, 79, 89, 157], which resamples blocks of consecutive data points with replacement from the observed time series, is used to reduce the risk of overestimating the confidence levels in our results when performing the Monte-Carlo test. Selecting an appropriate block length is not an easy task [89]. Ideally long blocks are desired to account for as much autocorrelation as possible. However with only 50 years of data the randomness of the synthetic data may be compromised if the blocks are too long. Our test of autocorrelation (see later) suggests that a block length of 7 years is sufficient.

The solar variability index used in our analysis is the monthly TSI reconstruction up to

2009 based on [87] and [153] and kindly provided to us by Dr. J. Lean, and the time series is extended to 2011 using the daily measurements from the Total Irradiance Monitor (TIM) on NASA's Solar Radiation and Climate Experiment (SORCE) available from `http://lasp.colorado.edu/sorce/data/tsi_data.htm` [74]. The years of 1958-2011 are objectively categorized into two groups between which the CMD results are calculated: the solar-max (-min) years in which the TSI values are 0.06 $\mathrm{Wm}^{-2}$ above (below) the mean value for the entire period. Two years immediately after each of the two large volcanic eruptions, El Chichón in March 1982 and Pinatubo in June 1991, are removed from our analysis to reduce the possible contamination of the solar signal.

The TSI time series has no trend in the period of 1958-2011, whereas non-uniform trends are present in the NCEP temperature records we are analyzing here. There are also systematic differences between the measured temperature data before and after 1979, when satellite measurements started to be incorporated in the reananlysis, which introduced a discontinuity at some locations (see Figure 4.1). Therefore the detrending process needs to be more complicated than simply removing a linear trend from the temperature time series. Low-order polynomial trends have the same drawbacks while high order polynomials could be too oscillatory and also pose a danger of over-fitting the trend. In this study we remove an adaptive trend obtained using the method of *locally weighted scatterplot smoothing* (LOWESS) [15, 16] from each zonally averaged monthly time series. The LOWESS technique defines the trend at each point in the data set individually by fitting a weighted linear least squares regression model to a localized subset of the data, giving more weight to points near the point whose response is being estimated and less weight to points further away. The weight function that we used for LOWESS is the traditional *tri-cube* weight function. And each subset to which we fit a local linear trend is selected to have a length of 20 years, which comprises approximately two complete solar cycles so that the fitted trend should contain minimal solar signal. Figure 4.1 shows some examples of the LOWESS trend fitted to the monthly temperature time series at several spatial locations. This trend removal is similar to using the 11-year *running mean* as the trend, except that the half cycles at the beginning and end of the time series are not sacrificed.

The bootstrap Monte-Carlo $\rho$-test mentioned above technically proves the existence of

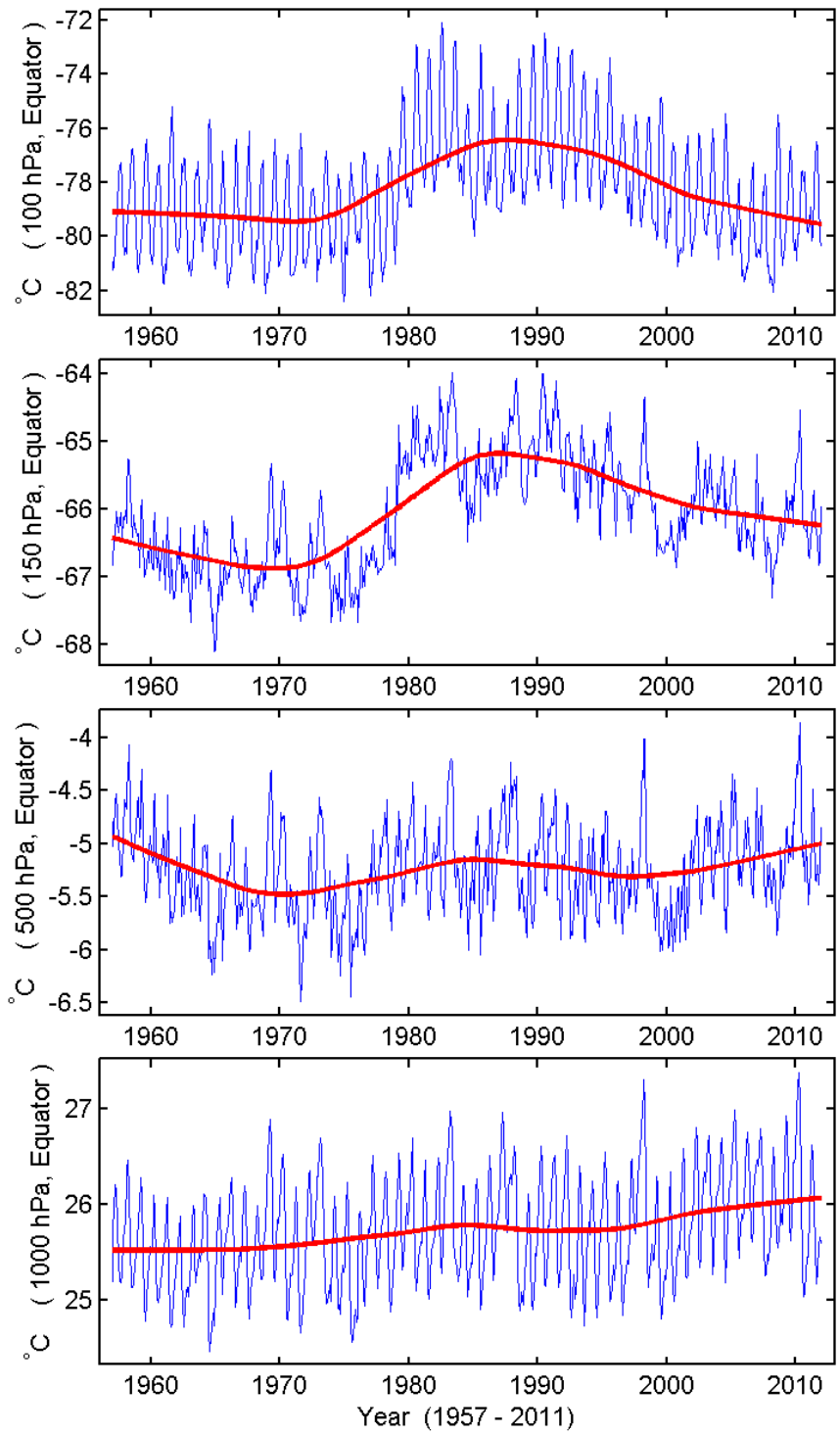


Figure 4.1: Raw data (in blue) at four selected latitude-height locations which are indicated along the left edge of the figure, and their smoothed trends (in red) defined by the LOWESS technique. The top two panels show clearly discontinuities around 1979 which are likely due to the inclusion of measurements from satellites.

the solar signal at a particular pressure level in the atmosphere. We are also interested in knowing whether the CMD-derived spatial characteristics of the solar cycle response are distinguishable from random noise at various latitudinal and vertical locations. Conventionally a two-sample Student's $t$-test (also known as Welch's $t$-test if the variances of the two populations are assumed to be unequal) is designed to answer such a question. However, the presence of serial correlations in the temperature data violates the fundamental assumption of the two-sample $t$-test. Here we perform a modified Student's $t$-test on our CMD results with autocorrelations being taken into account by assuming a regression-like model for the time series at each spatial grid point. Details about how to conduct the new Student's $t$-test are described in Section 3 of Chapter 2.

### *4.2.2 Results*

Figure 4.2 show the CMD spatial pattern at several pressure levels (left panels) along with the time series obtained by projecting the annual mean temperature data onto this pattern (middle panels). The number in the right panels indicates the statistical confidence level obtained using the Monte Carlo test described in the previous section. A number 95% or higher establishes the existence of the solar cycle signal at 95% confidence level. At 1000 hPa, the result here is consistent with surface warming obtained previously by Camp and Tung [11] (The negative value over the Antarctic continent below 700 hPa should be ignored, as that pressure level is probably underground), showing polar amplification of warming and tropical minimum of warming. The results above 1000 hPa are new. The response generally follows the TSI index, but the consistency of the two time series degrades as one goes up in height during the first cycle of the record (in the 1960s). The upper level data are probably not of good quality in early decades.

Figure 4.3 shows the latitude-height distribution of the CMD pattern. A second type of statistical test, the adaptive AR($p$)-$t$ test described in the previous section and Chapter 2, is performed and the result is expressed in contour form and superimposed on top of the color scale representing solar cycle temperature amplitude difference between the solar max and solar min groups. The dot-dash (solid) contours enclose regions where the amplitude of

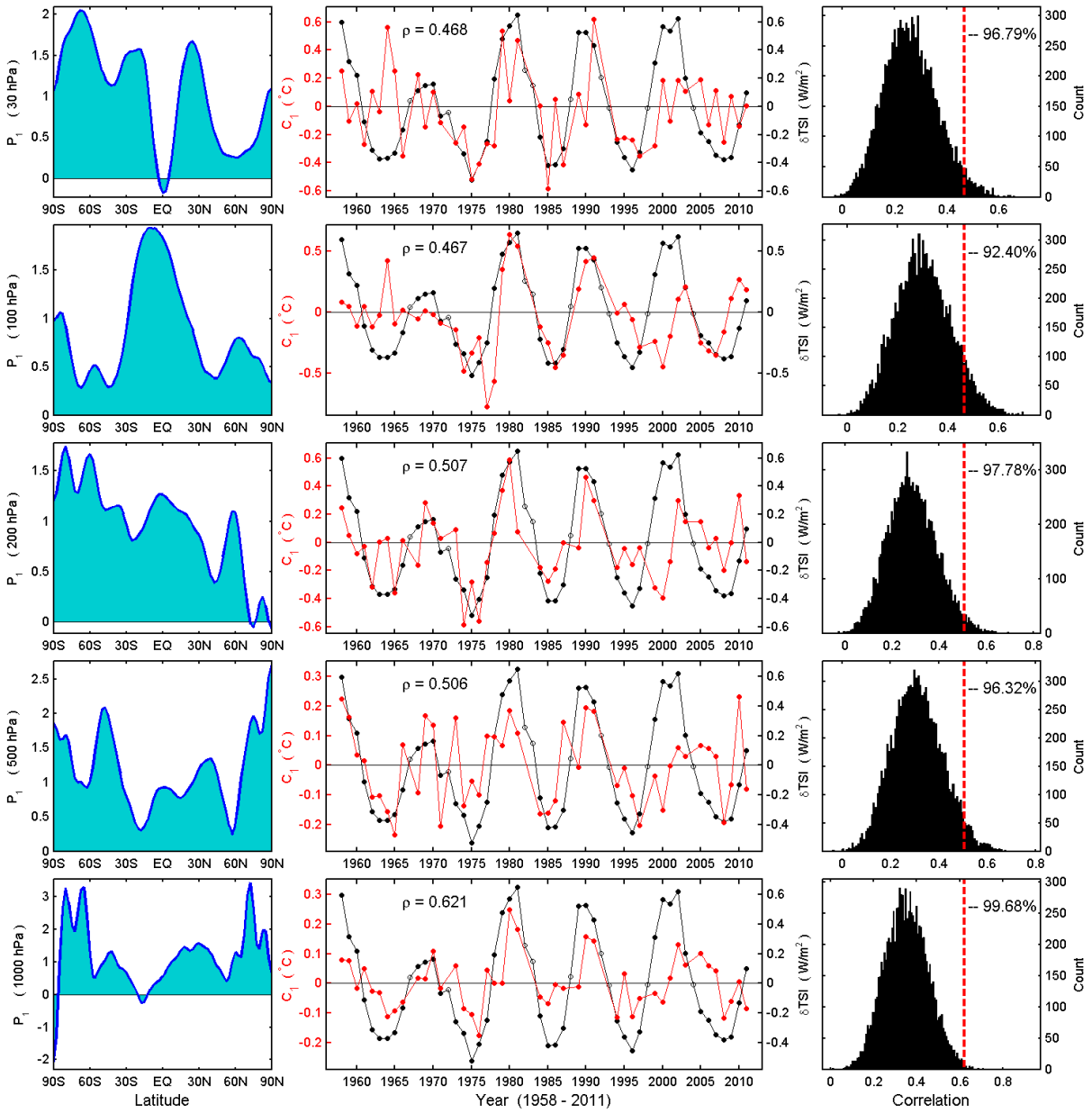


Figure 4.2: Composite-Mean Difference patterns as a function of latitude of the annual mean zonal mean NCEP temperature (left panels), and the time series obtained by projecting the temperature data onto this pattern (middle panels) with the TSI index superimposed. Each of the spatial patterns $P_1$ is normalized to have unit global mean and the projected time series $C_1$ is also normalized accordingly. The statistical confidence level using the Monte Carlo test is indicated for each pressure level in the right panels, along with the null distribution.

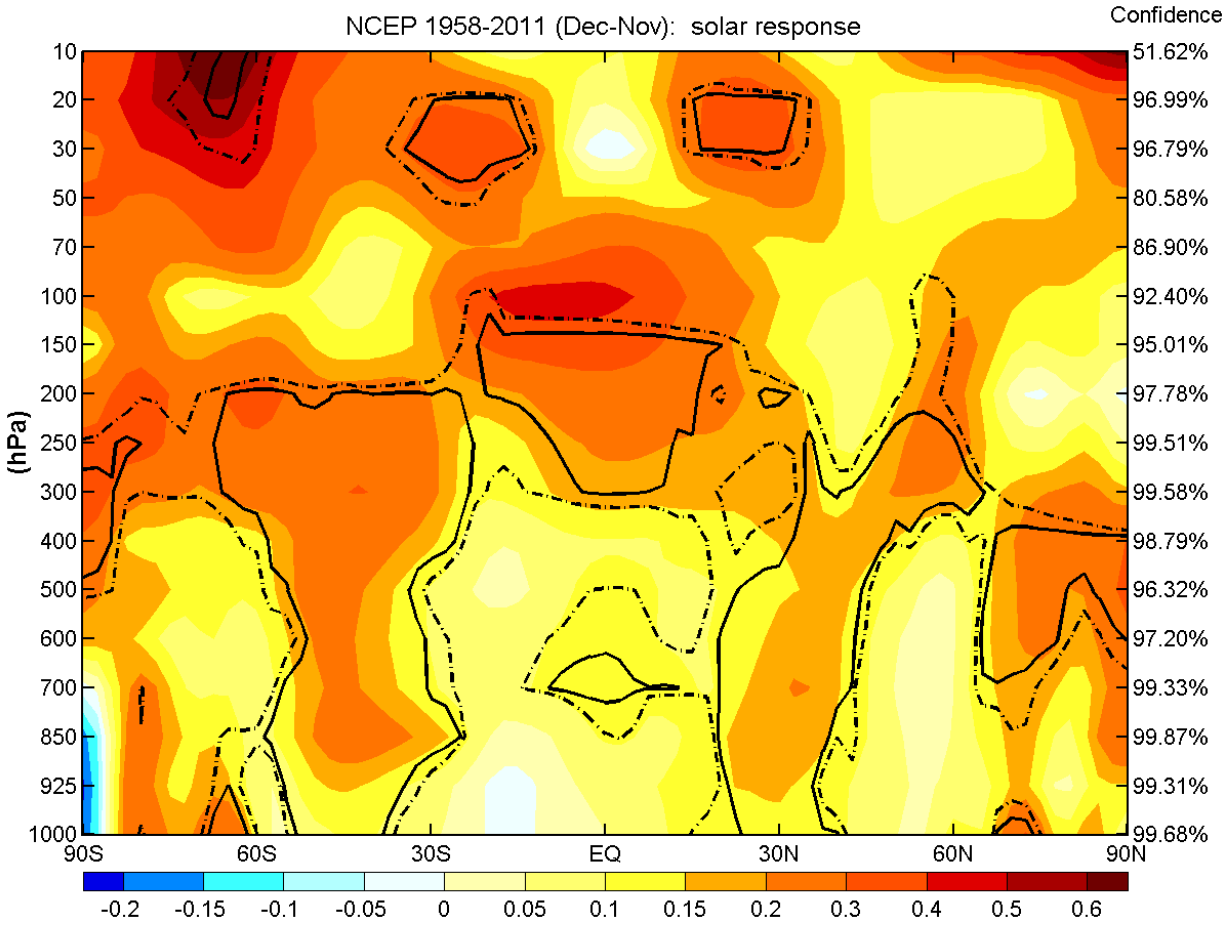


Figure 4.3: Composite-Mean Difference pattern between the solar max and solar min groups in annual mean zonal mean NCEP temperature. Color scale in degrees Celsius. 95% (90%) confidence level in dot-dashed contour and 97.5% (95%) confidence level in solid contour based on a modified Student's $t$-test based on a one-sided (two-sided) test. The existence of the 11-year solar cycle signal is established level by level, with confidence level listed on the right border of the figure.

the response is statistically significant at 95% (97.5%) confidence level based on a one-sided test (since we know that solar max warms). They become 90% (95%) respectively if we do not know a prior that it is warming and need to use a two-sided test. The regions that pass the statistical test at 95% confidence level are larger than in previous work because of the longer data record now available, and also because of the better trend removal process adopted. The Monte Carlo test results are appended to the right. Note that this test is for the whole level, and does not pick up particular features on a level. For that we rely on the contour confidence levels.

*Troposphere*

The existence of the solar cycle signal is established for the first time at 95% confidence for the entire troposphere in the annual mean (see confidence level on the right border of Figure 4.3). The height of the tropopause in the tropics is around 70-100 hPa. There is a warming region *below* the tropopause over the tropics. The part that is statistically significant lies is centered around 200 hPa. This feature is present in all four seasons. If this feature were due to ozone heating the warming would be centered in the stratosphere. This same warming center is found in the model result of [8], suggesting that this is caused by latent heating brought up from the surface by deep cumulus convection. At the surface, evaporative feedback should reduce the warming there, and indeed we find the minimum heating over the tropics at the earth's surface. At the 200 hPa level and below, there is a negative poleward temperature gradient, and so poleward and downward heat transport are possible, and the spatial patterns appear to show a poleward and downward pattern of warming as one goes from the equator to the pole. There are two vertical strips of warming outside the edge of the tropics. These are robust and statistically significant, as first shown by Haigh [45]. We interpret them as a consequence of the large-scale heat transport in the troposphere leading to the poleward shift of the Hadley circulation.

There is no indication of a solar peak - cold (La Niña) signature association at the equatorial Pacific, as proposed by van Loon et al. [148]; Meehl et al. [100]. Haam and Tung [43] recently found such association to be not statistically significant. See also [120, 169]

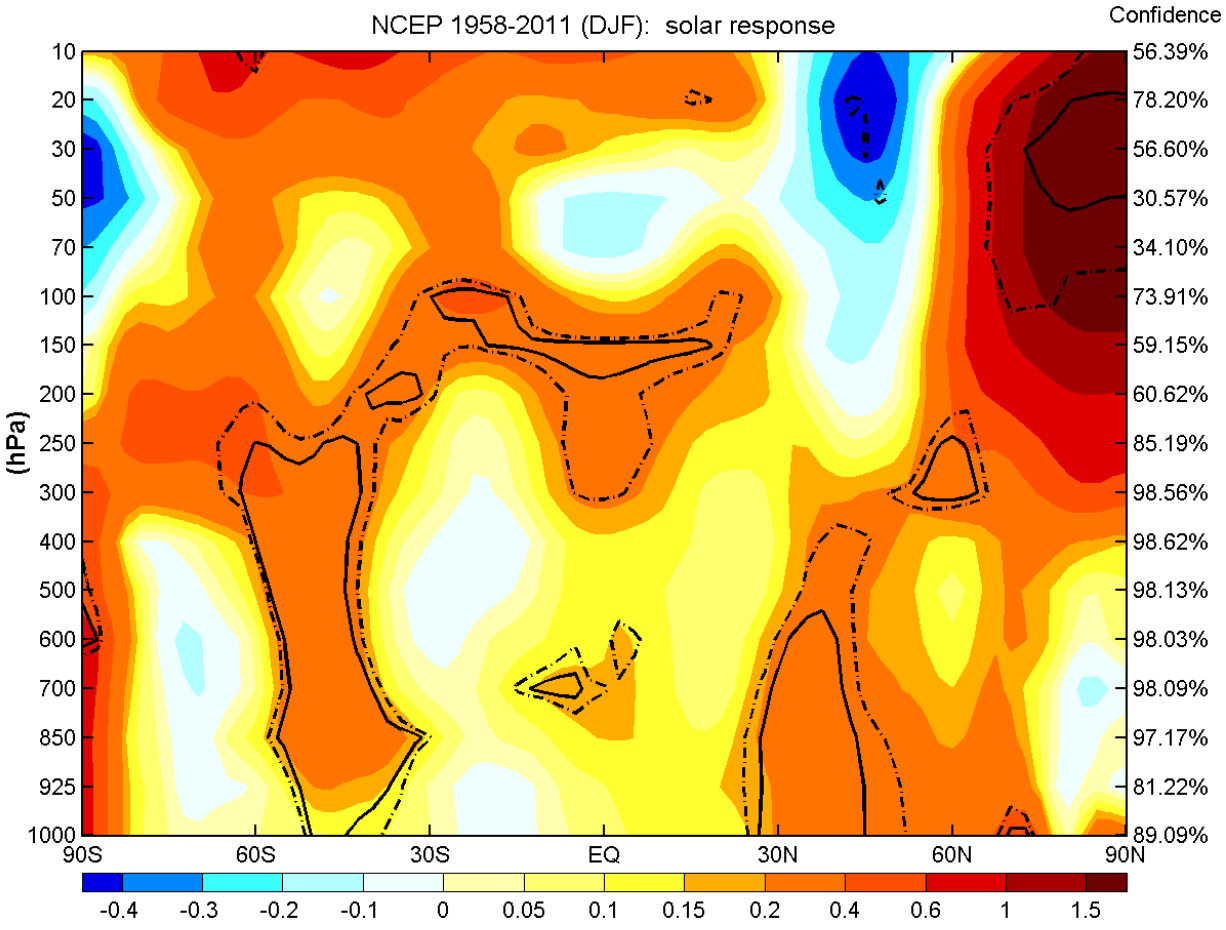


Figure 4.4: Same as Figure 4.3 except for boreal winter (DJF).

*Stratosphere*

The statistical significance of the solar cycle signal in the stratosphere is difficult to establish. The largest warming occurs over the polar stratosphere in the winter. NCEP data extends only to 10 hPa and catches only the lower portion of the warming hemisphere. Labitzke and van Loon [77] (and many of their papers on the subject) showed that the solar cycle signal over the polar region becomes apparent only when the data are segregated according to the phase of the equatorial Quasi-Biennial Oscillation (QBO). Since both solar max and easterly QBO can trigger a stratospheric Sudden Warming phenomenon [103], the solar signal is statistically significant only during the westerly phase of the QBO. The difference between solar max and solar min during the easterly phase of the QBO is small and not statistically significant [9]. The small difference can be of either sign; previously too much meaning was attached by some authors to a small cooling during some periods. There is not enough evidence to conclude that solar max warming reverses to cooling during the easterly phase of the QBO.

We are able to obtain a statistically significant solar cycle polar stratospheric warming during boreal winter (DJF) using the CMD-P method without having to segregate the data into QBO phases (Figure 4.4). The amplitude of the warming is however smaller than when the data are segregated, due to QBO contamination (e.g. warming due to easterly QBO during solar min, making the difference between solar max and solar min smaller). This result is consistent with the conclusion of [9] that solar max warms the polar stratosphere during winter, and that there is generally no reversal of the solar warming. The pattern of polar warming and cooling south of 60°N is suggestive of the meridonal pattern of the Sudden Warming phenomenon. The signal in austral winter (SON) stratosphere (see Figure 4.5) is not statistically significant, possibly due to poorer data quality, although the pattern of polar warming is consistent.

The two warming centers near 30 hPa on both sides of the equator (between 20-30 degrees of latitude) are statistically significant in the annual mean (Figure 4.3), but they do not show up clearly in winter and summer. Nor are they occurring in the same region in other seasons. They show up most clearly in July-August (Figure 4.6)) as the downward

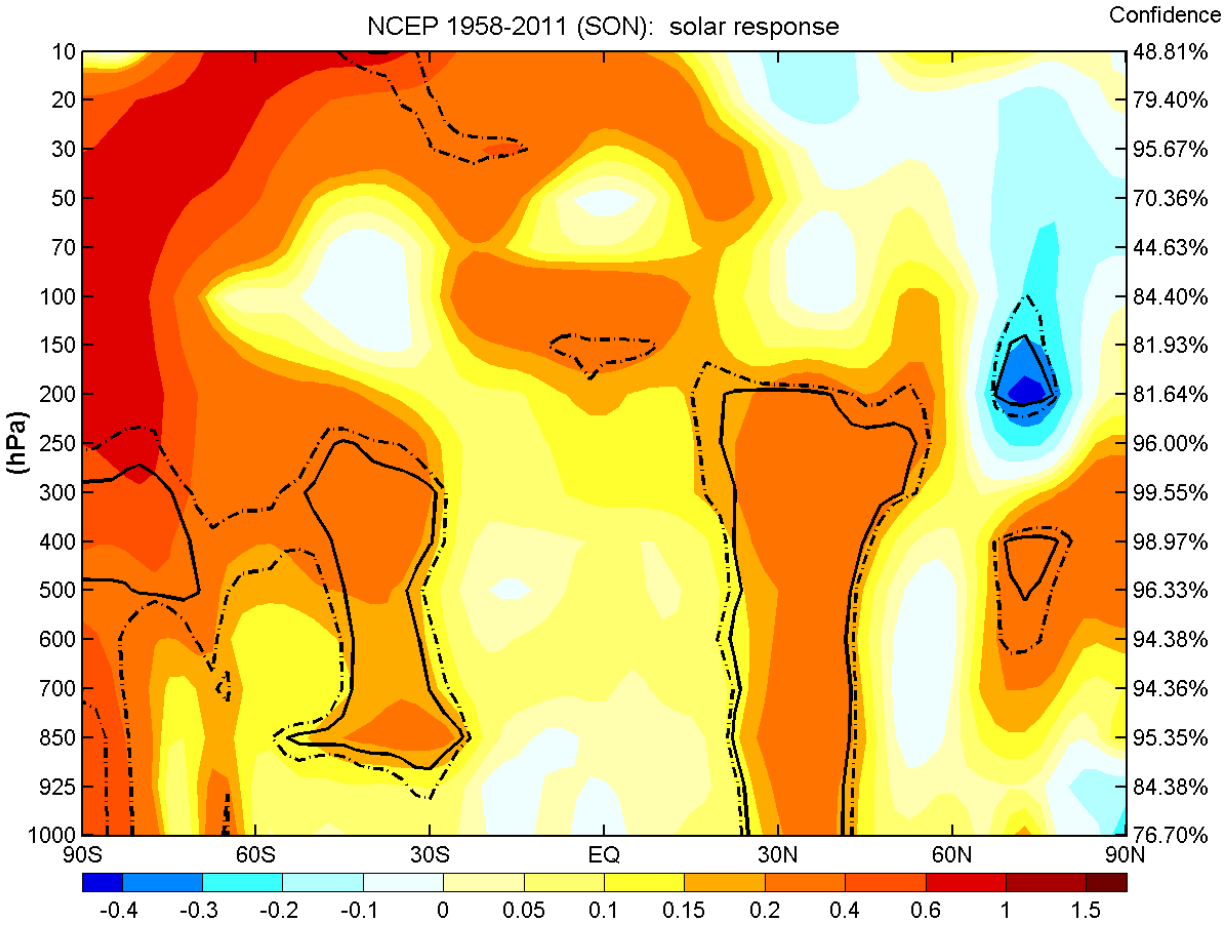


Figure 4.5: Same as Figure 4.3 except for austral winter (SON).

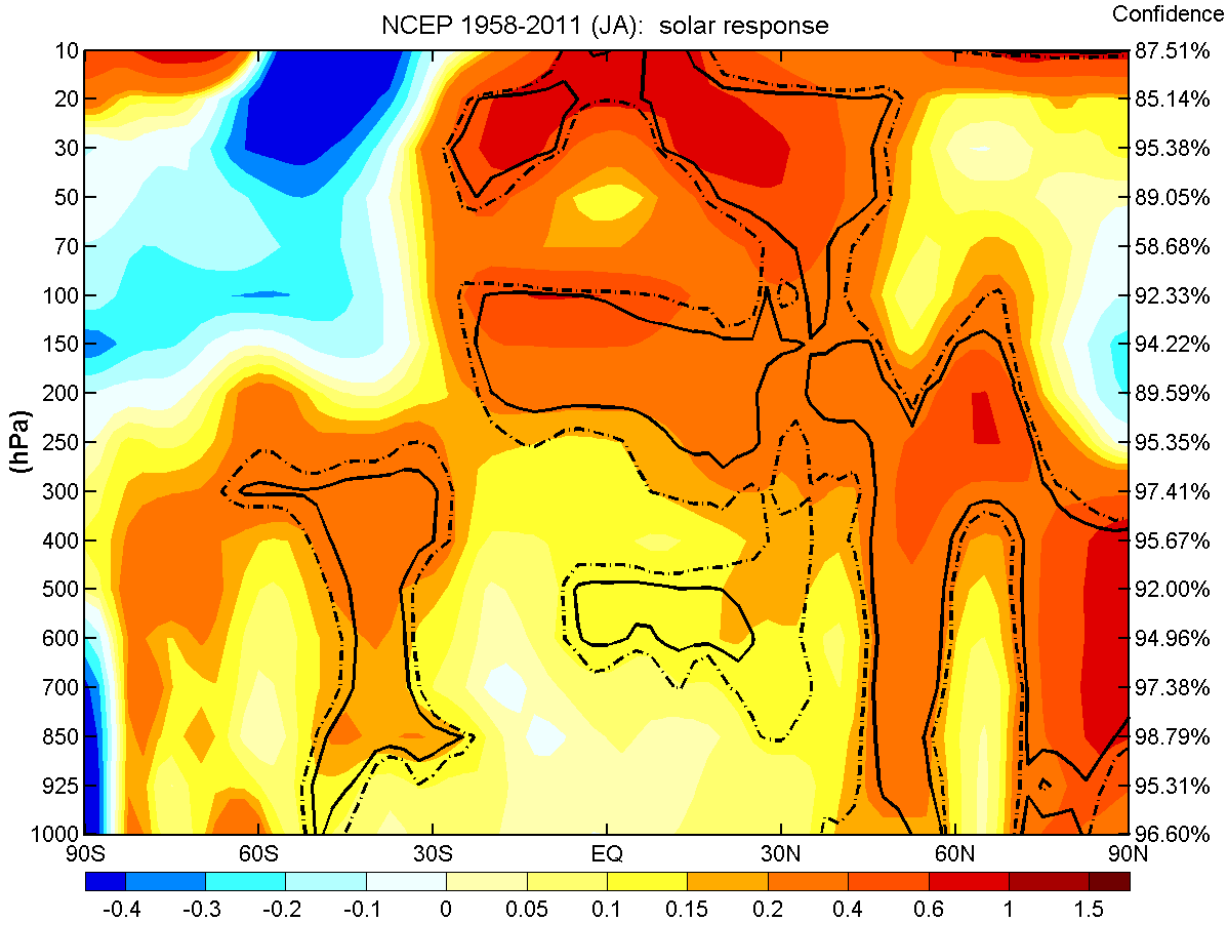


Figure 4.6: Same as Figure 4.3 except for July-August average.

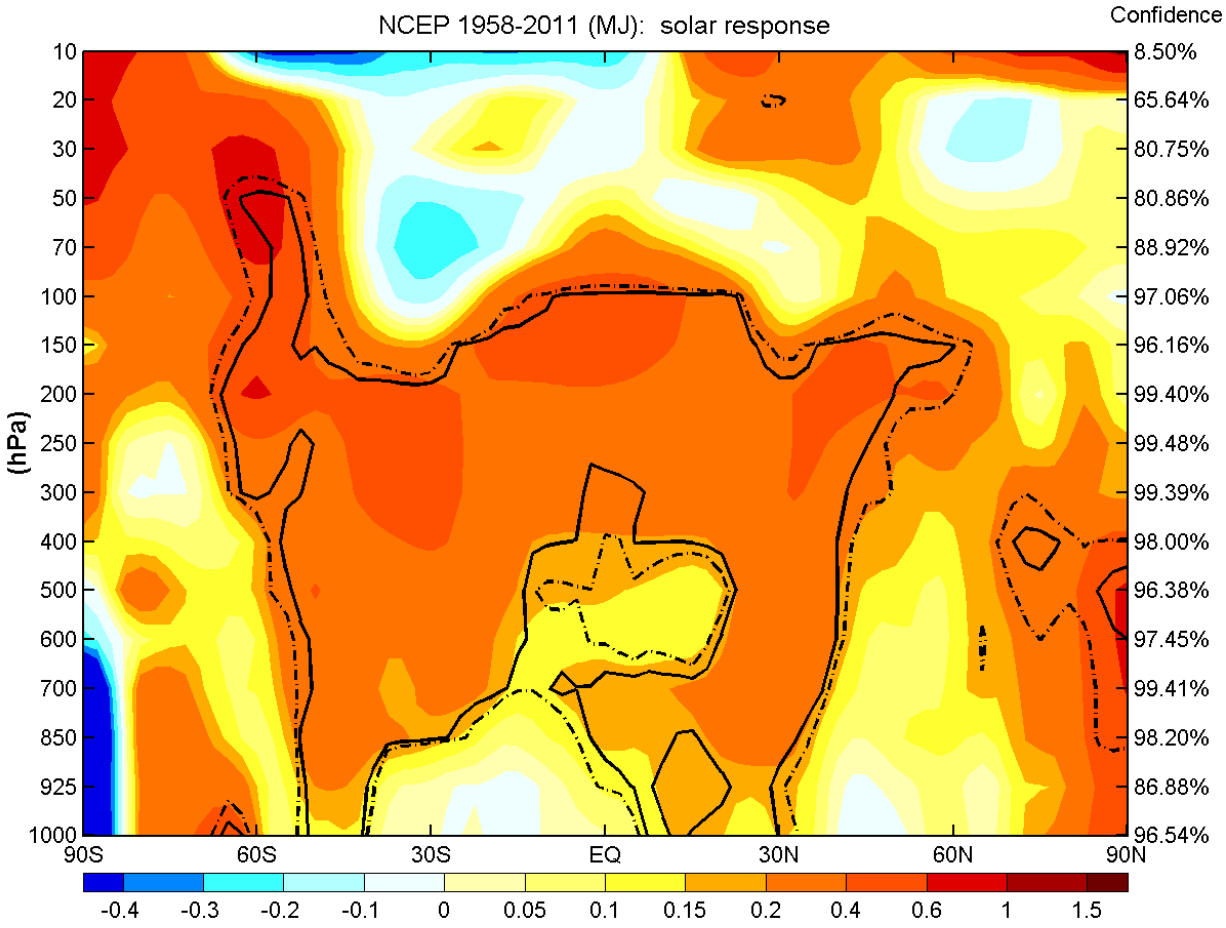


Figure 4.7: Same as Figure 4.3 except for May-June average.

legs of a heating centered at the equator with a maximum higher up (10 hPa and higher). This is likely caused by ozone heating (absorption of enhanced solar cycle UV radiation). However, because the heating is located so high above the tropopause and because it does not occur during all seasons – for example, there is cooling in the lower stratosphere in the two months prior (Figure 4.7) – it is not likely that they are the cause of the two vertical strips of warming in the troposphere, which are evident in all seasons.

### *4.3 Chapter Summary*

We have established statistically the existence of the 11-year solar cycle signal in temperature throughout the troposphere. There is robust heating centered over the tropics *below the tropopause* in all seasons, which cannot be interpreted as heating due to ozone absorption of solar UV radiation, since the tropospheric ozone concentration is extremely small. This heating is situated above a minimum in warming near the surface over the tropical oceans, suggestive of the role of deep cumulus convection caused by surface heating and evaporative feedback. There are also two vertical strips of warming outside the edge of the tropics in the troposphere that could be a result of poleward shift of an expanded and weakened Hadley circulation. The warming center in the stratosphere that can possibly be attributed to ozone heating occurs too high up (above 30 hPa) and it is too weak (less than 1 K) to be the likely cause of poleward shift in the tropospheric Hadley circulation. The evidence we present here is suggestive of a "bottom-up" mechanism similar to that for greenhouse warming, as discussed in [8].

Chapter 5

# THE PACIFIC'S RESPONSE TO SURFACE HEATING



Previously the pattern of the temperature response to increased radiative heating due to the buildup of greenhouse gases in the eastern equatorial Pacific Ocean has been reported in the forms of both a warm tongue and a cold tongue by different researchers, with corresponding supporting physical explanations. These contrasting points of view need to be reconciled. On the other hand, it has been difficult to unambiguously identify this response using observations due to the strong interference of the dominant pattern of natural variability in that region. Here we attempt to address the question of whether the enhanced anthropogenic or solar forcing produces a warmer or colder sea surface temperature in the tropical Pacific using a modified method of multiple linear regression.

## *5.1 Two Competing Theories*

The El Niño Southern Oscillation (ENSO) is a dominant mode of natural oscillation of the equatorial coupled atmosphere-ocean system in the Pacific Ocean. The question of whether the equatorial Pacific responds to the fluctuations in radiative heating in a La Niña-like (cold ENSO) pattern or an El Niño-like (warm ENSO) pattern is under debate in the context of global warming (see [149]). It has been argued that because of the tight coupling of the atmospheric Walker circulation with the thermocline depth in the eastern

equatorial Pacific Ocean, a larger radiative heating may not necessarily result in warmer sea surface temperature (SST). There are currently two competing theories, differing in the degree to which the atmosphere is coupled to the ocean. Clement et al. [14] presumed that the eastern Pacific SST is controlled by ocean cold-water upwelling, and therefore a basinwide heating increases only the SST in the western Pacific. The resulting east-west temperature gradient strengthens the atmospheric Walker circulation, whose easterly flow near the surface induces stronger ocean upwelling in the eastern Pacific, thus a cold-ENSO-like response. On the other hand, Held and Soden [59] and Vecchi and Soden [150] suggested that tropical circulations, especially zonal overturning circulations (such as the Walker circulation) would weaken in a warmer climate. The weakened surface easterlies lead to an El Niño (warm ENSO)-like SST, of a warm tongue in the eastern Pacific. Held and Soden [59] pointed out that this is a robust response of the current crop of coupled atmosphere-ocean general circulation models: As the SST warms, convection actually decreases, because the lower-tropospheric water vapor increases faster than the global mean precipitation. Xie et al. [167] suggested that the warming pattern should be less El Niño-like because of the strengthened southeasterlies south of the equator, which are due to hemispheric asymmetry in land-sea area [91]. Observational evidence is ambiguous. Vecchi et al. [149] showed that the trend in 1880-2005 has a cold-ENSO-like pattern in one dataset [The Met Office Hadley Centre Sea Ice and SST dataset (HadISST)] but a warm-ENSO-like pattern with asymmetry in another [the Extended Reconstructed SST dataset (ERSST)]. Karnauskas et al. [70] found that the zonal SST gradient strengthened in boreal autumn but weakened in spring. We hope to reconcile these disparate results in this chapter.

A related phenomenon is that of the 11-yr solar cycle. Does the tropical Pacific respond with a La Niña-like or an El Niño-like pattern during solar maximum, when the solar radiation is about 0.1% stronger than during solar minimum? Because the time scales for tropical convection and for ENSO responses are much shorter than both 11 yr and the multidecadal scale of greenhouse gas increases, these mechanisms in the equatorial Pacific should be equally applicable to the two phenomena. A recent series of papers by Meehl and Arblaster [99], Meehl et al. [100], van Loon and Meehl [147], and van Loon et al. [148] shows that the equatorial Pacific responds in a prominent symmetric cold-ENSO-like pattern

during the northern winter season of the peak solar year, as compared with climatology. The mechanism proposed by the authors is a variant of the "ocean thermostat" mechanism of [14], with an additional detail of positive cloud feedback: A basinwide radiative heating will preferentially heat the eastern Pacific, which is more cloud-free because of the colder SST. The authors suggested that the increased evaporation does not locally form clouds but is instead transported by the surface easterlies to the western Pacific and that the Walker circulation is strengthened instead of weakened, keeping the eastern Pacific cloud-free. Meehl et al. [100] proposed this as an amplifying mechanism for the response to solar forcing. The referenced work of [39] appears to support a strengthened Walker circulation. However, Coughlin and Tung [20] have showed that the results of [39] were problematic.

The van Loon-Meehl "solar response" pattern is not a response to solar forcing, because the level of total solar irradiance (TSI) for the 11 sunspot peaks used, relative to the 1968-96 "climatology" that was subtracted from it, is almost zero [the January-February (JF) mean difference is $-0.021$ Wm$^{-2}$, as compared with a typical value of 1 Wm$^{-2}$ variation from solar maximum to solar minimum]. Our second argument is that their response, if it were solar related, would have been opposite in peak sunspot minimum years compared to the peak maximum years. Yet they did not find a warm-ENSO-like pattern in the former, while a cold-ENSO-like pattern was found in the latter. Of the 11 sunspot peak years studied by van Loon et al. [148], eight (1883, 1893, 1917, 1937, 1957, 1968, 1979, and 1989) are cold-ENSO years, and three (1905, 1928, and 1947) belong to warm ENSO [with the classification defined by the cold tongue index (CTI) in January-February mean]. Van Loon and Meehl [147] mentioned 1989 as the only cold event, but 1893, not mentioned, was even colder. Van Loon and Meehl [147] later added three more solar peak years (1860, 1870, and 2000) to the time series, all of them belonging to cold ENSO. It is therefore not surprising that their "solar peak" patterns take the beautiful form of a La Niña pattern. In the solar-minimum peak years, distribution of warm- and cold-ENSO years is even, and hence no coherent pattern was found (see the previous argument of [120]).

Recently, Haam and Tung [43] pointed out that most of the sunspot peak years studied by [148] and [147] c ould coincide with cold ENSO just by chance. Such a strongly biased result could possibly have emerged due both to the short length and quasi-periodicity of

the two time series, even if they are physically independent. By examining the much longer Quinn El Niño Index [106, 113, 112, 110, 111] which dates back to 1525, the authors have verified that cold- and warm-ENSO years are almost equally distributed during solar peak years, implying that the two phenomena are not systematically connected. They further performed a sophisticated statistical test on the relationship between the peaks of the 11-year solar cycle and the cold stages of ENSO. The test result either does not support the argument that solar peak years trigger La Niña phenomena.

## *5.2 A Nested Multiple Linear Regression Analysis*

### *5.2.1 Data and Methodology*

We use the instrumented record of surface temperature from 1880 to 2008, in the form of the ERSST (NOAA_ERSST_V3 data, provided by the National Oceanic and Atmospheric Administration/Office of Oceanic and Atmospheric Research/Earth System Research Laboratory Physical Sciences Division and available online at `http://www.cdc.noaa.gov/`), as described in [129, 130] and [131]. We also use the Hadley Center's Global Ocean Surface Temperature dataset, which is part of the HadISST [117] and is available from 1870 to the present (available online at `http://hadobs.metoffice.com/hadisst/`).

In our analysis, the common method of multiple linear regression (MLR) is employed with a modification to disentangle responses to multiple phenomena. The MLR technique assumes that the SST variation in space $\mathbf{x}$ (a vector) and time $t$ can be modeled by

$$\mathrm{SST}(\mathbf{x},t) = p_{\mathrm{S}}((\mathbf{x}))\mathrm{TSI}(t) + p_{\mathrm{E}}((\mathbf{x}))\mathrm{CTI}(t) + p_{\mathrm{G}}((\mathbf{x}))\mathrm{G}(t) + p_{\mathrm{V}}((\mathbf{x}))\mathrm{V}(t) + \epsilon(\mathbf{x},t), \quad (5.1)$$

where G is the anthropogenic greenhouse emissions [53] (text file available online at `http://data.giss.nasa.gov/modelforce/`), V is the volcano aerosol index [121], and $\epsilon$ is the remainder and may or may not be Gaussian noise. TSI is from [87] and [153], extended to 2008 and kindly provided to us by J. Lean (2010, personal communication). CTI is the averaged SST over 6°N-6°S, 180°-90°W minus the global mean SST(offered online by the University of Washington at `http://jisao.washington.edu/data/cti/`).

To take into account the possibility that CTI itself could be forced by TSI, G, and V,

(5.1) is supplemented by

$$\mathrm{CTI}(t) = \beta_{\mathrm{S}}\mathrm{TSI}(t) + \beta_{\mathrm{G}}\mathrm{G}(t) + \beta_{\mathrm{V}}\mathrm{V}(t) + R(t). \tag{5.2}$$

The residual $R$ obtained from (5.2) represents the unforced ENSO index. The nested model (5.1) + (5.2) is equivalent to model (5.1) alone, with the explanatory variable CTI replaced by $R$. No lag is used here in the response relative to the forcing index, although Lean and Rind [86] incorporated various lags in their multiple regression. The spatial patterns are little changed when we repeated the calculation with their lags.

### *5.2.2 Results*

Figure 5.1 (Figure 5.1) shows the various responses $p$ for the ERSST (HadISST) data. Figures 5.1(a) and 5.2(a) show the result of model (15.1) using CTI as the ENSO index, and Figures 5.1(b) and 5.2(b) use the unforced ENSO index $R$. A two-tailed Student's $t$-test (yellow contour) answers the question of whether the signal is different from zero, while the one-tailed $t$-test (black contour) tests whether the signal is positive or negative. Some of the previous applications of Student's $t$-tests may have overestimated the statistical significance by assuming $\epsilon$ to be white noise, while in reality serial correlations are usually present in the residuals from a regression analysis, as confirmed by the Durbin-Watson test in our case [27, 122] (see Figure 5.3(a)). Here the $t$-test is applied only over the locations that pass the Durbin-Watson test. In other regions, an autoregressive model of order one (AR(1)) is fitted to the residual $\epsilon$ and then the MLR analysis is performed after a prewhitening process (see Section 3 of Chapter 2). It appears appropriate to assume an AR(1) model for $\epsilon$ because the Durbin-Watson test also verifies that the prewhitened residuals are not autocorrelated at almost all spatial locations (see Figure 5.3(b)).

*Anthropogenic response*

Anthropogenic response is warming in almost the entire Pacific, and the signal is statistically significant above the 95% confidence level. There is very little projection of G onto CTI, and so there is not much difference between the left and right panels of either Figure 5.1 or Figure 5.2. The anthropogenic warming is consistent with the sea level pressure result

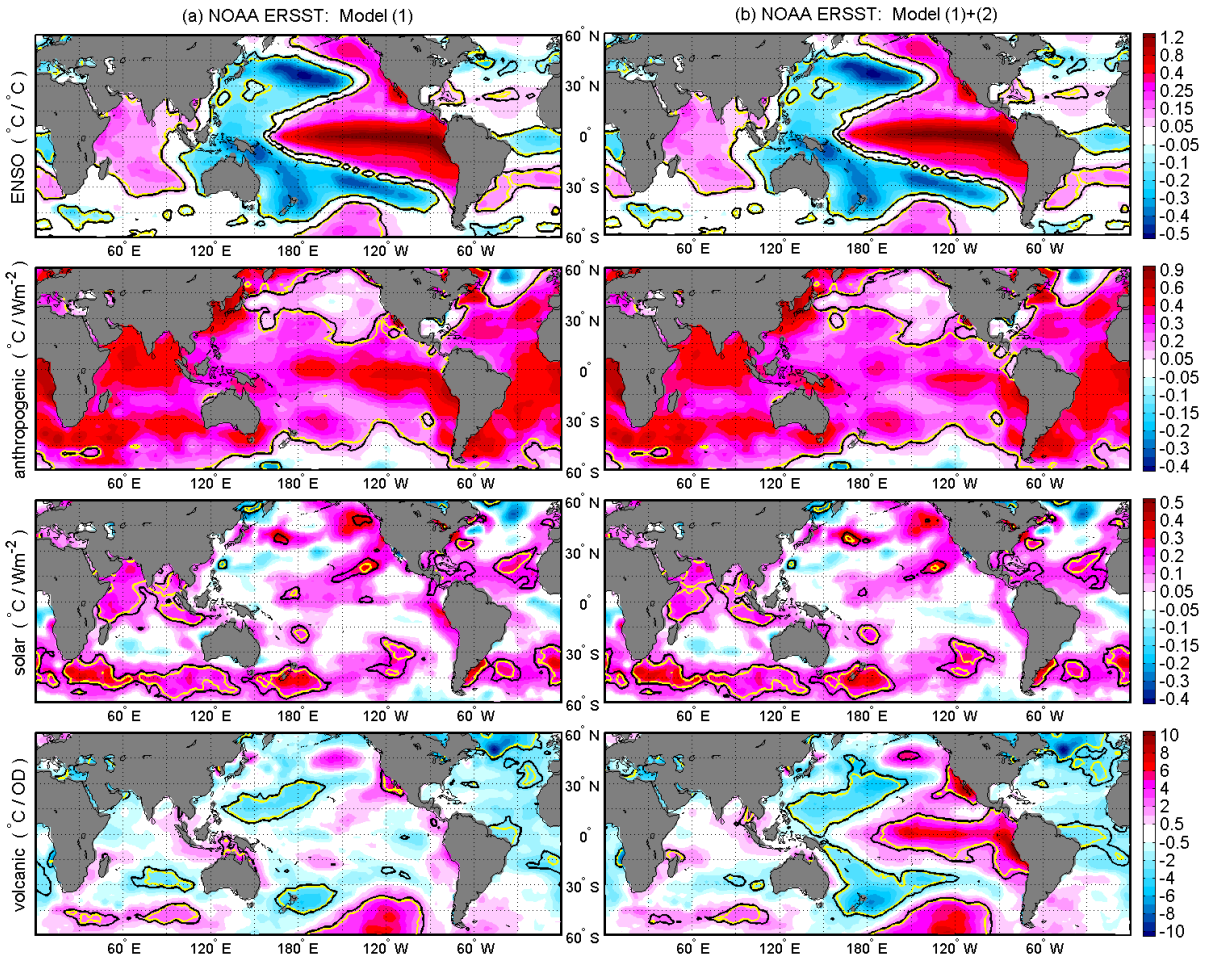


Figure 5.1: Spatial pattern in annual mean SST response obtained using (a) multiple linear regression method (5.1) and (b) model (5.1) + (5.2) using ERSST data for responses to (top-bottom) ENSO, net anthropogenic forcing, solar forcing, and volcano aerosols. Yellow (black) contours enclose regions of 95% confidence level in a two-tailed (one-tailed) test after prewhitening. The solar response is in degrees Celsius per watt per square meter of variation in the solar constant (TSI), which in recent decades (since direct satellite measurement) varies by about 1 $\mathrm{Wm}^{-2}$ between solar maximum and solar minimum. The anthropogenic response is in degrees Celsius per watt per square meter of net radiative forcing (RF; at the top of the troposphere). The RF change since 1880 is about 1.8 $\mathrm{Wm}^{-2}$. The ENSO response is in degrees Celsius per degrees Celsius of the CTI index. The aerosol response is in degrees Celsius per optical depth variation of the aerosol index.

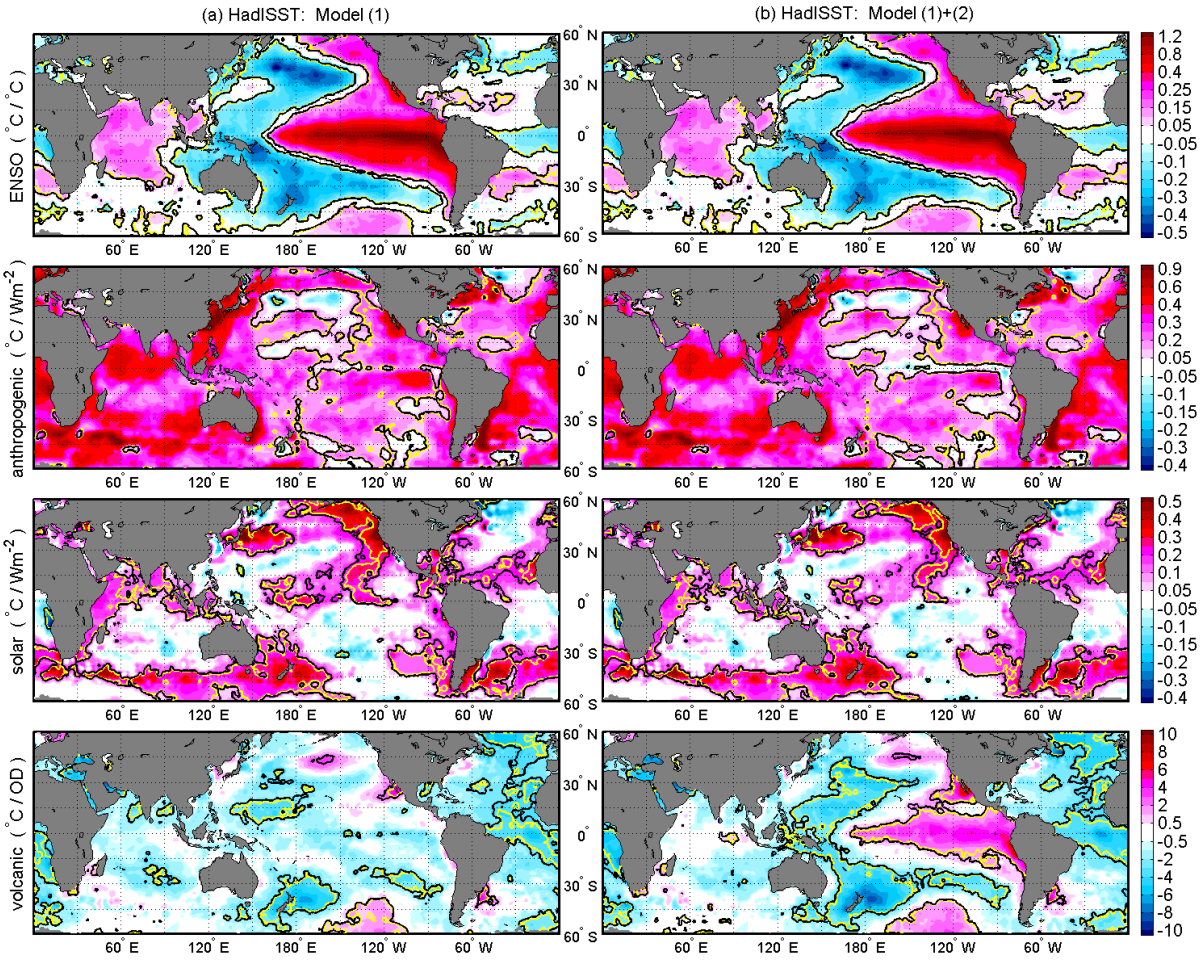


Figure 5.2: As in Figure 5.1, but using HadISST data.

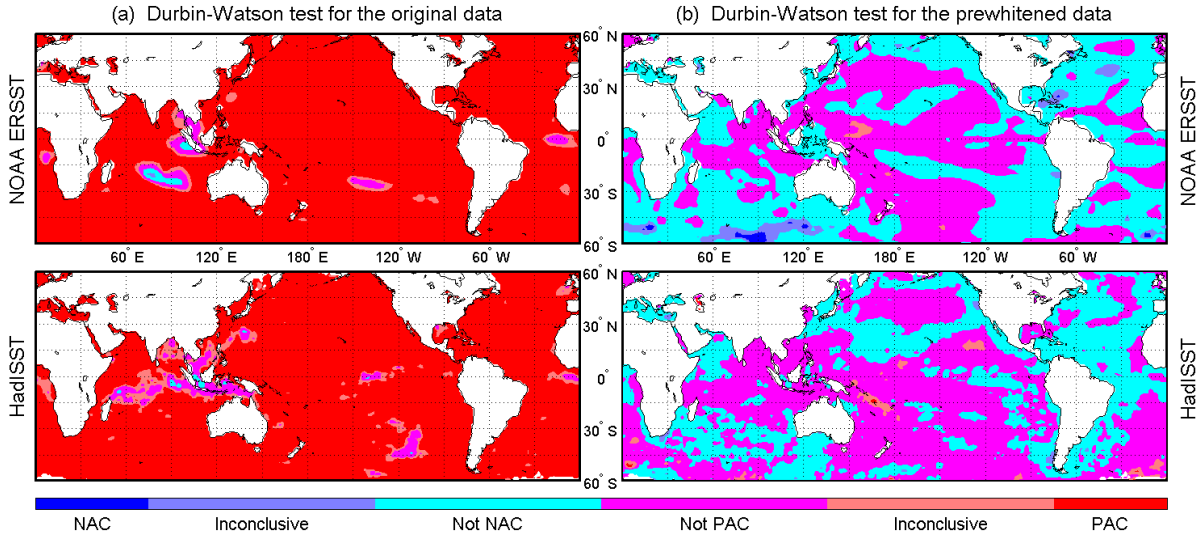


Figure 5.3: Durbin-Watson test on the residual (a) in the original data, which are shown to be positively autocorrelated (PAC), and (b) after one stage of prewhitening, which then become neither positively autocorrelated nor negatively autocorrelated (NAC), for (top) ERSST data and (bottom) HadISST data.

of [151] that showed a weakening of the Walker circulation. The results for the two datasets are consistent. There is, however, a very thin strip over the equator in HadISST where the small anthropogenic response is not statistically significant. Such a narrow strip of different behaviour likely indicates a data-quality problem. Our statistical test also shows a larger area of insignificance in the mid-Pacific in HadISST than in the ERSST case. Seasonal behaviour for the anthropogenic response is shown in Figure 5.4 for the two datasets using model (5.1) + (5.2), and it is consistent with the annual mean result. The contrasting behavior between the autumn and spring seasons mentioned by Karnauskas et al. [70] is not seen in our analysis. There is again a thin strip of possibly bad data (not statistically significant) in the HadISST along the equator, which is more confined to the mid- and eastern Pacific during boreal autumn, and this may have contributed to the strengthened zonal temperature gradient in [70].

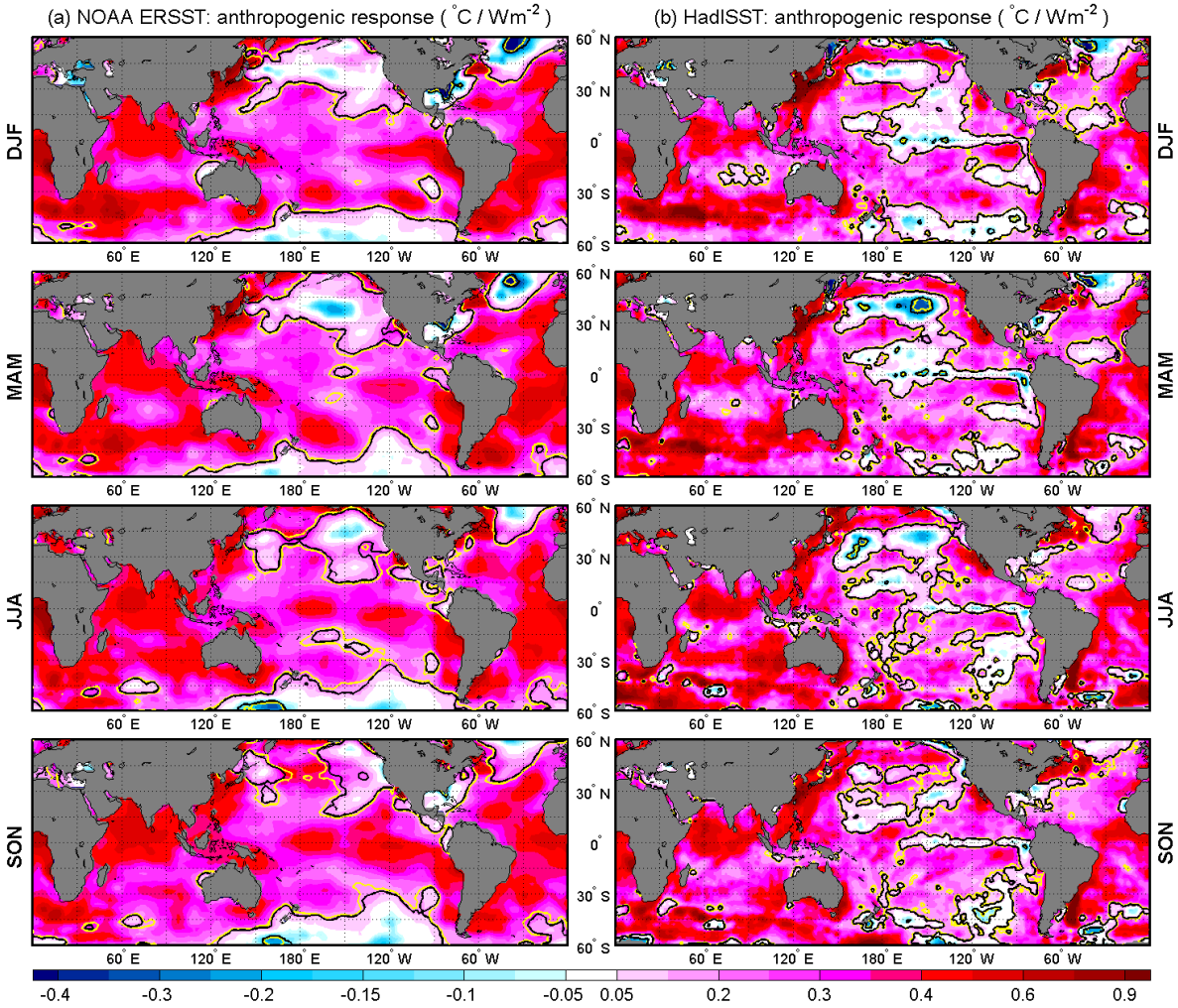


Figure 5.4: Seasonal anthropogenic responses from model (5.1) + (5.2), using (a) ERSST data and (b) HadISST data for 1882-2008. Monthly data for SST and the indices were used in the multiple regression, except that G was only available as annual means. Monthly data are not available before 1882 for the TSI.

*Solar response*

Solar response is weakly warm in the equatorial Pacific, as was previously found by Roy and Haigh [120] using a multiple regression model similar to our model (5.1) and by Zhou and Tung [169] using a different method. This warming is not statistically significant, in annual mean or in seasonal mean, when we use model (5.1) + (5.2) (Figure 5.5). It is neither El Niño-like nor La Niña-like. It looks nothing like the large ($\sim -1°$C) cold tongue found by van Loon et al. [148]. This does not mean that there is no solar response; it just means that the amplitude of the response in the equatorial Pacific is too weak. Zhou and Tung [169] established that the global SST pattern is related to solar forcing.

*Volcanic response*

Volcanic aerosol forcing projects partly onto the CTI. While the response obtained using CTI as the explanatory variable is weakly negative in the Pacific, it becomes warm-ENSO-like when CTI is replaced by $R$ in the analysis (Figure 5.6). The result is statistically significant and appears consistent with the earlier suggestion of [48]. However, there is uncertainty in the choice of the volcano index, and the time behavior of the response a few years after the eruption may not follow the aerosol optical depth.

We repeated the calculation without using V as one of the explanatory variables, and our previous responses to anthropogenic, solar, and ENSO indices are little changed; so, these responses are unaffected by the choice of volcano index.

## *5.3 Chapter Summary*

We have demonstrated that in the tropical Pacific the anthropogenic or solar forcing produces mostly a warmer sea surface temperature, but the spatial pattern is not in the form of an ENSO-like warm tongue or cold tongue. There is consistency in such responses in two different sea surface temperature datasets of long duration. The warming response in the tropical Pacific to greenhouse forcing is consistent with recent Intergovernmental Panel on Climate Change model results [149, 167]. The magnitude of solar warming is found in this region to be about $0.1° \pm 0.3°$C. The much larger response (of 1°C cooling in a beautiful

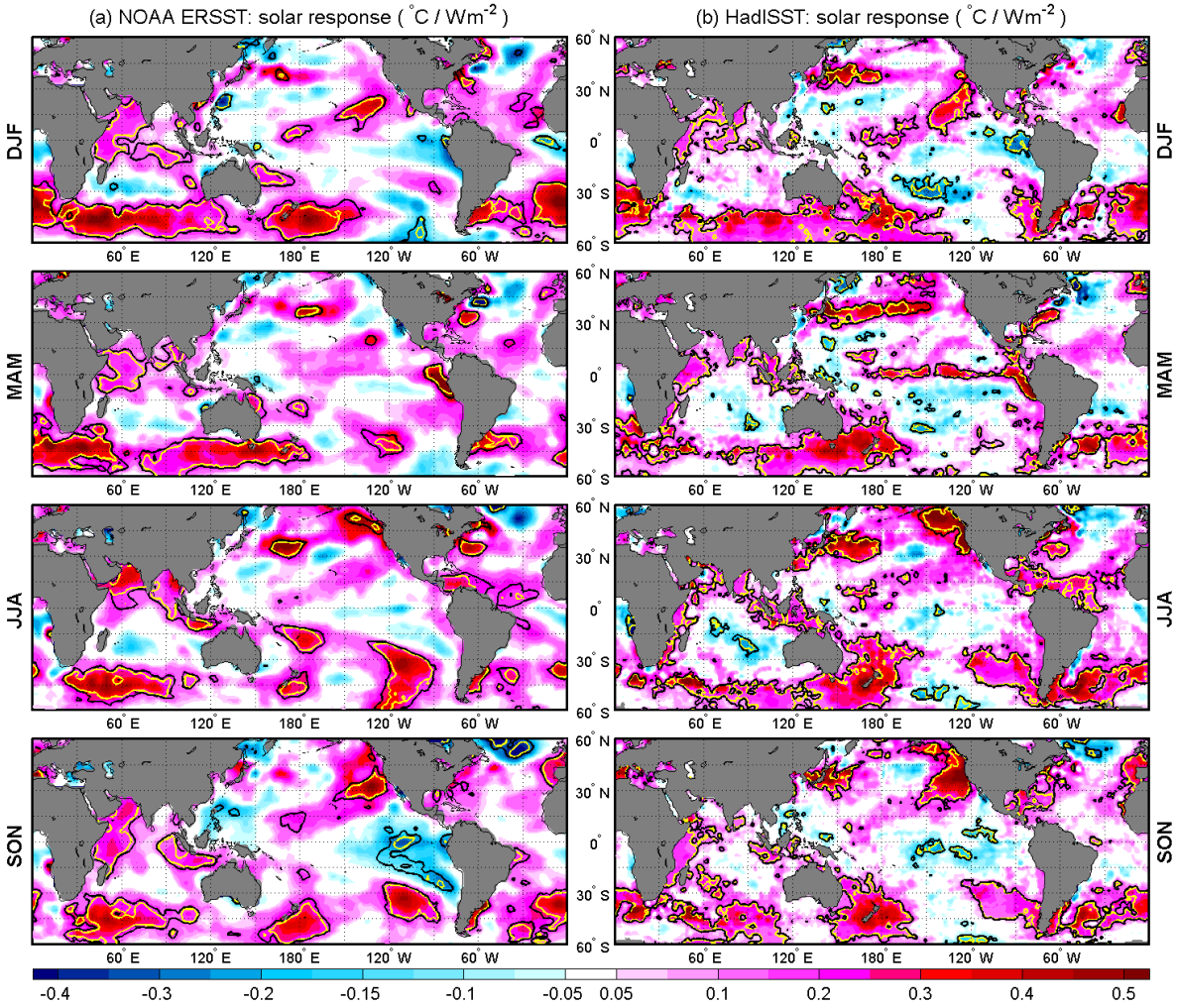


Figure 5.5: As in 5.4, but for the solar response.

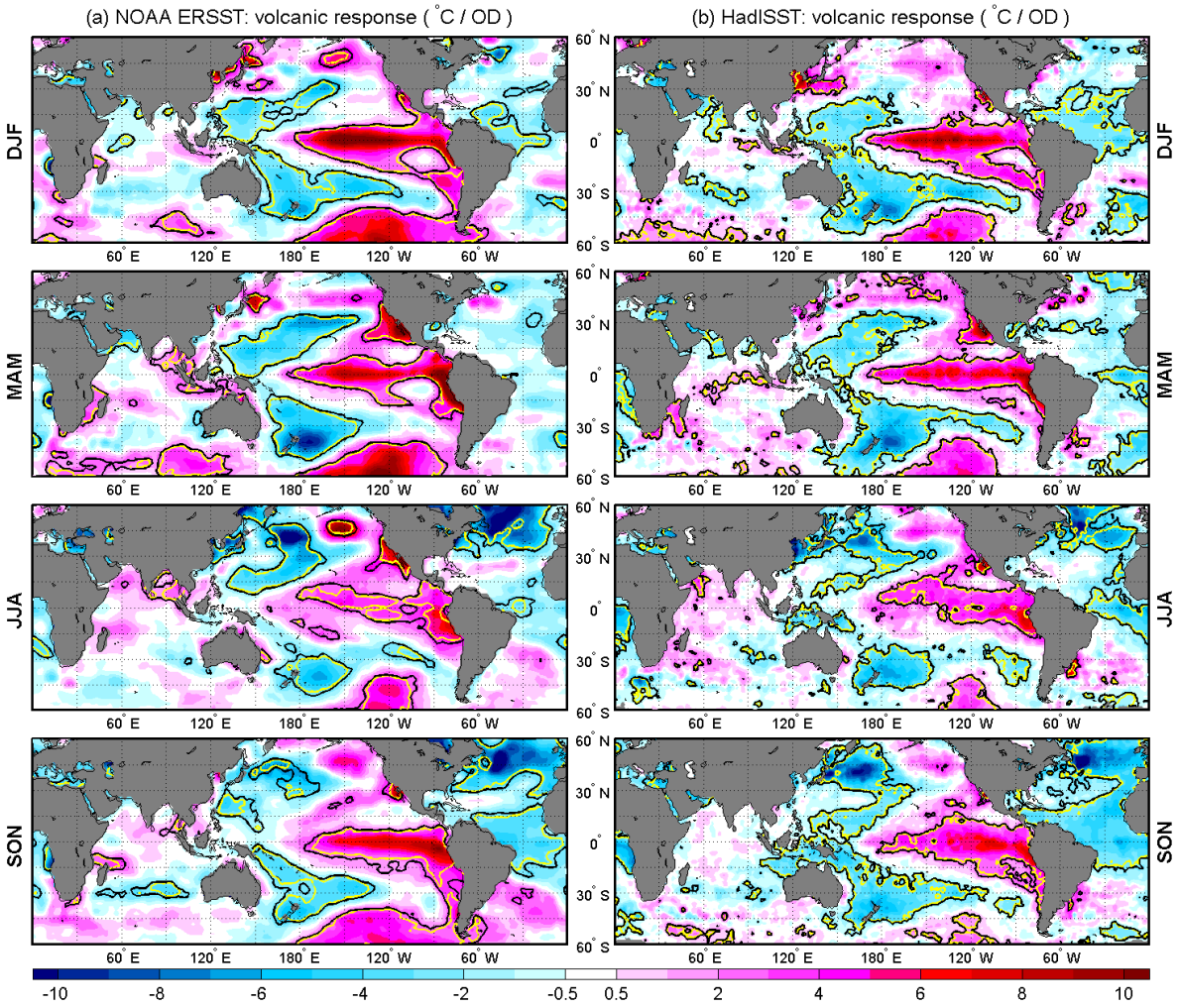


Figure 5.6: As in 5.4, but for the volcanic response.

cold tongue in the equatorial Pacific) found by van Loon and Meehl [147] is likely due to ENSO and not to the amplifying effect of positive cloud feedback.

# Chapter 6

# THE RECENT ANTHROPOGENIC GLOBAL WARMING

The common Detection and Attribution approach [56, 136] uses model-derived spatial patterns to apportion the observed recent warming into responses to various external forcing. Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4)[56] found that most of the observed increase in global temperatures since the mid-20th century was due to the observed increase in anthropogenic greenhouse gas concentrations, but what "most" means is controversial [21, 57] and needs further quantification. Models of intermediate complexity often used for attribution do not have ocean internal variability because of the slab ocean used. The most recent and comprehensive attribution study [64] underestimates internal variability. Coupled atmosphere-ocean models do have multi-decadal internal variability but not necessarily of the same amplitude and frequency as the observed. The phase in each realization may be different due to different initiation of the ocean state (except by chance [26, 25]). Ensemble averaging also tends to reduce multi-decadal internal variability. We use exclusively observational analysis: the latest global dataset HadCRUT4 [102] (1850-2010), which has additional stations in the Arctic, and the regional record of Central England Temperature (CET) [107, 67] (1659-present).

Our analysis is performed using both Fourier and Wavelet methods. Low-frequency components of CET and the global mean HadCRUT4 [102] time series are first extracted by means of Fourier filtering. Statistical significances of the Fourier spectral peaks of these data are examined against the red-noise background from an autoregressive process of order one (AR(1)). We also estimate the errors of the extracted modes using bootstrap Monte Carlo simulations. A continuous wavelet analysis is then performed to add another piece of evidence.

## *6.1 Fourier Analysis and Chi-squared Test*

### *6.1.1 Linear detrending*

The Fourier analysis is always performed on the linearly detrended monthly time series, by fitting a linear regression model using the ordinary least-squares approach. By assuming that the regression residual is a red noise time series obeying an AR(1) process, the 95% confidence intervals for the linear trends are estimated. These trends are robust and not sensitive to dropping one to 5 years at the end of the records (Since England experienced three consecutive cold winters recently due in part to regional blocking, sensitivity to end points was a concern). The Fourier components of the detrended data remain quite robust while allowing the linear trend to vary within its 95% confidence limits. Note that the annual cycle is removed from the monthly CET series before doing the error analysis of the linear trend because it is known to be a natural component rather than random noise. Due to the fact that the annual cycle contains most of the data variance, the estimated errors are falsely large without de-seasoning. Here the mean of a particular month over the whole time domain when CET data are available is taken to be the value of the annual cycle for that particular month. It is worth pointing out that the linear trend and its error bars are insensitive to the definition of the annual cycle. Changing the reference period over which the mean is calculated, for example, only imposes negligible effects on our results, and our analysis is also not much influenced by removing a Modulated Annual Cycle [166] obtained using some sophisticated techniques such as the Empirical Mode Decomposition [62, 163].

### *6.1.2 Errors in the Band-passed and Low-passed Results*

An "ideal" Fourier filter in the frequency sense is used in this study. It perfectly has a gain of 1 within the desired passband and a gain of 0 for all stopband frequencies. This filter in the time domain is in the form of a *sinc* function (for lowpass) or the difference of two *sinc* functions (for bandpass). Generally it is not feasible to apply the "ideal" filter to any practical problem because it is noncausal and has an infinite length. Therefore, in real applications it is commonly truncated to a finite length and optionally windowed to reduce the Gibbs phenomenon which is the penalty incurred for truncating. However, the filter

length required to create a good frequency response for our target lowpass is much longer than the length of the world's longest instrumental data record (352 years of CET). A lot of undesired high frequencies must be leaking into the filtered components if the lowpass filter is windowed to be short enough such that one will not suffer a severe data shortage when implementing it on the available CET and global-mean HadCRUT4 series. Considering this difficulty, here the filter is not truncated; instead the time domain is artificially extended by padding an infinite number of zeros in both positive and negative time directions. By this means the filter itself is flawless, but do data beyond the real time domain contribute to the band-passed and low-passed results? Since we do not have these data, the effects are then investigated using a Monte Carlo simulation.

The unavailable data in the past and the future are filled in using bootstrap resampling with replacement of the observed time series. The time domain before the beginning and after the end of the real data is divided into many time slots of fixed length (say $L$ years). For each time slot we randomly select a block of $L$ years from the instrumental record and assign it after taking out a linear trend to an empty time slot. The trend removal from each data block is performed to reduce artificial oscillations resulting from the discontinuous trends between blocks. A synthetic time series as long as desired with the real observations residing at the center can be produced using this so-called *moving-block bootstrap* method [29, 79, 89, 157] which has also been used in previous chapters. We generate a large number (e.g., 1000) of synthetic monthly time series, each of $2^{18}$ data points (totaling approximately 21845 years), to mimic the natural temperature data over such a long period and to estimate the contribution of the unobtainable data to the band-passed and low-passed results. The least-squares trend and its upper and lower 95% confidence limits are taken out, separately, from the central part of each synthetic time series, which consists of the real data, to account for the influence of the regression errors in Fourier analysis, and then added back after band pass of the Fourier modes. For our problem the impulse response of the ideal filter decays to be sufficiently small when it is more than $2^{17}$ months away from the center of the filter. Therefore the ideal filter is truncated to have a length of 21845 years when we carry out Monte Carlo experiments to study the errors in the results. The truncation is equivalent to a multiplication by a rectangular window function in the time domain. Alternatively we can

also multiply the impulse response by some other window function (Gaussian, Hamming, etc.) of the same length to approximate the ideal filter. The changes to our results caused by choosing different window functions are within the error bands shown.

The block length $L$ needs to be large enough to take into account as many inherent characteristics of the climate time series as possible for the purpose of simulating the real-world data. On the other hand, the choice of data blocks will be very limited if $L$ is too large so that we lose randomness in the synthetic time series. As a compromise between these two considerations, the block length $L$ is chosen to be 200 (100) years in the analysis of the 352 (161)-year data. These blocks are time-reversed when used to create the future data, to avoid large discontinuities near 2010. These form the range of possible errors in the present analysis arise from the use of Fourier frequency windows, including the end-point effects.

### *6.1.3 Chi-squared test*

We choose the AR(1) red noise as the background to determine the significance levels for the Fourier power spectra. As is typically assumed in the analysis of geophysical time series, we take the Gaussian probability density for the noise. The *null hypothesis* for the significance test of a Fourier spectral peak then states that there is zero signal power at the corresponding harmonic. Gilman et al. [37] estimated the discrete Fourier power spectrum of an AR(1) red noise time series with lag-1 autocorrelation $\rho$, after normalizing, to be

$$P_k = \frac{1-\rho^2}{1+\rho^2-2\rho\cos\left(\frac{2\pi k}{N}\right)}, \tag{6.1}$$

where $N$ is the number of points in the frequency domain and $k = 0, 1, \cdots, N/2$ is the frequency index up to the Nyquist frequency ($k = N/2$). Given a time series $\{x_j\}$ with variance $\sigma^2$ and denote its discrete Fourier transform (DFT) by $\hat{x}_k$, in some literature such as [140] it is claimed that under the null hypothesis the power in the $k^{\text{th}}$ harmonic normalized by the average background noise power at the corresponding frequency index is distributed as $\chi^2_2$ (the chi-squared distribution with 2 degrees of freedom):

$$\frac{2N|\hat{x}_k|^2}{\sigma^2 P_k} \sim \chi^2_2. \tag{6.2}$$

This is true if we assume the background spectrum is a Gaussian white noise. However, in the case of red noise, (6.2) fails because the real and imaginary parts of $\hat{x}_k$ are not statistically independent, although they are both normally distributed. The statistical significance of the Fourier power spectrum can still be established using a chi-squared test after doing slightly more mathematical manipulations.

Suppose a time series $\{x_j\}$ is the sum of the signal $\{y_j\}$ and a sample $\{\boldsymbol{\xi}_j\}$ of random noise where $\boldsymbol{\xi}_j = \rho\boldsymbol{\xi}_{j-1} + \boldsymbol{\varepsilon}_j$ is an AR(1) process with mean zero and variance $\sigma^2_{\boldsymbol{\xi}}$, and $\{\boldsymbol{\varepsilon}_j\}$ is a Gaussian white noise time series for $j = 0, 1, \cdots, n-1$. Based on the DFT the time series can be represented as follows:

$$\begin{aligned} x_j &= \hat{a}_0 + \sum_{k=1}^{N-1} \hat{a}_k \cos\left(\frac{2\pi jk}{N}\right) + \hat{b}_k \sin\left(\frac{2\pi jk}{N}\right) \\ y_j &= \tilde{a}_0 + \sum_{k=1}^{N-1} \tilde{a}_k \cos\left(\frac{2\pi jk}{N}\right) + \tilde{b}_k \sin\left(\frac{2\pi jk}{N}\right) \\ \boldsymbol{\xi}_j &= \boldsymbol{a}_0 + \sum_{k=1}^{N-1} \boldsymbol{a}_k \cos\left(\frac{2\pi jk}{N}\right) + \boldsymbol{b}_k \sin\left(\frac{2\pi jk}{N}\right). \end{aligned} \tag{6.3}$$

Since the Fourier transform is a linear operation, we have $\hat{a}_k = \tilde{a}_k + \boldsymbol{a}_k$ and $\hat{b}_k = \tilde{b}_k + \boldsymbol{b}_k$. Direct calculations yield

$$\begin{aligned} \boldsymbol{a}_k &= \left[\ \boldsymbol{\xi}_0,\ \ \tfrac{\varepsilon_1}{\sqrt{1-\rho^2}},\ \ \cdots,\ \ \tfrac{\varepsilon_{n-1}}{\sqrt{1-\rho^2}}\ \right] \cdot \mathbf{A} \\ \boldsymbol{b}_k &= \left[\ \boldsymbol{\xi}_0,\ \ \tfrac{\varepsilon_1}{\sqrt{1-\rho^2}},\ \ \cdots,\ \ \tfrac{\varepsilon_{n-1}}{\sqrt{1-\rho^2}}\ \right] \cdot \mathbf{B}. \end{aligned} \tag{6.4}$$

Both $\boldsymbol{a}_k$ and $\boldsymbol{b}_k$ are normally distributed random variables with the $\ell^{\text{th}}$ ($\ell = 0, 1, \cdots, n-1$) elements of the column vectors $\mathbf{A}$ and $\mathbf{B}$ being, respectively,

$$\begin{aligned} \mathbf{A}_\ell &= \frac{s\left(\cos\left(\frac{2\pi k\ell}{N}\right) - \rho\cos\left(\frac{2\pi k(\ell-1)}{N}\right) - \rho^{n-\ell}\cos\left(\frac{2\pi kn}{N}\right) + \rho^{n-\ell+1}\cos\left(\frac{2\pi k(n-1)}{N}\right)\right)}{N\left(1+\rho^2-2\rho\cos\left(\frac{2\pi k}{N}\right)\right)} \\ \mathbf{B}_\ell &= \frac{s\left(\sin\left(\frac{2\pi k\ell}{N}\right) - \rho\sin\left(\frac{2\pi k(\ell-1)}{N}\right) - \rho^{n-\ell}\sin\left(\frac{2\pi kn}{N}\right) + \rho^{n-\ell+1}\sin\left(\frac{2\pi k(n-1)}{N}\right)\right)}{N\left(1+\rho^2-2\rho\cos\left(\frac{2\pi k}{N}\right)\right)} \end{aligned} \tag{6.5}$$

where the scale factor $s = 1$ for $\ell = 0$ and $s = \sqrt{1-\rho^2}$ for $\ell = 1, \cdots, n-1$. For the two special cases of $k = 0$ and $N/2$, we have $\boldsymbol{b}_k = 0$ and $\boldsymbol{a}_k^2$ normalized by the variance of $a_k$ is a chi-squared random variable with one degree of freedom. For $k = 1, \cdots, N/2-1$, $a_k$ and $b_k$

are not necessarily independent. But we can always make a transform $\boldsymbol{\alpha}_k = \boldsymbol{a}_k + w\boldsymbol{b}_k$ and $\boldsymbol{\beta}_k = u\boldsymbol{a}_k + v\boldsymbol{b}_k$ such that $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ are independent, standard normal random variables. Therefore $\boldsymbol{\alpha}_k^2 + \boldsymbol{\beta}_k^2 \sim \chi_2^2$. A nice property of this newly defined chi-squared random variable is that it is independent of the choices of the parameters $w$,$u$ and $v$:

$$\begin{aligned} \boldsymbol{\alpha}_k^2 + \boldsymbol{\beta}_k^2 &= \frac{1}{\sigma_{\boldsymbol{\xi}}^2} \cdot \frac{\mathbf{B}^{\mathrm{T}}\mathbf{B} \cdot \boldsymbol{a}_k^2 - 2\mathbf{A}^{\mathrm{T}}\mathbf{B} \cdot \boldsymbol{a}_k\boldsymbol{b}_k + \mathbf{A}^{\mathrm{T}}\mathbf{A} \cdot \boldsymbol{b}_k^2}{\mathbf{A}^{\mathrm{T}}\mathbf{A} \cdot \mathbf{B}^{\mathrm{T}}\mathbf{B} - \left(\mathbf{A}^{\mathrm{T}}\mathbf{B}\right)^2} \\ &= \frac{1}{\sigma_{\boldsymbol{\xi}}^2} \cdot \frac{\mathbf{B}^{\mathrm{T}}\mathbf{B} \cdot (\hat{a}_k - \tilde{a}_k)^2 - 2\mathbf{A}^{\mathrm{T}}\mathbf{B} \cdot (\hat{a}_k - \tilde{a}_k)\left(\hat{b}_k - \tilde{b}_k\right) + \mathbf{A}^{\mathrm{T}}\mathbf{A} \cdot \left(\hat{b}_k - \tilde{b}_k\right)^2}{\mathbf{A}^{\mathrm{T}}\mathbf{A} \cdot \mathbf{B}^{\mathrm{T}}\mathbf{B} - \left(\mathbf{A}^{\mathrm{T}}\mathbf{B}\right)^2}. \end{aligned} \tag{6.6}$$

This result gives us a way to test for the presence of signals at particular harmonic frequencies. In this case, under the null hypothesis that the Fourier coefficients of the $k^{\text{th}}$ harmonic of the signal are zero, i.e., $\tilde{a}_k = 0$ and $\tilde{b}_k = 0$, (6.7) simplifies to

$$\frac{1}{\sigma_{\boldsymbol{\xi}}^2} \cdot \frac{\mathbf{B}^{\mathrm{T}}\mathbf{B} \cdot \hat{a}_k^2 - 2\mathbf{A}^{\mathrm{T}}\mathbf{B} \cdot \hat{a}_k\hat{b}_k + \mathbf{A}^{\mathrm{T}}\mathbf{A} \cdot \hat{b}_k^2}{\mathbf{A}^{\mathrm{T}}\mathbf{A} \cdot \mathbf{B}^{\mathrm{T}}\mathbf{B} - \left(\mathbf{A}^{\mathrm{T}}\mathbf{B}\right)^2} \sim \chi_2^2. \tag{6.7}$$

Since the signal $y_j$ is assumed to be noise-free, the variance $\sigma_{\boldsymbol{\xi}}^2$ of the noise is equal to the variance $\sigma^2$ of $x_j$. When performing the chi-squared test according to (6.7), $\sigma_{\boldsymbol{\xi}}^2$ is estimated as the sample variance obtained for the observed time series $\{x_j\}$. In addition, the lag-1 autocorrelation $\rho$ of the background red noise is estimated from $\left(\rho_1 + \sqrt{\rho_2}\right)/2$, where $\rho_1$ and $\rho_2$ are the lag-1 and lag-2 autocorrelations of $\{x_j\}$.

The Fourier power spectrum of the global-mean HadCRUT4 temperature and that of the difference between it and CET are shown in Figure 6.1. Chi-squared tests were performed against AR(1) red noise and the statistically significant spectral peaks at a 5% level of significance are indicated in red.

### *6.2 Continuous Wavelet Transform*

The Fourier transform which represents the signal as a sum of sinusoids is a global analysis in time and thus could create a global oscillation through the entire time domain when one does not exist. Its another drawback is that it is only applicable to stationary time series. As a comparison, the wavelet transform possesses the ability to construct a time-frequency representation of a signal that offers very good time and frequency localization. It can be used to analyze time series that contain nonstationary power at many different

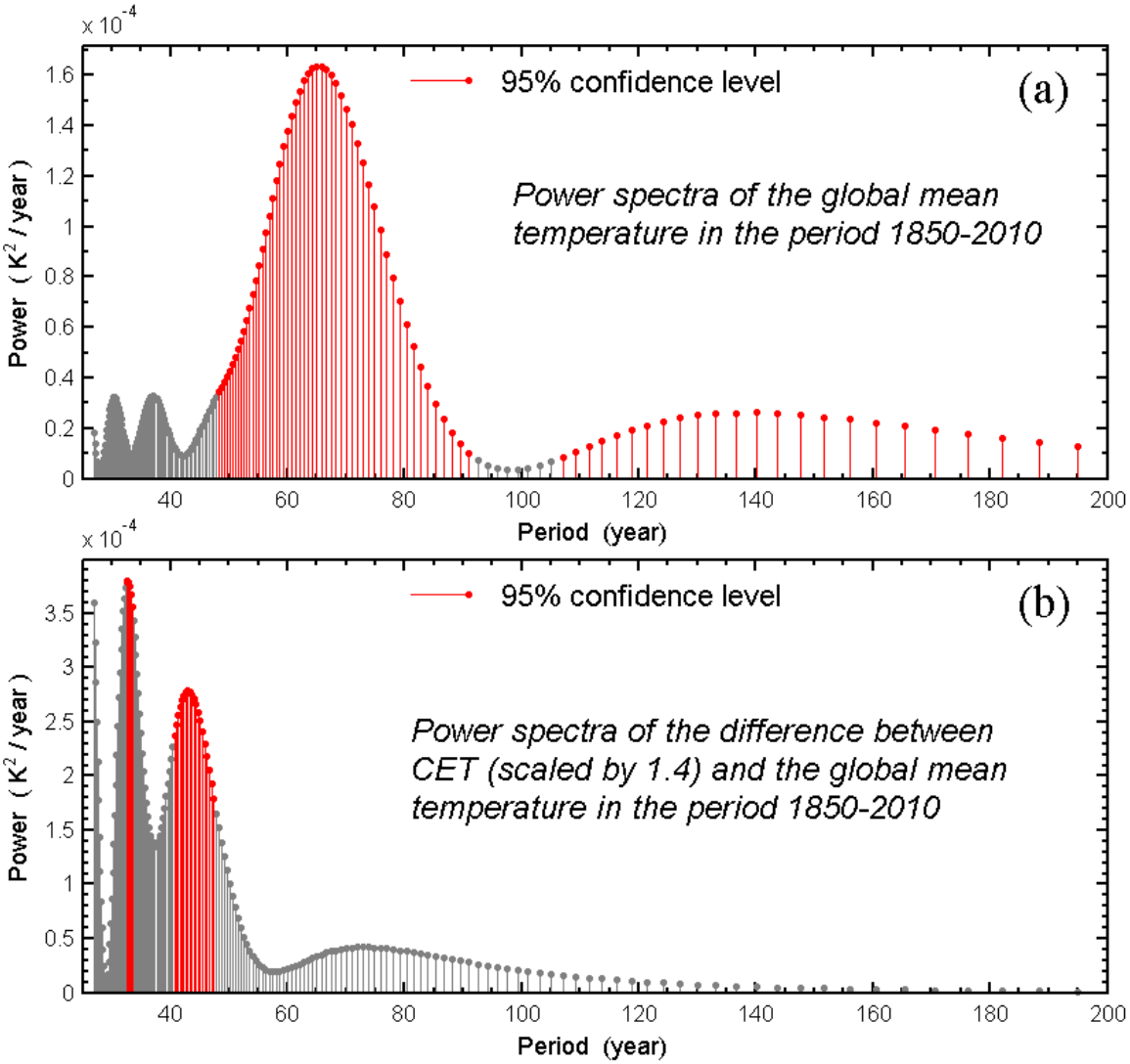


Figure 6.1: **(a)** Fourier spectrum of the global-mean HadCRUT4 temperature for the period 1850-2010. The time series is padded with zeros before 1850 and after 2010. Red indicates the part of the spectrum that is statistically significant at 95% confidence level again red noise. **(b)** The Fourier spectrum of the difference of the scaled CET and global-mean temperature for the period 1850-2010, when they overlap. The difference is not statistically significant for periods lower than 50 years. The conclusion is unchanged using the unscaled CET data.

frequencies [22]. Given a mother wavelet funciton $\psi(t)$ which is a continuous function in both the time domain and the frequency domain, with zero mean and unit energy, the continuous wavelet transform of a continuous, square-integrable function $x(t)$ at a scale $s > 0$ and translational value $\tau$ is expressed by the following integral

$$W(\tau, s) = \int_{-\infty}^{\infty} x(t) \cdot \frac{1}{\sqrt{|s|}} \psi^* \left( \frac{t-\tau}{s} \right) dt, \tag{6.8}$$

where $*$ represents operation of complex conjugate. Varying the scale $s$ provides information about the power spectrum of the signal $x(t)$ at any localized time $\tau$. The inverse continuous wavelet transform can be exploited to recover the original signal $x(t)$:

$$x(t) = \frac{1}{C_\psi} \int_0^{\infty} \int_{-\infty}^{\infty} \frac{1}{s^2} W(\tau, s) \cdot \frac{1}{\sqrt{|s|}} \psi \left( \frac{t-\tau}{s} \right) d\tau ds, \tag{6.9}$$

where

$$C_\psi = \frac{1}{2} \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega, \tag{6.10}$$

is called the *admissibility constant* and $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(t)$. To be "admissible" as a wavelet, the admissibility constant has to satisfy the admissibility condition: $0 < C_\psi < +\infty$. In this situation, the inverse transform is successful and thus can be used to design band-pass and low-pass filters. The Morlet wavelet is employed to be the mother wavelet in our analysis, which consists of a plane wave modulated by a Gaussian and takes the following form:

$$\psi(t) = \left( 1 + \kappa_\sigma^2 - 2\kappa_\sigma^{\frac{3}{2}} \right)^{-\frac{1}{2}} \pi^{-\frac{1}{4}} e^{-\frac{1}{2}t^2} \left( e^{i2\pi\sigma t} - \kappa_\sigma \right), \tag{6.11}$$

where the constant $\kappa_\sigma = e^{-2\pi^2\sigma^2}$ is defined to satisfy the admissibility condition [140].

Due to the finite length of the time series to be analyzed, errors will occur at the beginning and end of the wavelet power spectrum. Therefore, in wavelet analysis, the *cone of influence* (COI) is defined to indicate the region of the wavelet spectrum in which edge effects become strong. The COI in our analysis is determined by the $e$-folding time for the wavelet power of a single spike (i.e., the Dirac delta function $\delta(t)$) at the edge to drop by a factor $e^{-2}$ [140]. Figure 6.2 illustrates the Morlet wavelet spectrum of the CET time series, with the COI indicated by masked areas.

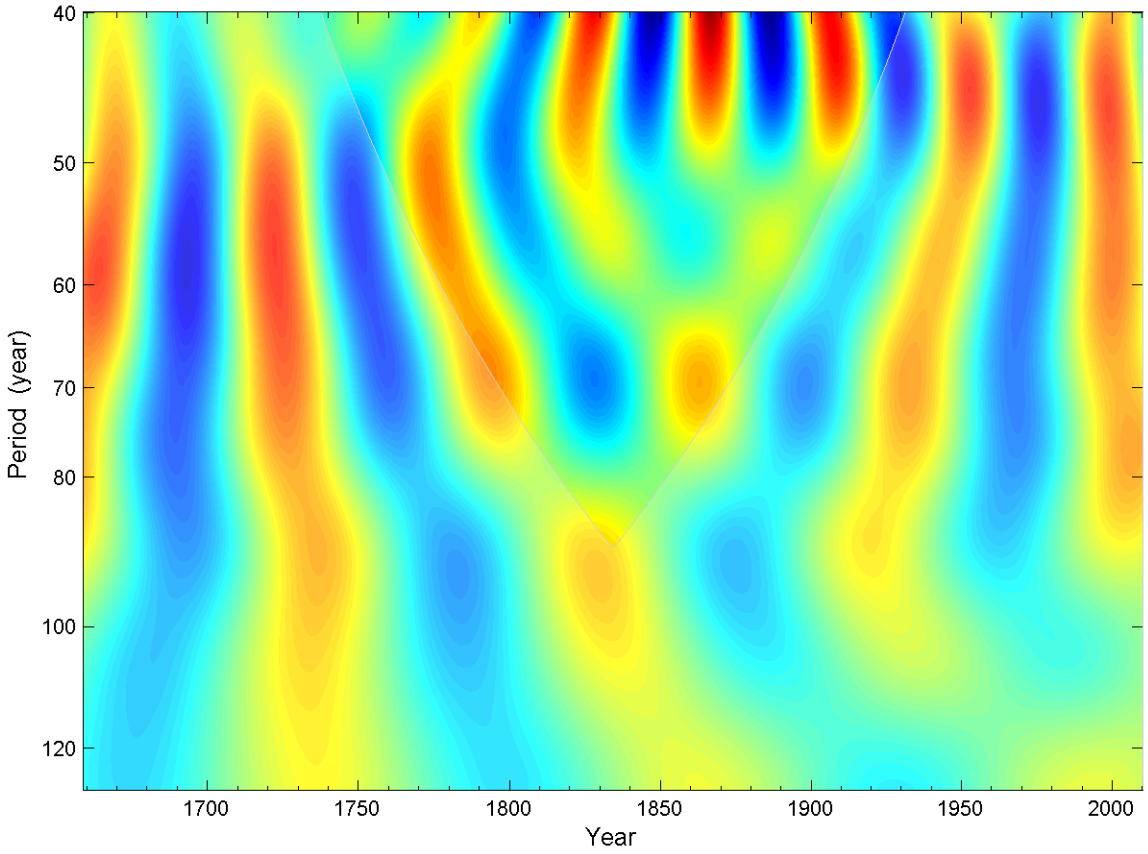


Figure 6.2: Morlet wavelet spectrum of linearly detrened CET (1659-2010). Wavelet transform is applicable to nonstationary time series so detrending is actually not necessary. Applying the wavelet transform to the original CET time series without a trend removal yields almost the same power spectrum. The parameter $\sigma$ in (6.11) is chosen to be 1.4. The masked areas indicate the cone of influence. The shown wavelet spectrum is normalized to satisfy the energy conservation property, i.e., the integral of the power spectrum over the whole time-frequency plane is equal to the energy in the original time series.

### *6.3 Global Warming in HadCRUT4 Data*

Our main result will first be obtained using the recently available global temperature from HadCRUT4 [102]. The authors quoted a fitted linear trend of 0.07°C per decade from 1901 to 2010, and 0.17°C per decade from 1979 to 2010 globally. Various fitted linear trends in global mean temperature up to 2005 were presented in IPCC AR4 [142], with the recent 25-year trend (at 0.177°C per decade) larger than the 50-year trend (0.128°C per decade), which is in turn larger than the 100-year and 150-year trends (0.074 and 0.045°C per decade). The result is updated in Figure 6.3a. The phenomenon of "accelerating warming trends" is still present. The fitted 25-year trend is not robust, being sensitive to the addition or subtraction of a single-year end-point data, and so will not be discussed further here. A 50-year low-pass filter is applied to the data points. It is seen by eye that the smoothed curve captures the main episodes of warming and cooling in the past 160 years, as it agrees with the simple running mean. In particular, one can see that there is a low-frequency oscillation present in the data. Since linear trends should not be applied to a time series with an imbedded oscillation whose half period is comparable to the interval over which the trends are fitted, we reprocess the data after removing the oscillatory component, defined by the 50-80-year band pass. This component will be called here "Atlantic Multi-decadal Oscillation" (AMO); justification will follow. Linear trends are then fitted to the resulting data points in Figure 6.3b in the same way as in AR4. It is visually apparent in Figure 6.3b that removing the oscillatory AMO from the raw data organizes the data points into a monotonic band and yields a more stable linear trend, converging to the 50-year trend of 0.08°C per decade. We will later argue that this is the long-term anthropogenic trend, forced by greenhouse gas increases offset by tropospheric aerosol cooling, which also increased along with industrialization. Although aerosol forcing from tropospheric pollution is not expected to be smoothly increasing, the AR4 methodology of fitting linear trends using intervals longer than 25 years, also adopted here, yields only a smoothed version, and so our results should be interpreted accordingly. Comparing Figure 6.3b with 6.3a, we see that the AMO accounts for 40% of the observed 50-year trend. This is essentially the conclusion of [165], except that they could not rule out time-varying aerosol forcing as a possible cause,

which we will pursue more later. They used an entirely different method of Empirical Mode Decomposition [62, 163]. This adds confidence to the methods used by the two groups.

Foster and Rahmstorf [34] considered the period 1979-2010 and obtained an estimate of anthropogenic warming after removing ENSO, individual volcano eruptions and the solar forcing through multiple regression analysis. Their adjusted data also showed no change in the linear warming rate over the period, but their estimate for anthropogenic warming, at 0.17°C per decade for HadCRUT3v [7], is twice as high as ours for the same dataset, due to their inability to remove multi-decadal internal variability in the short record: the oscillatory component of AMO, which happens to be in a positive half cycle, appears as a positive trend in their data since 1979. This is demonstrated in Figure 6.4, where Foster and Rahmstorf's analysis is repeated, but for the longer record 1850-2010 and using HadCRUT4 (HadCRUT3v results are similar. Their index for ENSO is not available over the longer period, and so instead we use the Cold Tongue Index.) The residual plus the regressed linear trend is shown in Figure 6.4a. The same 32-year "anthropogenic trend" of 0.17°C per decade is obtained. However, one can see clearly that a 70-year oscillation is still present in the residual (see the orange running mean). In Figure 6.4b, we add the AMO Index [30] in the multiple linear regression analysis. The anthropogenic warming rate obtained, at 0.07°C per decade, is less than half of theirs. Given the error bar of the short record, it is still consistent with the conclusion of no change from the 50-year trend. The 50-year trend is 0.08°C per decade, same as in Figure 6.3b. The AMO is thus seen to be the most important internal variability as far as its effect on the deduced anthropogenic warming trend is concerned.

We next give a mathematical and a physical justification to the removal of AMO to arrive at a robust underlying trend. We adopt the mathematical definition of Secular Trend as the *least monotonic* function of time. It is obtained here by adding low-frequency Fourier modes to the linear trend for the period under consideration until the monotonicity is about to fail. This definition, as all such definitions of trend, is dependent on the length of the record. For the 160-year period of global-mean data, the Secular Trend is obtained with 80-year low pass. The threshold, 80-years, then separates the Secular Trend from the Oscillations. It is seen in Figure 6.3b that this monotonic function gives a good smoothed fit to the

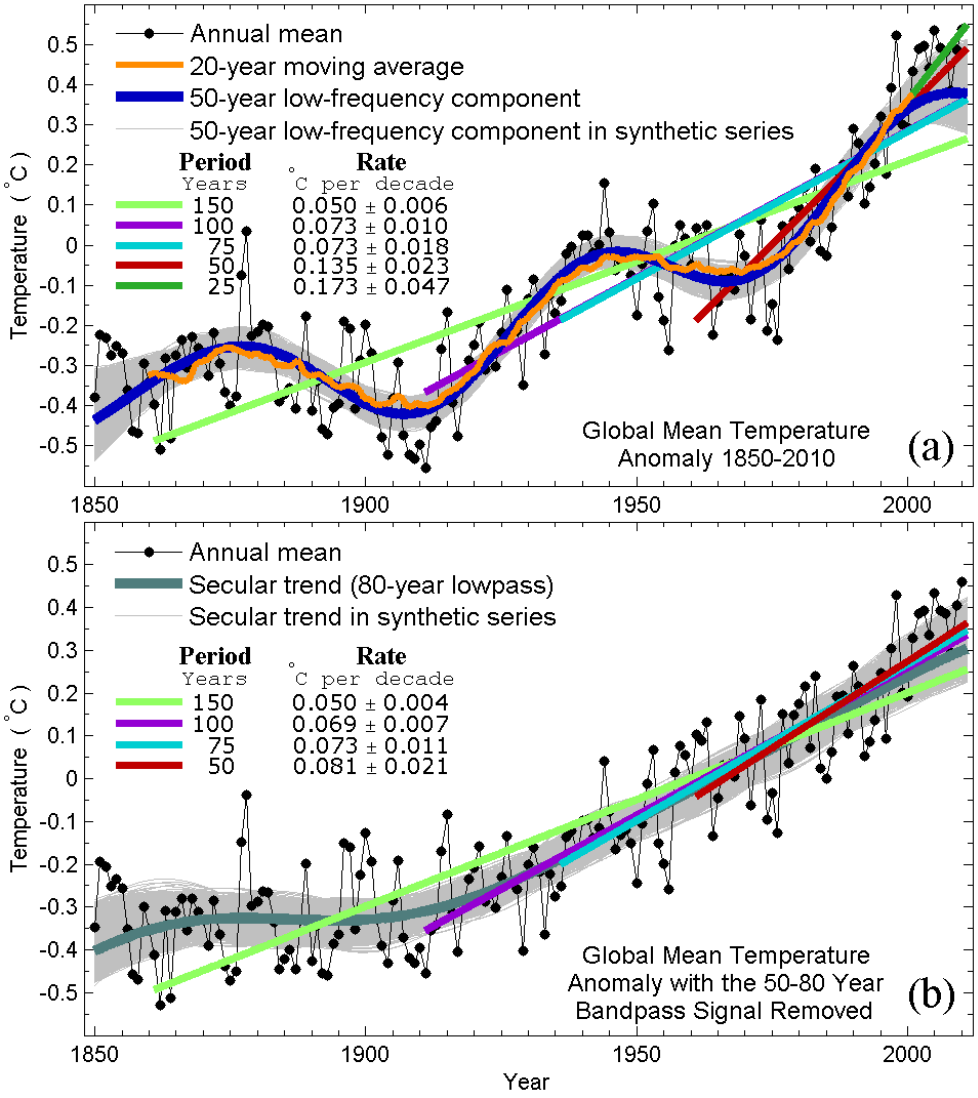


Figure 6.3: **(a)** the result in IPCC AR4 is updated using HadCRUT4 from 1850 to 2010, and their 150-year, 100-year, 50-year and 25-year linear trends (with labeled magnitudes) are shown. Even though the annual means are shown, the linear trends are calculated using the raw monthly mean data. The 50-year low-passed global mean is superimposed (in blue). It gives a smoothed fit to the raw data and agrees with the running-time mean. The blue curve consists entirely of the sum of the Secular Trend and the AMO mode. **(b)** The dots represent adjusted annual-mean global-mean temperature data with the 50-80 year Fourier band removed. The removed band will be shown in Figure 6.3. The linear trends are recalculated. While the 150-year trend is the same as in (a), trends for the shorter periods are reduced and become more stable. The 80-year low pass data (in dark green) yields the least monotonic fit to the 161-year adjusted data, showing a slow start of the anthropogenic warming in the first 50 years and almost constant warming trend since 1920. The thin grey curves are the 80-year low pass of 3,000 Monte-Carlo synthetic data, obtained by random resampling in 100-year blocks, that yield the extent of possible errors associated with low pass, end-point effects and linear regression (at 95% confidence level).

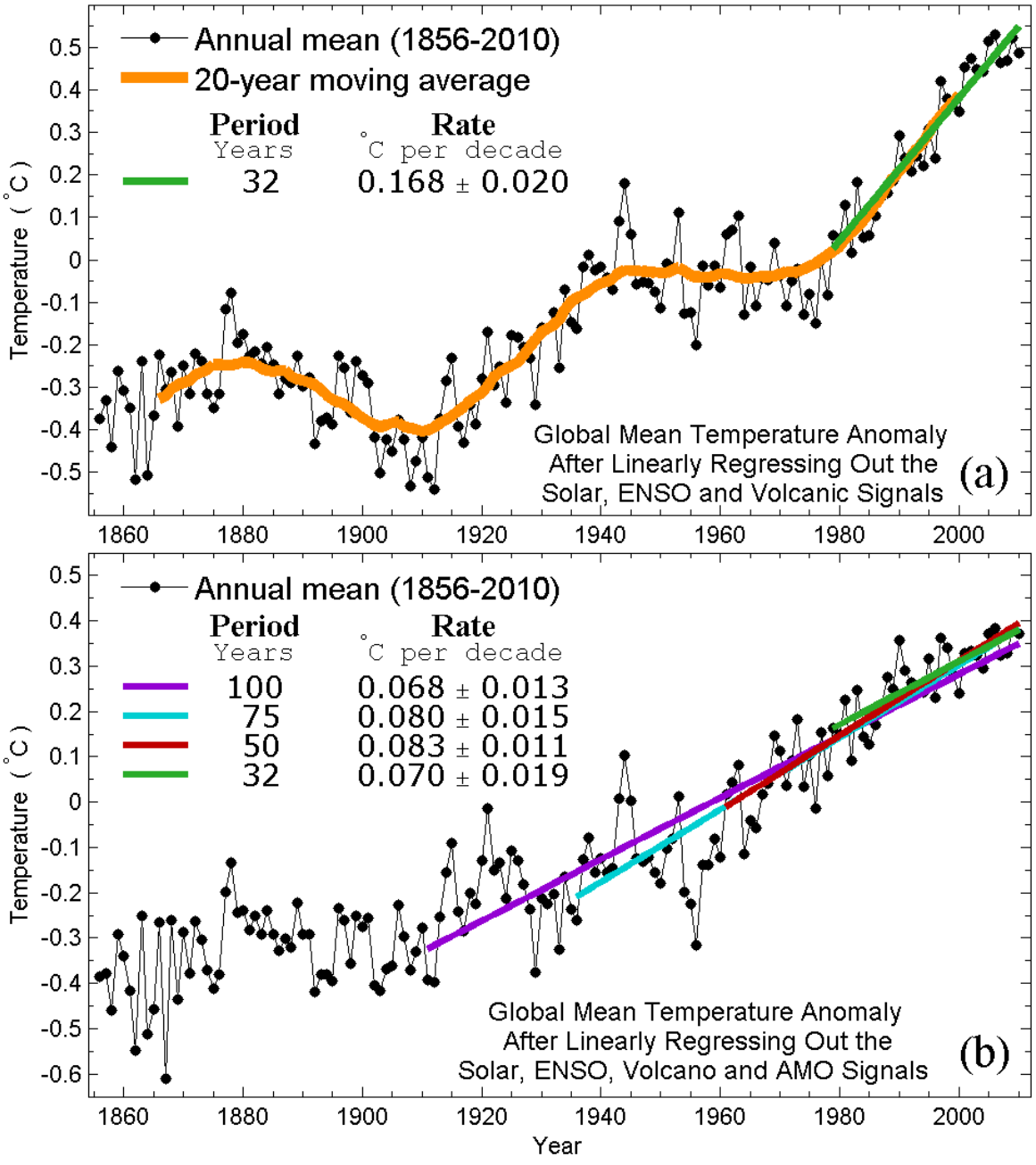


Figure 6.4: **(a)** adjusted data after the removal of ENSO, volcano and solar influence. A running-time mean (in orange) shows a long-period oscillation still present in the residual. **(b)** adjusted data after the removal of ENSO, volcano and solar influence, plus the AMO. The AMO index [30] is shown in Figure 6.5. The correlation between solar secular forcing trend and the global forcing trend is not small, and so there is a concern of collinearity in the explanatory variables in this method, but the influence on the results is negligible.

data. The same trend can also be obtained using Empirical Mode Decomposition [164], by removing centered oscillations until what is left is monotonic, thus providing an additional perspective. Global warming (the Secular Trend in dark green in 6.3b) is seen to begin after the (Second) Industrial Revolution, first slowly and then more rapidly after the end of the $19^{\text{th}}$ century, reaching 0.08°C per decade after 1920. There is little acceleration of the trends after 1920, except for the superimposed 70-year oscillation, which we will now discuss.

### *6.4 Atlantic Multi-decadal Oscillation*

The world's longest instrumental record, CET, started in the Maunder Minimum [28], in the midst of the Little Ice Age (LIA), and covers subsequent episodes of natural and anthropogenic warming. After the invention of thermometers, almost continuous temperature records were kept in Central England. Manley [96] painstakingly compiled most of the monthly CET series starting from 1659. Parker et al. [107] compiled the daily data from 1772. Both are updated to present by the Hadley Centre, Met Office. This record has previously been analyzed to study inter-annual and inter-decadal variability up to the 25-year period [109], inter-annual winter variability and its association with solar forcing [92], and its variance at inter-annual, inter-decadal and 50-years compared with a general circulation model output [71]. CET has a 160-year overlap with the global-mean instrumental record. During this period it can be verified that frequencies lower than 50 years in CET represent global averages after scaling by the factor $\sim 1.4$. This result however has not been confirmed for the period prior to the availability of global data and so CET is used here only in a supporting role, as one of several pieces of evidence that we will present to argue for the likely existence of a 70-year oscillation that has been recurrent in history.

Figure 6.5 shows a comparison of the AMO mode in HadCRUT4 and CET, both obtained using the 50-80-year band pass. The global-mean oscillation varies by about 0.3-0.4°C with a period of about 70 years. The Northern Hemispheric average is slightly larger than the global average. They agree remarkably well with CET after 1850. Prior to 1850, the CET AMO agrees with the multi-proxy data [26] in phase (see Figure 6.6) – the amplitude in the proxy data is only relative. However, the multi-proxy record does not agree with the recent

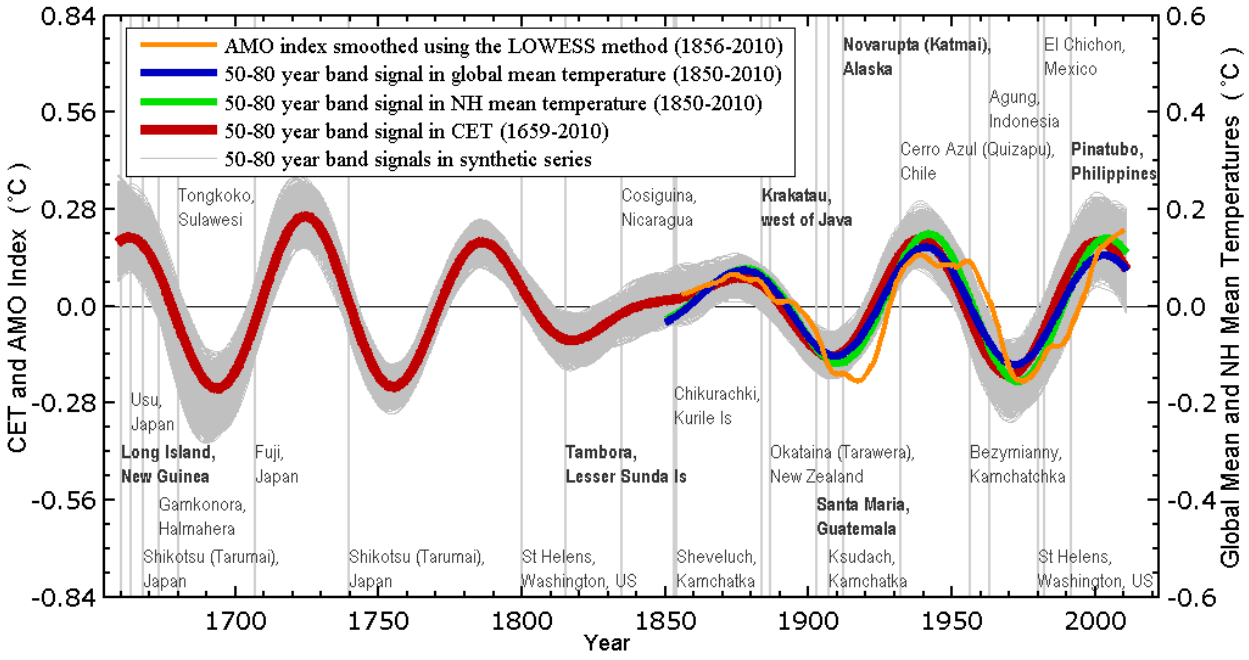


Figure 6.5: The multi-decadal variability (50-80 year band-passed) in CET (in red), in global-mean temperature (in blue) and Northern Hemisphere mean (in green) using HadCRUT4 are compared. The AMO Index [30] smoothed by Locally Weighted Scatterplot Smoothing (LOWESS) [15, 16] is superimposed. The LOWESS is a modified running-time mean and its use allows the mean to extend to the beginning and end of the record. It uses quadratic fit to 25-year subintervals. The light grey curves are the corresponding band pass of 3,000 synthetic data obtained by randomly resampling 200-year blocks of CET data to extend the record before 1659 and after 2010. To show that the troughs of the AMO are not systematically aligned with large volcano eruptions, years of large volcano eruptions (with Volcano Explosivity Index 5 and higher) are indicated in light grey vertical lines. Bold letters are used for the names of major volcano eruptions with VEI of 6 and 7.

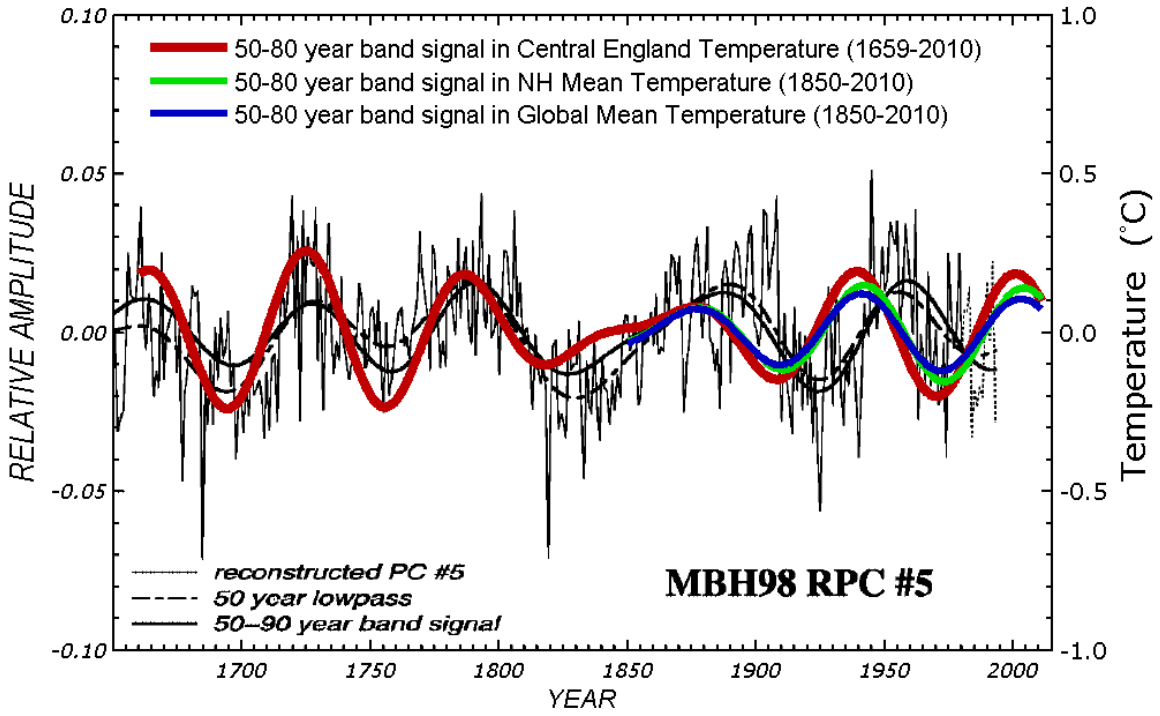


Figure 6.6: The multi-decadal variability (50-80 year band-passed) in CET (in red), in global-mean temperature (in blue) and Northern Hemisphere mean (in green) using HadCRUT4 are superimposed on the multi-proxy data of Delworth and Mann [26] (in black, with line legend shown in the lower left corner in their Figure 3a, with relative amplitude scale of the proxy data on the left axis ). It shows agreement before the availability of global data. After 1900, agreement with proxy data degrades, but proxy data are not needed for comparison with CET when there is instrumental global data.

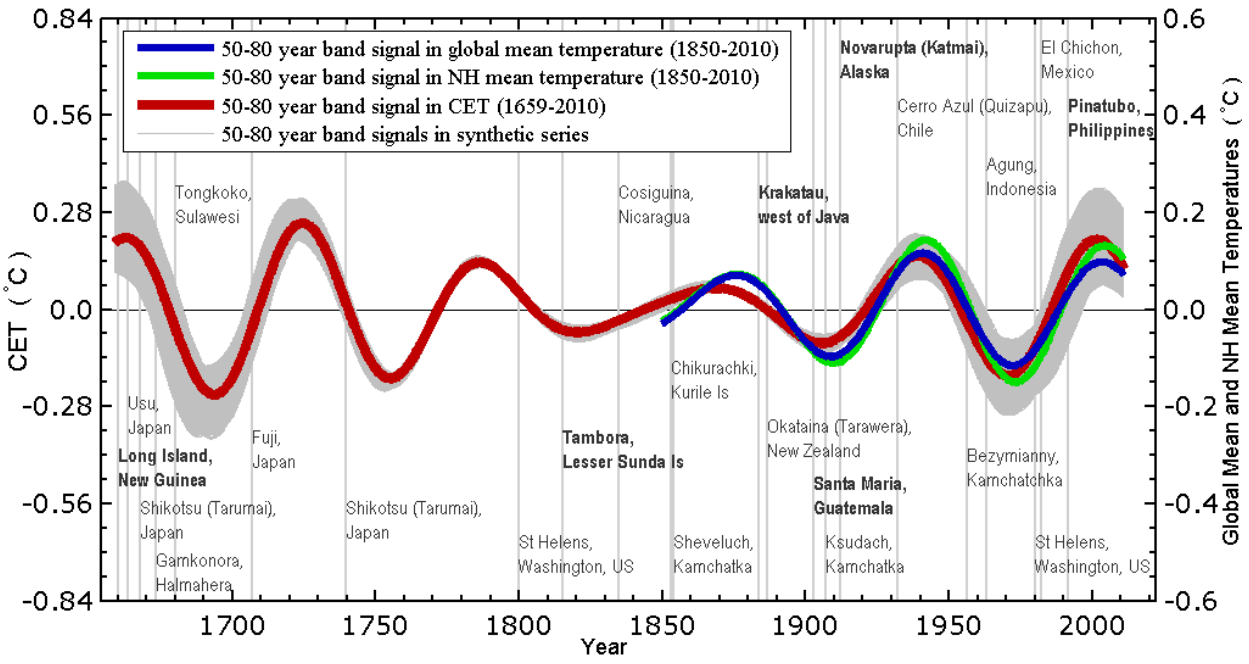


Figure 6.7: Same as Figure 6.5, except that the AMO signals are obtained using a filter based on continuous wavelet transform. The Morlet wavelet with the parameter value $\sigma = 1.4$ is chosen to be the mother wavelet. The light grey curves are again the corresponding band pass of 3,000 synthetic data obtained by randomly resampling 200-year blocks of CET data to extend the record before 1659 and after 2010. The error bars are small in the middle of the time domain which is not surprising, as revealed by the cone of influence in Figure 6.2.

global instrumental record, and therefore is not an appropriate proxy for the AMO. This divergence is known but the cause is unclear [4]. CET, being an instrumental record, does not suffer such a divergence. England, being near the Atlantic coast, is strongly influenced by the Atlantic sea-surface temperature. Indeed, Figure 6.5 shows that CET coincides with the oft-used AMO Index [30], which is defined as the North Atlantic sea-surface temperature average, linearly detrended, over the past 150 years. The wavelet analysis yields very similar AMO signals in CET, global mean and NH mean temperatures, as shown in Figure 6.6. The Morlet wavelet spectrum of the CET time series (Figure 6.2) clearly demonstrates that there exits oscillatory components with periods ranging from 50 to 80 years during the entire period of 1659-2010. This is consistent with the result obtained from Fourier analysis. Although the cone of influence covers a large area for these low-frequency oscillations, the error bars obtained by extending the available temperature record to the past and the future with reasonable synthetic data should be able to account for the edge effects. The error bars shown in Figure 6.6 suggests that the limited data length does not have a crucial effect on the oscillatory nature of the AMO signal.

We therefore argue that CET can be used as a proxy AMO index, extending it back to 1659. The amplitude of the global AMO can also be calibrated against CET AMO, because the two agrees for 160 years after scaling by 1.4.

Using 800 years of ice-core data from Greenland (1880-1975), Broecker [6] in 1975 extracted two recent cycles of an oscillation with a period of 70 years, and correctly predicted the end of cooling in 1970s and the ensuring decades of accelerating warming. Wu et al. [165], using global-mean data, revealed 2.5 cycles with 65-year period. DelSole et al. [24] also found 2.5 cycles by extracting the spatial pattern in the AR4 model control runs that best characterizes internal variability in AR4 control runs, and projecting the observed global data onto this pattern. Wu et al., DelSol et al. and Delworth and Mann [26] related the global-mean oscillation to a spatial pattern emphasizing the North Atlantic, with Pacific and global extensions. The 5 cycles found here argue that this oscillation is natural and recurrent (see also [73]). The phenomenon likely involves thermohaline circulation variability in the Atlantic Ocean [165, 4, 24, 126] ([26, 73]). Recently a 55-80-year AMO has been model simulated as arising from the variability of the meridional overturning circulation

in the Atlantic [154]. Using CET as supporting evidence, we have shown here that these same 2.5 cycles in the global data are a part of a recurrent oscillation going back at least 350 years and therefore unlikely that they can be attributed to volcanic aerosols, whose eruptions were not periodic nor aligned with the troughs (see Figure 6.5).

On the other hand, Wood et al. [160] considered a 200-year record of surface air temperature reconstructed from four stations in the Atlantic-Arctic boundary. The extra years from 1802-1850 coincided with the weaker AMO cycle shown in the middle of our Figure 6.5. Without the benefit of seeing the prior cycles shown in our figure, they conjectured that the Early Twentieth Century Warming is a singular internal variability event without a regular period. Recently there have been debates about the slowing of the warming rates since 2005, with explanations [132, 133, 153] ranging from increases in stratospheric water vapor, background aerosol, to increased coal burning in the emergent economy of China of the last 20 years. If one accepts the conclusion that the AMO is recurrent, and since this period coincides with the start of the descending phase of the AMO, one can suggest that the AMO is a more likely explanation, although it is too early to tell definitively.

### *6.5 Solar and Volcano Contributions to Global Warming*

The solar contribution to the 50-year linear trend is minimal simply because the trend in solar forcing for the past 50 years is minimal [153, 169] (Figure 6.8). The debate on satellite calibration yielding an upward trend [125] in the Total Solar Irradiance (TSI) has now been settled in favor of no trend in the satellite era [41]. Before the satellite era, solar models differ on the degree of irradiance increase. After reviewing these solar models in light of additional evidence on facular areas, Foukal [35] ruled out the possibility of a rapid brightening of the Sun that is required to produce the Early Twentieth Century Warming. We previously [169] found the spatial fingerprint for the combined 11-year solar cycle and secular solar trend after the global warming trend was removed cycle by cycle while allowing solar changes between cycles. The observed temperature was projected onto this pattern to yield a response to both cyclic and secular solar forcing. Thus, without relying on a solar model but relying only on this observationally determined spatial pattern we estimated the solar contribution over the period 1854-2007. For the first half of the twentieth century,

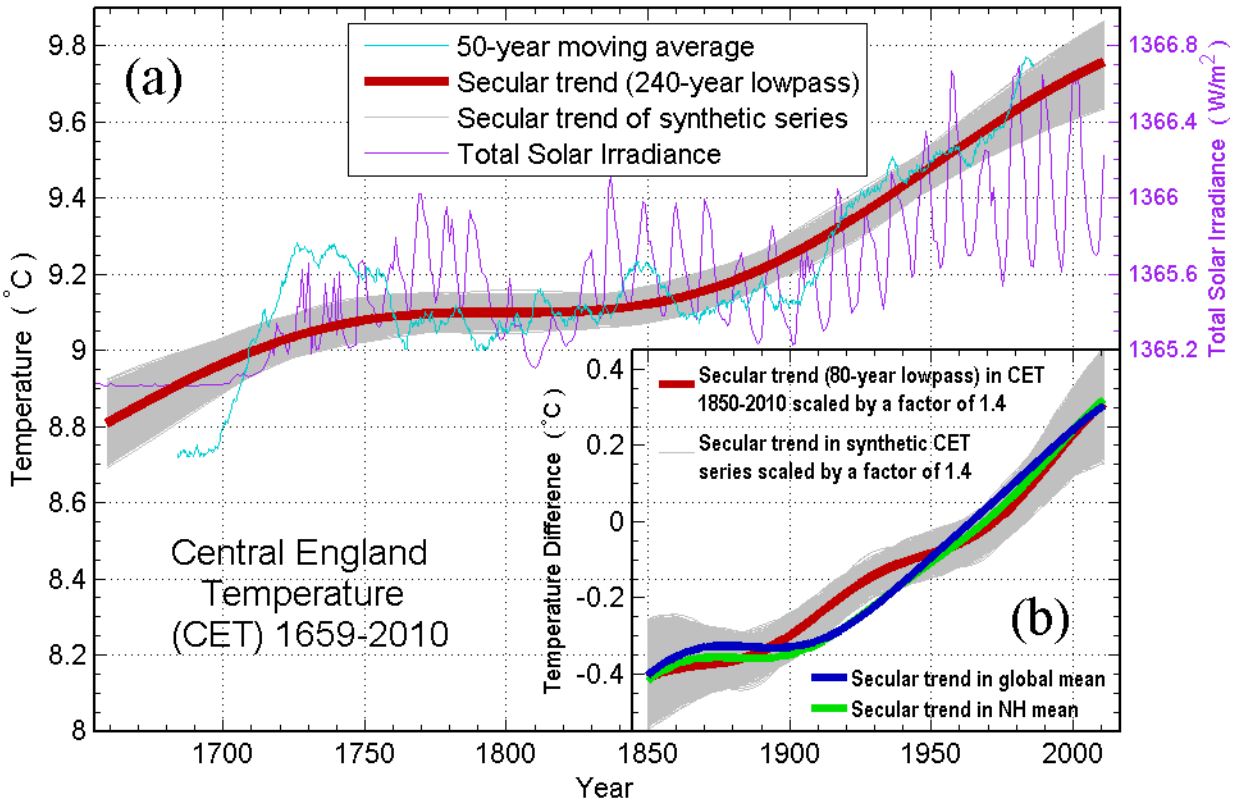


Figure 6.8: Secular Trend is objectively defined as a least monotonic function in a time series. **(a)** For the 352-year Central England Temperature monthly record it is obtained by a 240-year low-pass filter (in red). The 50-year running mean of the raw CET data is shown in cyan. **(b)** The Secular Trend over the 161-year period (1850-2010) is shown (in this case defined by the 80-year low pass), divided by the scaling factor, along with that of the global-mean HadCRUT4 data (in blue) and Northern Hemisphere mean (in green). The solar TSI index, courtesy of Judith Lean, is plotted in purple, with axis on the right. The light grey curves form the error bars of the CET data in the corresponding periods. They are obtained by random resampling 200 (100)-year blocks for the longer (shorter) record.

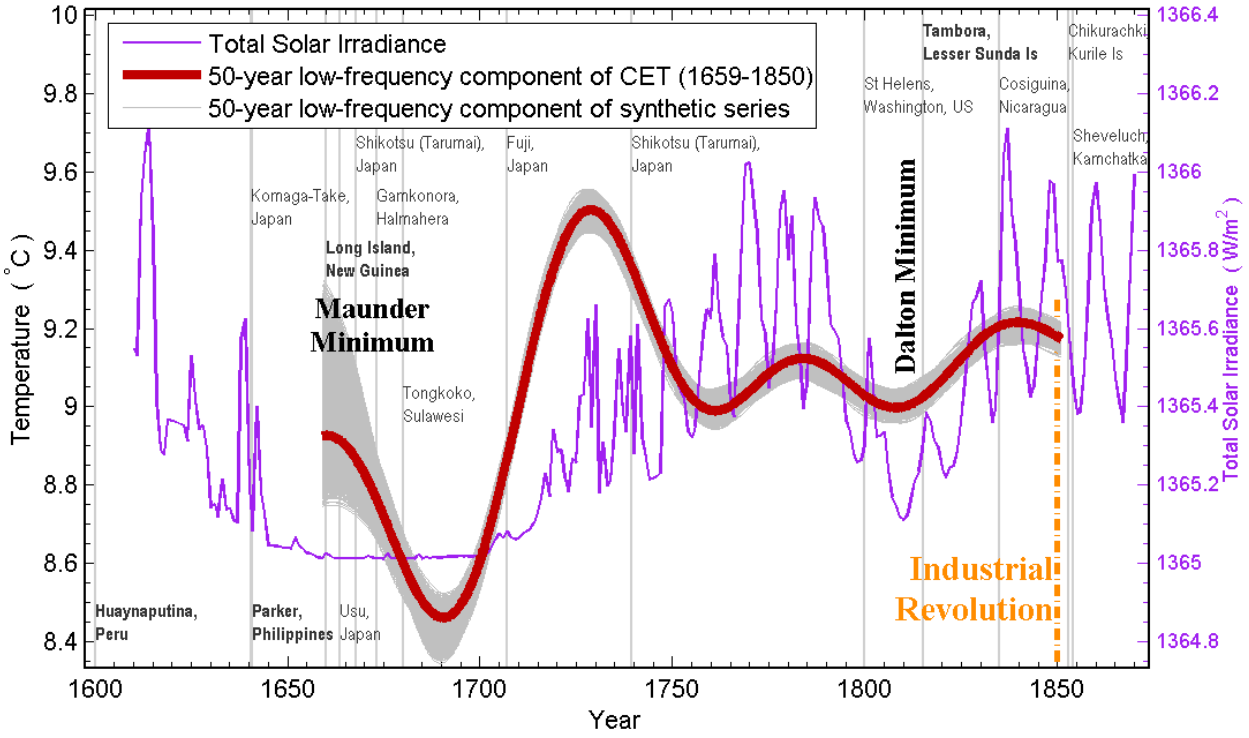


Figure 6.9: The low-frequency portion of the monthly Central England Temperature (the left scale axis) is in red. We propose it as a proxy for the global-mean temperature variation when divided by the scaling factor (not done in this figure). Total Solar Irradiance (TSI) (scale on the right axis), courtesy of Judith Lean, is shown in purple. Fourier and error analyses are done for the entire 352-year record although only the preindustrial portion is shown.

the solar contribution to the linear trend is less than 10%. Our results do not support the much larger role (69%) for the Sun to the observed warming, obtained by Scafetta and West [123] by attributing Early Twentieth Century Warming to solar forcing (see also [3]). Our result is consistent with that of Lean and Rind [86] using multiple regression analysis, which assumed the correctness of Lean's TSI. Arguments on the importance of solar influence on climate often point to the period during the Maunder Minimum [28]: If the low solar activity in the 17$^{\text{th}}$ century caused the Little Ice Age, it should be able to cause significant warming now that the Sun is in its Modern Maximum in the 20$^{\text{th}}$ century. As we will see in Figure 6.9, the timing of the climate events fits volcano aerosols better than solar variability:

Figure 6.8 shows the Secular Trend in the 352-year CET. It reveals a century (1750-1850)

without trend. This pre-industrial period can be interpreted as the background climate separating two *perturbed climates*: the cold of LIA caused probably by volcano aerosols (to be argued), and the modern anthropogenic era of warming. The Secular Trend over the shorter period 1850-2010 is shown in the inset for global-mean HadCRUT4 temperature and CET. They agree in both end-point increments and in linear trends. In the absence of global data, the 50-year low-pass CET record serves to shed some light on the perturbed global climate in LIA (Figure 6.9). Almost no Sunspots were observed during a 70-year interval (1645-1715) called the Maunder Minimum [28]. Large volcano eruptions, Huaynaputina (1600), Parker (1641) and Long Island (1660), contributed to the cold LIA as the CET record began. An unusual series of 5 large volcano eruptions from 1660 to 1680 probably prolonged the cold into the Late Maunder Minimum. A negative phase of the AMO accentuated the cold further in Late Maunder Minimum, reported in Europe [28], although it was thought that the cold CET was only "locally representative" [66]. Our current work argues that it is probably global since the AMO has inphase global manifestations (Figure 6.5). There were no major known volcanoes from 1680 to 1707 (though there were some unknown ones [38]), and it started to warm. Though commonly attributed to the Sun [28], the rapid warming of 0.8°C at the end of Maunder Minimum (after dividing by 1.4) is 8 times larger than our understanding of the solar radiation change [153] can explain, but is within the range of a speculative theory [36], if we remove 0.4°C as due to the AMO. The timing of the warming however appears to precede the increase in TSI [153, 87] by 20-30 years, and favors the reduction of volcano aerosols as the main cause for the warming. The 20-year small dip in temperature near 1810 coincides with the solar Dalton Minimum, but is probably caused by the negative phase in the AMO. The rising AMO cycle in the first half of the $19^{\text{th}}$ century produced a warming, despite the eruption of Tambora (1815), the largest in the past four centuries. The next rising AMO [165, 164, 6, 24] led to the oft-cited Early Twentieth Century Warming in the global mean (1910-1940) of 0.4°C (Figure 6.3a), which happened to occur during a period of increasing mean solar irradiance, leading some to attribute it, incorrectly, to solar forcing [123]. It lies above the range of all model responses to combined anthropogenic and natural forcing compiled by IPCC AR4 [134], even after correcting a discontinuity in the wartime data [139], corroborating the suggestion here that

it is mostly caused by an internal variability. The cooling experienced in the 1970s is seen occurring in the negative phase of the AMO. The period after 1970s shows a secular increase in global-mean temperature. The rising AMO half cycle gives the appearance of an accelerated warming that lasted until 2005.

The anthropogenic greenhouse warming started after the mid-19$^{th}$ century of Industrial Revolution. After a slow start, the smoothed version of the warming trend then stayed almost constant since 1920. Superimposed on the secular trend is a natural multi-decadal oscillation of average period of 70 years of significant amplitude, which could explain many historical episodes of warming and cooling, and account for 40% of the observed warming since mid-20$^{th}$ century, and 50% of the previous attributed anthropogenic warming trend [34]. Prior to the Second Industrial Revolution, the Europe experienced a century (1750-1850) of mild temperature with little trend. This flat inflection serves as reference background for the subsequent anthropogenic warming of about 0.75°C since 1850 and 0.55°C in the last 75 years, and rules out the speculation that the current global warming trend started in the Little Ice Age [1]. The extra length of the CET record provides the perspective for viewing the origin of the modern period of warming, and the recurrent nature of the multi-decadal oscillation.

### *6.6 Chapter Summary*

The observed global warming rate has been non-uniform, and each episode of slowing in the expected warming rate is subjected to intense debates as to its cause. Singular, non-recurrent events have commonly been invoked as explanation for each episode separately. After reviewing evidence in both the latest global data (HardCRUT4) and the longest instrumental record (CET), a substantially revised picture is emerging that gives a consistent attribution for each multi-decadal episode of warming and cooling in recent history. Five cycles of a recurrent multi-decadal oscillation is found to extend to the pre-industrial era in the 352-year CET, and is likely natural. The underlying anthropogenic warming rate is found to be steady since 1920, through the ups and downs that included the Early Twentieth Century Warming, the cooling of the 1960s and 70s, the accelerated warming since 1978 and the recent slowing of the warming rates claimed by some. Quantitatively, the recurrent

multi-decadal internal variability, often underestimated in attribution studies, accounts for 40% of the observed recent 50-year trend. The solar contribution is minimal for the past 50 years. The anthropogenic global warming is 0.08°C per decade, less than half of the previously deduced values. The perspective from a long record not only helps quantifying the contribution from internal variability, especially one with a period so long that it is often confused with secular trends in shorter records, but also confirms that the anthropogenic warming started after the Industrial Revolution, with no warming trend in the century prior to that.

# Chapter 7

# CONCLUSIONS

Using the CMD projection method, we are able to find the temperature response to the decadal solar cycle variation at Earth's surface in various datasets including the modern Reanalysis and in-situ data and the long-term sea surface temperature data dating back to 1854. Bootstrap Monte Carlo $\rho$-tests are conducted to demonstrate that the extracted decadal solar cycle signals are all statistically significant at 95% confidence level. The 14 cycles extracted from the longest global instrumental data record suggests that the decadal signal is solar-induced and unlikely caused by varying volcano aerosols. The observed solar cycle response in temperature indicates that the global mean temperature is warmer when the TSI is higher consistently through the past 150 years. A phase reversal before 1920 as reported in some literature (e.g., [61, chapter 5]) is not seen. Based on a reasonable assumption that the climate response to secular solar forcing and to solar cycle forcing has the same spatial pattern, we estimated the warming trend due to solar variations during the $20^{\text{th}}$ century is approximately $0.01^{\circ}$ C per decade, accounting for only about 15% of the total warming trend in the same period. By examining the largest trend between a particular solar min of 1913 and another solar min of 2005, we conclude that no more than 26% of the observed global warming can be explained by solar forcing since 1880.

The CMD projection method and the associated bootstrap $\rho$-test have been also used to successfully establish the existence of the solar cycle signal in temperature for the entire troposphere. The resulting latitude-height patterns of the solar cycle response show that a "horse shoe" pattern of warming with a horizontal region near the tropical tropopause and two vertical strips at midlatitudes from the surface to the tropopause is present in all seasons while the ozone heating centers in lower stratosphere are located so high above the tropopause and even not seen in some seasons. This discovery may suggest it is the "bottom-up" mechanism [8] similar to the one responsible for the response to greenhouse

forcing that produces a temperature response to increasing solar forcing.

The spatial pattern of the temperature response to radiative heating in the eastern tropical Pacific is also investigated in some detail using a nested multiple linear regression analysis. It is revealed that the anthropogenic forcing warms almost everywhere over the oceans. However, the form of the anthropogenic warming is neither El Niño-like nor La Niña-like. This warm response implies that the Walker circulation weakened in a greenhouse warmed world. The solar response in the eastern tropical Pacific is found to be very weak after the ENSO contamination is removed. This result suggests that solar peaks are not necessarily related to La Niña events.

In addition to the solar response, we have put some effort into understanding the anthropogenic global warming as well. Based on both Fourier and wavelet analyses, a multi-decadal internal variability with a period of about 70 years is found in the global mean temperature record and in the much longer central England temperature starting from the year of 1659, which confirms that it is a natural and recurrent oscillation. Due to its considerable amplitude, this multi-decadal variability could explain many historical episodes of warming and cooling, including the cooling in 1970s and the Early Twentieth Century Warming. Possibility also exists that one might overestimate the rate of the anthropogenic global warming in a short record by mistakenly treating this oscillation as a trend. After removing the long-term multi-decadal oscillation, the residual global mean surface air temperature data demonstrate steady trends around 0.07-0.08° C per decade for the past 100 years, which is likely the "true" anthropogenic warming rate.

Our studies completed using a variety of analysis techniques are devoted to better understanding the natural and anthropogenic influences on the Earth's climate system.

# BIBLIOGRAPHY

[1] S. Akasofu. On the recovery from the Little Ice Age. *Natural Science*, 2:1211–1224, 2010.

[2] R. J. Allen. ENSO and climate variability in the past 150 years. *El Niño and the Southern Oscillation: Multiscale Variability and Global and Regional Impacts*, H. F. Diaz and V. Markgraf, Eds., Cambridge University Press:1–55, 2000.

[3] R. E. Benestad and G. A. Schmidt. Solar trends and global warming. *J. Geophys. Res.*, 114, 2009.

[4] K. R. Briffa, F. H. Schweingruber, J. D. Jones, T. J. Osborn, S. G. Shiyatov, and E. A. Vaganov. Reduced sensitivity of recent tree-growth to temperature at high northern latitudes. *Nature*, 391:678–682, 1998.

[5] P. J. Brockwell and R. A. Davis. *Introduction to Time Series and Forecasting.* Springer, second edition, 2002.

[6] W. S. Broecker. Are we on the brink of a pronounced global warming? *Science*, 189:460–463, 1975.

[7] P. Brohan, J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones. Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *J. Geophys. Res.*, 111:D12106, 2006.

[8] M. Cai and K. K. Tung. Robustness of dynamical feedbacks from radiative forcing: 2versus 2xCO2 experiments in an idealized GCM. *J. Atmos. Sci.*, to appear, 2012.

[9] C. D. Camp and K. K. Tung. The influence of the solar cycle and QBO on the late winter stratospheric polar vortex. *J. Atmos. Sci.*, 64:1267–1283, 2007.

[10] C. D. Camp and K. K. Tung. Stratospheric polar warming by ENSO in winter: A statistical study. *Geophys. Res. Lett.*, 34:L04809, 2007.

[11] C. D. Camp and K. K. Tung. Surface warming by the solar cycle as revealed by the composite mean difference projection. *Geophys. Res. Lett.*, 34:L14703, 2007.

[12] G. Casella and R. L. Berger. *Statistical Inference.* Thomson Learning, second edition, 2002.

[13] H. H. Clayton. The 11-year and 27-day solar periods in meteorology. *Smithson. Misc. Collect.*, 39:1–20, 1940.

[14] A. C. Clement, R. Seager, M. A. Cane, and S. E. Zebiak. An ocean dynamical thermostat. *J. Climate*, 9:21902196, 1996.

[15] W. S. Cleveland. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74:(368): 829836, 1979.

[16] W. S. Cleveland and S. J. Devlin. Locally-weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83:(403): 596610, 1988.

[17] Cochrane and Orcutt. Application of least squares regression to relationships containing autocorrelated error terms. *Journal of the American Statistical Association*, 44:32–61, 1949.

[18] K. T. Coughlin and K. K. Tung. The 11-year solar cycle in the lower stratosphere extracted by the empirical model decomposition method. *Adv. Space Res.*, 34:323–329, 2004.

[19] K. T. Coughlin and K. K. Tung. Eleven-year solar cycle signals throughout the lower atmosphere. *J. Geophys. Res.*, 109:D21105, 2004.

[20] K. T. Coughlin and K. K. Tung. Misleading patterns in correlation maps. *J. Geophys. Res.*, 111:D24102, 2006.

[21] J. A. Curry and P. J. Webster. Climate Science and the Uncertainty Monster. *Bulletin of the American Meteorological Society*, 92:1667–1682, 2011.

[22] I. Daubechies. The wavelet transform time-frequency localization and signal analysis. *IEEE Trans. Inform. Theory*, 36:961–1004, 1990.

[23] D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. DelSol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J. J. Morcrette, B. K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J. N. Thépaut, and F. Vitart. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, 137:553–597, 2011.

[24] T. DelSol, M. K. Tippett, and J. Shukla. A significant component of unforced multidecadal variability in the recent acceleration of global warming. *J. Climate*, 24:909–926, 2011.

[25] T. L. Delworth and T. R. Knutson. Simulation of Early 20th Century Global Warming. *Science*, 287:2246–2250, 2000.

[26] T. L. Delworth and M. E. Mann. Observed and simulated multidecadal variability in the Northern Hemisphere. *Climate Dynamics*, 16:661–676, 2000.

[27] J. Durbin and G. S. Watson. Testing for serial correlation in least squares regression, III. *Biometrika*, 58:1–19, 1971.

[28] J. A. Eddy. Maunder Minimum. *Science*, 192:1189–1202, 1976.

[29] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Chapman and Hall, 1993.

[30] D. B. Enfield, A. M. Mestas-Nunez, and P. J. Trimble. The atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U. S. 28:2077–2080, 2001.

[31] M. Feldstein. Errors in variables: A consistent estimator with smaller MSE in finite samples. 69(348):990–996, 1974.

[32] C. E. Forest, P. H. Stone, and A. P. Sokolov. Estimated pdfs of climate system properties including natural and anthropogenic forcings. *Geophys. Res. Lett.*, 33:L01705, 2006.

[33] P. M. Forster and Coauthors. Changes in atmospheric constituents and in radiative forcing. *Climate Change 2007: The Physical Science Basis*, S.Solomonetal.,Eds.,Cambridge University Press, 2007.

[34] G. Foster and S. Rahmstorf. Global temperature evolution 19792010. *Environmental Research Letters*, 6:044022, 2011.

[35] P. Foukal. A new look at solar irradiance variation. *Solar Physics*, 1, 2012.

[36] P. Foukal, A. Ortiz, and R. Schnerr. Dimming of the 17th Century sun. *The Astrophysical Journal Letters*, 733:L38, 2011.

[37] D. L. Gilman F.J. Fuglister and J. M. Mitchel Jr. On the power spectrum of "red noise". *Journal of Atmospheric Sciences*, 20(2):182–184, 1963.

[38] C. Gao, A. Robock, and C. Ammann. Volcanic forcing of climate over the past 1500 years: An improved ice core-based index for climate models. *J. Geophys. Res.*, 113:D23111, 2008.

[39] H. Gleisner and P. Thejll. Patterns of tropospheric response to solar variability. *Geophys. Res. Lett.*, 30(13):1711, 2003.

[40] L. G. Godfrey. Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica*, 46, 1978.

[41] L. J. Gray, J. Beer, M. Geller, J. D. Haigh, M. Lockwood, K. Matthes, U. Cubasch, D. Fleitmann, G. Harrison, L. Hood, J. Luterbacher, G. A. Meehl, D. Shindell, B. van Geel, and W. White. Solar influences on climate. *Rev. Geophys.*, 48:RG4001, 2010.

[42] L. J. Gray and et al. A review of the influence of solar changes on the earths climate. *Tech. Note 62*, 82 pp., 2005. Met Off. Hadley Cent., Exeter, U. K.

[43] E. Haam and K. K. Tung. Statistics of solar cycle-la niña connection: Correlation of two autocorrelated time series. *J. Atmospheric Sciences*, Submitted, 2012.

[44] J. D. Haigh. A GCM study of climate change in response to the 11-year solar cycle. *Q. J. R. Meteorol. Soc.*, 125:871–892, 1999.

[45] J. D. Haigh. The effects of solar variability on the Earth's climate. *Philos. Trans. R. Soc. London, Ser. A*, 361(1802):95–111, 2003.

[46] J. D. Haigh. The Sun and the Earth's Climate. *Living Rev. Solar Phys.*, 4, 2007.

[47] J. D. Haigh, M. Blackburn, and R. Day. The response of tropospheric circulation to perturbations in lower-stratospheric temperature. *Journal of Climate*, 18:3672–3685, 2005.

[48] P. Handler. Possible association of stratospheric aerosols and el niño type events. *Geophys. Res. Lett.*, 11:11211124, 1984.

[49] J. Hansen and et al. Climate response times: Dependence on climate sensitivity and ocean mixing. *Science*, 229:857859, 1985.

[50] J. Hansen, A. Lacis, D. Rind, G. L. Russell, P. Stone, I. Fung, R. Ruedy, and J. Lerner. Climate sensitivity: analysis of feedback mechanisms. *Geophysical Monograph*, 29:130–163, 1984.

[51] J. Hansen, R. Ruedy, J. Glascoe, and M. Sato. Giss analysis of surface temperature change. *J. Geophys. Res.*, 104:30,99731,022, 1999.

[52] J. Hansen, M. Sato, A. Lacis, R. Ruedy, I. Tegen, and E. Matthews. Climate forcings in the industrial era. *Proc. Natl. Acad. Sci. USA*, 95:1275312758, 1998.

[53] J. Hansen, M. Sato, R. Ruedy, P. Kharecha, A. Lacis, R. Miller, L. Nazarenko, K. Lo, G. A. Schmidt, G. Russell, I. Aleinov, S. Bauer, E. Baum, B. Cairns, V. Canuto, M. Chandler, Y. Cheng, A. Cohen, A. Del Genio, G. Faluvegi, E. Fleming, A. Friend, T. Hall, C. Jackman, J. Jonas, M. Kelley, N. Y. Kiang, D. Koch, G. Labow, J. Lerner, S. Menon, T. Novakov, V. Oinas, Ja. Perlwitz, Ju. Perlwitz, D. Rind, A. Romanou, R. Schmunk, D. Shindell, P. Stone, S. Sun, D. Streets, N. Tausnev, D. Thresher, N. Unger, M. Yao, and S. Zhang. Climate simulations for 1880-2003 with GISS modelE. *Climate Dyn.*, 29:661–696, 2007.

[54] J. Hansen, M. Sato, R. Ruedy, L. Nazarenko, A. Lacis, G. A. Schmidt, G. Russell, I. Aleinov, M. Bauer, S. Bauer, N. Bell, B. Cairns, V. Canuto, M. Chandler, Y. Cheng, A. Del Genio, G. Faluvegi, E. Fleming, A. Friend, T. Hall, C. Jackman, M. Kelley, N. Kiang, D. Koch, J. Lean, J. Lerner, K. Lo, S. Menon, R. Miller, P. Minnis, T. Novakov, V. Oinas, Ja. Perlwitz, Ju. Perlwitz, D. Rind, A. Romanou, D. Shindell, P. Stone, S. Sun, N. Tausnev, D. Thresher, B. Wielicki, T. Wong, M. Yao, , and S. Zhang. Efficacy of climate forcings. *J. Geophys. Res.*, 110:D18104, 2005.

[55] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion.* SIAM Monogr. on Mathematical Modeling and Computation, No. 4, Society for Industrial and Applied Mathematics, 1997.

[56] G. C. Hegerl and Co-authors. *in Climate Change 2007: The Physical Science Basis.* Cambridge University Press, s. solomon, ed. edition, 2007.

[57] G. C. Hegerl, P. Stott, S. Solomon, and F. Zwiers. Comment on "Climate Science and the Uncertainty Monster", j. a. curry and p. j. webster. *Bulletin of the American Meteorological Society*, 92:1683–1685, 2011.

[58] G.C. Hegerl, T.J. Crowley, W.T. Hyde, and D.J. Frame. Constraints on climate sensitivity from temperature reconstructions of the past seven centuries. *Nature*, 440:1029–1032, 2006.

[59] I. M. Held and B. J. Soden. Robust responses of the hydrological cycle to global warming. *J. Climate*, 19:56865699, 2006.

[60] J. R. Holton and H. C. Tan. The Quasi-Biennial Oscillation in the Northern Hemisphere lower stratosphere. *Journal of the Meteorological Society of Japan*, 60:140–148, 1982.

[61] D. V. Hoyt and K. H. Schatten. *The Role of the Sun in Climate Change.* Oxford University Press, 1997.

[62] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the Hilbert spectrum

for nonlinear and non-stationary time series analysis. *Proc. R. Soc. London, Ser. A*, 454:903–995, 1998.

[63] N. E. Huang and Z. Wu. A review on Hilbert-Huang transform: method and its applications to geophysical studies. *Reviews of Geophysics*, 46:RG2006, 2008.

[64] M. Huber and R. Knutti. Anthropogenic and natural warming inferred from changes in earth's energy balance. *Nature Geoscience*, 5:3136, 2011.

[65] R. Jenne. Global observations for reanalysis, 1948-on. In *Proc. Second WCRP Int. Conf. on Reanalyses*, pages 5–9. WCRP Series Rep. 109, Waterfield Park, Reading, United Kingdom, World Meteorological Organization, 1999.

[66] P. D. Jones and Co-authors. High-resolution palaeoclimatology of the last millennium: a review of current status and future prospects. *The Holocene*, 19:3–49, 2009.

[67] P. D. Jones and M. Hulme. *in Climates of the British Isles: Present, Past and Future.* (Routledge, London, 1997), m. a. b. hulme, e., ed. edition, 1997.

[68] D. J. Shea K. E. Trenberth. Atlantic hurricanes and natural varibility in 2005. *Geophys. Research. Lett.*, 33:L12704, 2006.

[69] E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, R. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, Roy Jenne, and Dennis Joseph. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.*, 77:437–471, 1996.

[70] K. B. Karnauskas, R. Seager, A. Kaplan, Y. Kushnir, and M. A. Cane. Observed strengthening of the zonal sea surface temperature gradient across the equatorial Pacific Ocean. *J. Climate*, 22:43164321, 2009.

[71] D. Karoly and P. A. Stott. Anthropogenic warming of central England temperature. *Atmospheric Science Letters*, 7:81–85, 2006.

[72] R. K. Kaufmann, H. Kauppi, M. L. Mann, and J. H. Stock. Reconciling anthropogenic climate change with observed temperature 1998-2008. *Proc. Natl. Acad. Sci., USA*, 108:11790–11793, 2011.

[73] J. R. Knight, R. J. Allan, C. K. Folland, and M. Vellinga. A signature of persistent natural thermohaline circulation cycles in observed climate. *Geophys. Research. Lett.*, 32:L20708, 2005.

[74] G. Kopp and J. Lean. A new, lower value of Total Solar Irradiance: evidence and climate significance. *Geophys. Res. Letters Frontier article*, 38:L01706, 2011.

[75] K. Labitzke. The solar signal of the 11-year sunspot cycle in the stratosphere: Differences between the northern and southern summers. *J. Meteor. Soc., Japan*, 80:963–971, 2002.

[76] K. Labitzke, J. Austin, N. Butchart, J. Knight, M. Takahashi, M. Nakamoto, T. Nagashima, J. Haigh, and V. Williams. The global signal of the 11-year solar cycle in the stratosphere: Observations and models. *J. Atmos. Sol. Terr. Phys.*, 64:203–210, 2002.

[77] K. Labitzke and H. van Loon. Associations between the 11-year solar-cycle, the QBO and the atmosphere .1. the troposphere and stratosphere in the northern hemisphere in winter. *J. Atmos. Terr. Phys.*, 50:197–206, 1988.

[78] K. Labitzke and H. van Loon. On the association between the QBO and the extratropical stratosphere. *Journal of Atmospheric and Terrestrial Physics*, 54:1453–1463, 1992.

[79] S. N. Lahiri. *Resampling Methods for Dependent Data*. Springer, 2003.

[80] A. Larkin and et al. The effect of solar uv irradiance variations on the earths atmosphere. *Space Sci. Rev.*, 94:199214, 2000.

[81] J. Lean. Solar ultraviolet irradiance variations: A review. *J. Geophys. Res.*, 92:839–868, 1987.

[82] J. Lean. Variations in the Sun's radiative output. *Rev. Geophys.*, 29:505–535, 1991.

[83] J. Lean. Living with a variable sun. *Phys. Today*, 58:3238, 2005.

[84] J. Lean, J. Beer, and R. Bradley. Reconstruction of solar irradiance since 1610: Implications for climate change. *Geophys. Res. Lett.*, 22:31953198, 1995.

[85] J. Lean and D. Rind. Climate forcing by changing solar radiation. *J. Climate*, 11:30693094, 1998.

[86] J. Lean and D. Rind. How natural and anthropogenic influences alter global and regional surface temperatures: 1889-2006. *Geophys. Res. Lett.*, 35:L18701, 2008.

[87] J. Lean, G. J. Rottman, J. Harder, and G. Kopp. SORCE contributions to new understanding of global change and solar variability. *Sol. Phys.*, 230:27–53, 2005.

[88] J. L. Lean, G. J. Rottman, H. L. Kyle, T. N. Woods, J. R. Hickey, and L. C. Puga. Detection and parameterization of variations in solar mid- and near-ultraviolet radiation (200400 nm). *J. Geophys. Res.*, 102:29,93929,956, 1997.

[89] C. Leger, D. N. Politis, and J. P. Romano. Bootstrap technology and applications. *Technometrics*, 34:378–398, 1992.

[90] S. S. Leroy, J. G. Anderson, and J. A. Dykema. Testing climate models using gps radio occultation: A sensitivity analysis. *J. Geophys. Res.*, 111:D17105, 2006.

[91] Z. Liu, S. Vavrus, F. He, N. Wen, and Y. Zhong. Rethinking tropical mean response to global warming. *J. Climate*, 18:46844700, 2005.

[92] M. Lockwood, R. G. Harrison, T. Woollings, and S. K. Solanki. Are cold winters in Europe associated with low solar activity? *Environ. Res. Lett.*, 313(5):1–7, 2010.

[93] Smith T. M. and et al. Reconstruction of historical sea surface temperature using empirical orthogonal functions. *J. Clim.*, 9:14031420, 1996.

[94] S. Manabe. Carbon dioxide and climate change. *Adv. Geophys.*, 25:39–82, 1983.

[95] S. Manabe and R. T. Wetherald. Effects of doubling CO2 concentration on climate of a general circulation model. *J. Atmos. Sci.*, 32:3–15, 1975.

[96] G. Manley. Central England Temperatures: monthly means 1659 to 1973. *Quart. J. Roy. Meteor. Soc.*, 100:389–405, 1974.

[97] G. J. Marshall. Trends in the Southern Annular Mode from observations and reanalyses. *J. Clim.*, 16:4134–4143, 2003.

[98] J. Marshall and et al. Mean climate and variability of the atmosphere and ocean on an aquaplanet. *J. Atmos. Sci.*, 64:42704286, 2007.

[99] G. A. Meehl and J. M. Arblaster. A lagged warm event-like response to peaks in solar forcing in the pacific region. *J. Clim.*, 22(13):3647–3660, 2009.

[100] G. A. Meehl, J. M. Arblaster, F. Sassi, K. Matthes, and H. van Loon. Amplifying the Pacific climate system response to a small 11 year solar cycle forcing. *Science*, 325:1114–1118, 2009.

[101] G. A. Meehl and et al. Factors affecting climate sensitivity in global coupled models. *J. Clim.*, 17:15841596, 2004.

[102] C. P. Morce, 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. *J. Geophys. Res.*, 2012.

[103] Y. Naito and I. Hirota. Interannual variability of the northern winter stratospheric circulation related to the QBO and the solar cycle. *J. Meteorol. Soc. Jpn.*, 75:925–937, 1997.

[104] G. R. North and M. J. Stevens. Detecting climate signals in the surface temperature record. *J. Clim.*, 11:563–577, 1998.

[105] G. R. North, J. Wang, and M. G. Genton. Correlation models for temperature fields. *J. Climate*, 24:5850–5862, 2011.

[106] L. Ortlieb. The documented historical record of El Niño events in Peru: An update of the Quinn record (Sixteenth through Ninententh centuries). *El Niño and the Southern Oscillation, Multiscale Variability and Global and Regional Impacts*, H. F. Diaz and V. Markgraf, Eds., Cambridge Univ. Press:207–295, 2000.

[107] D. E. Parker, T. P. Legg, and C. K. Folland. A new daily Central England Temperature series. *International Journal of Climatology*, 12:317–342, 1992.

[108] T. C. Peterson and R. S. Vose. An overview of the global historical climatology network temperature database. *Bull. Am. Meteorol. Soc.*, 78:28372849, 1997.

[109] G. Plaut, M. Ghil, and R. Vautard. Interannual and interdecadal variability in 335 years of Central England Temperatures. *Science*, 268:710713, 1995.

[110] W. H. Quinn. A study of Southern Oscillation-related climatic activity for A.D. 622-1900 incorporating Nile River flood data. *El Niño historical and paleoclimatic aspects of the Southern Oscillation*, H. F. Diaz and V. Markgraf, Eds., Cambridge Univ. Press:119–149, 1992.

[111] W. H. Quinn. The large-scale ENSO event, the El Niño and other important regional features. *Bull. Inst. Fr. Études Andines*, 22:13–34, 1993.

[112] W. H. Quinn and V. T. Neal. The historical record of El Niño events. *Climate since A.D. 1500*, R. S. Bradley nad P. D. Jones, Eds., Routledge:623–648, 1992.

[113] W. H. Quinn, V. T. Neal, and S. E. Antunez de Mayolo. El Niño occurrences over the past four and a half centuries. *J. Geophys. Res.*, 92:14449–14461, 1987.

[114] D. A. Randall and et al. *Climate models and their evaluation, in 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 Univ. Press, Cambridge, U. K., edited by s. solomon et al. edition, 2007.

[115] S. C. B. Raper and et al. The role of climate sensitivity and ocean heat uptake on AOGCM transient temperature response. *J. Clim.*, 15:124130, 2002.

[116] N. A. Rayner and et al. Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: The HadSST2 dataset. *J. Clim.*, 19:446469, 2006.

[117] N. A. Rayner, D. E. Parker, E. B. Norton, C. K. Folland, L. V. Alaxander, D. P. Rowell, E. C. Kent, and A. Kaplan. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, 108:4407, 2003.

[118] R. W. Reynolds and T. M. Smith. Improved global sea surface temperature analyses. *J. Clim.*, 7:929948, 1994.

[119] B. D. Ripley. *Pattern Recognition and Neural Networks.* Cambridge University Press, 1996.

[120] I. Roy and J. D. Haigh. Solar cycle signals in sea level pressure and sea surface temperature. *Atmos. Chem. Phys.*, 10:3147–3153, 2010.

[121] K. Sato, J. E. Hansen, M. P. McCormick, and J. B. Pollack. Stratospheric aerosol optical depths, 1850-1990. *J. Geophys. Res.*, 98:22987–22994, 1993.

[122] N. E. Savin and K. J. White. The Durbin-Watson test for serial correlation with extreme sample sizes or many regressors. *Economitrica*, 45:1989–1996, 1977.

[123] N. Scafetta and B. J. West. Estimated solar contribution in the global mean surface warming using ACRIM TSI satellite composite. *Geophys. Research. Lett.*, 32:L18713, 2005.

[124] N. Scafetta and B. J. West. Phenomenological reconstructions of the solar signature in the Northern Hemisphere, surface temperature records since 1600. *J. Geophys. Res.*, 112:D24S03, 2007.

[125] N. Scafetta and R. Willson. ACRIM-gapand TSI trend issue resolved using a surface magnetic flux TSI proxy model. *Geophys. Res. Lett.*, 36:L05701, 2009.

[126] M. E. Schlesinger and N. Ramankutty. An oscillation in the global climate system of period 65-70 years. *Nature*, 367:723–726, 1994.

[127] T. Schneider and I. M. Held. Discriminants of twentiethcentury changes in earth surface temperatures. *J. Climate*, 14:249–254, 2001.

[128] R. Seager and et al. Mechanisms of hemispherically symmetric climate variability. *J. Clim.*, 16:29602978, 2003.

[129] T. M. Smith and R. W. Reynolds. Extended reconstruction of global sea surface temperatures based on COADS data (18541997). *J. Climate*, 16:14951510, 2003.

[130] T. M. Smith and R. W. Reynolds. Improved extended reconstruction of SST (18541997). *J. Climate*, 17:24662477, 2004.

[131] T. M. Smith, R. W. Reynolds, T. C. Peterson, and J. Lawrimore. Improvements to NOAAs historical merged landocean surface temperature analysis (18802006). *J. Climate*, 21:22832296, 2008.

[132] S. Solomon and Co-authors. Contributions of stratospheric water vapor to decadal changes in the rate of global warming. *Science*, 327:1219–1223, 2010.

[133] S. Solomon and Co-authors. The persistently variable "background" stratospheric aerosol layer and global climate change. *Science*, 333:866–870, 2011.

[134] S. Solomon, D. Qin, M. Manning, M. Marquis, K. Averyt, M. M. B. Tignor, H. L. Miller Jr., and Z. Chen. *Climate Change 2007: The Physical Sciences Basis*. Cambridge University Press, eds., 2007 edition.

[135] M. J. Stevens and G. R. North. Detection of the climate response to the solar cycle. *J. Atmos. Sci.*, 53:25942608, 1996.

[136] D. A. Stone, M. R. Allen, and P. A. Stott. A multimodel update on the detection and sttribution of global surface warming. *J. Climate*, 20:3551–3565, 2007.

[137] P. A. Stott and et al. Observational constraints on past attributable warming and predictions of future global warming. *J. Clim.*, 19:30553069, 2006.

[138] K. L. Swanson, G. Sugihara, and A. A. Tsonis. Long-term natural variability and 20th century climate change. *Proc. Natl. Acad. Sci., USA*, 106:16120–16130, 2009.

[139] D. W. J. Thompson, J. J. Kennedy, J. M. Wallace, and P. D. Jones. A large discontinuity in the mid-twentieth century in observed global mean surface temperature. *Nature*, 453:646–649, 2008.

[140] C. Torrence and G. P. Compo. A practical guide to wavelet analysis. *Bull. Amer. Meteor. Soc.*, 79:61–78, 1998.

[141] Y. M. Tourre, B. Rajagopalan, Y. Kushnir, M. Barlow, and W. B. White. Patterns of coherent decadal and interdecadal climate signals in the Pacific Basin during the 20th century. *Geophys. Res. Lett.*, 28:20692072, 2001.

[142] K. E. Trenberth and Co-authors. *in Climate Change 2007: The Physical Science Basis*. Cambridge University Press, s. solomon, ed. edition, 2007.

[143] K. K. Tung and C. D. Camp. Solar-cycle warming at the Earth's surface in NCEP and ERA-40 data: A linear discriminant analysis. *J. Geophys. Res.*, 113:D05114, 2008.

[144] K. K. Tung and J. Zhou. The Pacific' response to surface heating in 130 years of SST: La Niña-like or El Niño-like? *J. Atmospheric Sciences*, 67:2649–2657, 2010.

[145] K. K. Tung, J. Zhou, and C. D. Camp. Constraining model transient climate response using independent observations of solar-cycle forcing and response. *Geophys. Research Lett.*, 35:L17707, 2008.

[146] S. M. Uppala, P. W. Kallberg, A. J. Simmons, U. Andrae, V. D. Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Holm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P. McNally, J. F. Mahfouf, J. J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen. The ERA-40 re-analysis. *Q.J.R. Meteorol. Soc.*, 131:29613012, 2005.

[147] H. van Loon and G. A. Meehl. The response in the Pacific to the Sun's decadal peaks and contrasts to cold events in the southern oscillation. *J. Atmos. Sol. Terr. Phys.*, 70:1046–1055, 2008.

[148] H. van Loon, G. A. Meehl, and D. J. Shea. Coupled air-sea response to solar forcing in the Pacific region during northern winter. *J. Geophys. Res.*, 112:D02108, 2007.

[149] G. A. Vecchi, A. Clement, and B. J. Soden. Examining the tropical Pacifics response to global warming. *Eos, Trans. Amer. Geophys. Union*, 89:8183, 2008.

[150] G. A. Vecchi and B. J. Soden. Global warming and the weakening of tropical circulation. *J. Climate*, 20:43164340, 2007.

[151] G. A. Vecchi, B. J. Soden, A. T. Wittenberg, and I. M. Held. Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing. *Nature*, 441:7376, 2006.

[152] A. Wald. The fitting of straight lines if both variables are subject to error. *Ann. Math. Statist.*, 11(3):284–300, 1940.

[153] Y.-M. Wang, J. Lean, and N. R. Sheeley Jr. Modeling the Sun's magnetic field and irradiance since 1713. *Astrophys. J.*, 625:522–538, 2005.

[154] W. Wei and Lohmann. Simulated Atlantic Multidecadal Oscillation during the Holocene. *J. Climate*, 2012.

[155] W. B. White, J. Lean, D. R. Cayan, and M. D. Dettinger. Response of global upper ocean temperature to changing solar irradiance. *J. Geophys. Res.*, 102:32553266, 1997.

[156] W. B. White and Y. M. Tourre. Global SST/SLP waves during the 20th century. *Geophys. Res. Lett.*, 30:1651, 2003.

[157] D. S. Wilks. Resampling hypothesis tests for autocorrelated fields. *J. Climate*, 10:65–82, 1997.

[158] D. S. Wilks. *Statistical Methods in the Atmospheric Sciences*. Academic Press, second edition, 2006.

[159] R. C. Willson, S. Gulkis, M. Janssen, H. S. Hudson, and G. A. Chapman. Observation of solar irradiance variability. *Science*, 211:700–702, 1981.

[160] K. R. Wood, J. E. Overland, T. Jonsson, and B. V. Smoliak. Air temperature variations on the Atlantic-Artic boundary since 1802. *Geophys. Research. Lett.*, 2010.

[161] S. D. Woodruff, H. F. Diaz, J. D. Elms, and S. J. Worley. COADS Release 2 data and metadata enhancements for improvements of marine surface flux fields. *Phys. Chem. Earth*, 23:517526, 1998.

[162] Z. Wu and N. E. Huang. A study of the characteristics of white noise using the empirical mode decomposition method. *Proc. Roy. Soc. London*, 460A:1597–1611, 2004.

[163] Z. Wu and N. E. Huang. Ensemble Empirical Mode Decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 1(1):1–41, 2009.

[164] Z. Wu, N. E. Huang, S. R. Long, and C. K. Peng. On the trend, detrending and variability of nonlinear and non-stationary time series. *Proc. Natl. Acad. Sci., USA*, 104:14889–14894, 2007.

[165] Z. Wu, N. E. Huang, J. M. Wallace, B. Smoliak, and X. Chen. On the time-varying trend in global-mean surface temperature. *Climate Dynamics*, 2011.

[166] Z. Wu, E. K. Schneider, B. P. Kirtman, E. S. Sarachik, N. E. Huang, and C. J. Tucker. The modulated annual cycle: an alternative reference frame for climate anomalies. *Climate Dynamics*, 31:823–841, 2008.

[167] S.-P. Xie, C. Deser, G. A. Vecchi, J. Ma, H. Teng, and A. T. Wittenberg. Global warming pattern formation: Sea surface temperature and rainfall. *J. Climate*, 23:966986, 2010.

[168] J. Zhou and K. K. Tung. On the CMD Projection method and the associated statistical tests in climate data analysis. *Advances in Adaptive Data Analysis*, 2010.

[169] J. Zhou and K. K. Tung. Solar cycle in 150 years of global sea surface temperature data. *J. Climate*, 23:3234–3248, 2010.

[170] F. W. Zwiers. Statistical considerations for climate experiments. Part II: Multivariate tests. *J. Climate Appl. Meteor.*, 26:477487, 1987.

[171] F. W. Zwiers. The effect of serial correlation on statistical inferences made with resampling procedures. *J. Climate*, 3:14521461, 1990.

Appendix A

# SPATIAL PATTERNS OF THE DECADAL SOLAR CYCLE RESPONSE IN VARIOUS SURFACE TEMPERATURE DATA RECORDS

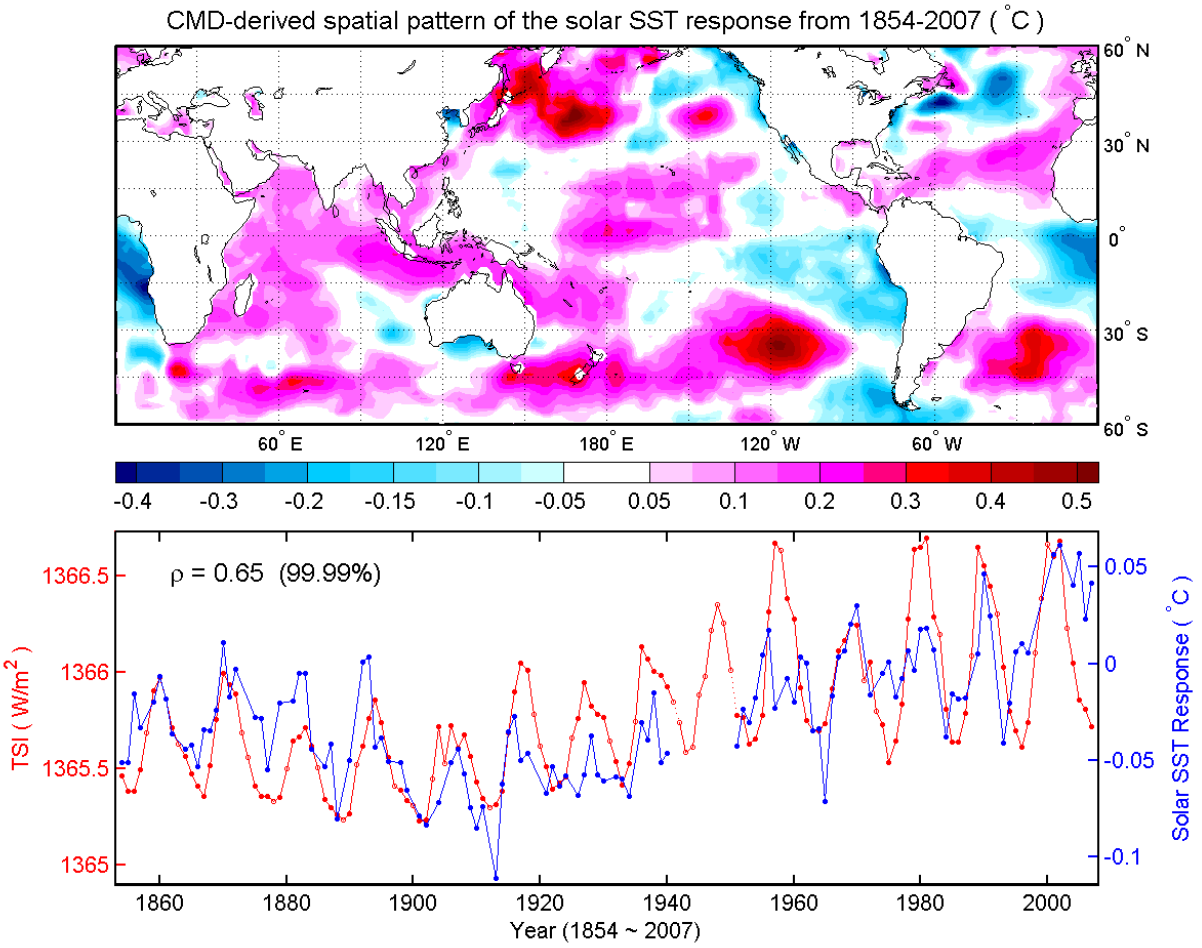


Figure A.1: Reproduction of Figure 3.2 using the same color scale as the one for the solar response in Figure 5.1. This change is made to facilitate the comparison of the solar responses in different datasets obtained using different methods (Composite Mean Difference and Multiple Linear Regression). This CMD-derived spatial pattern is normalized by the composite mean difference value of the TSI time series between the solar max group and the solar min group so that it has the same dimension as the one obtained using multiple linear regression.

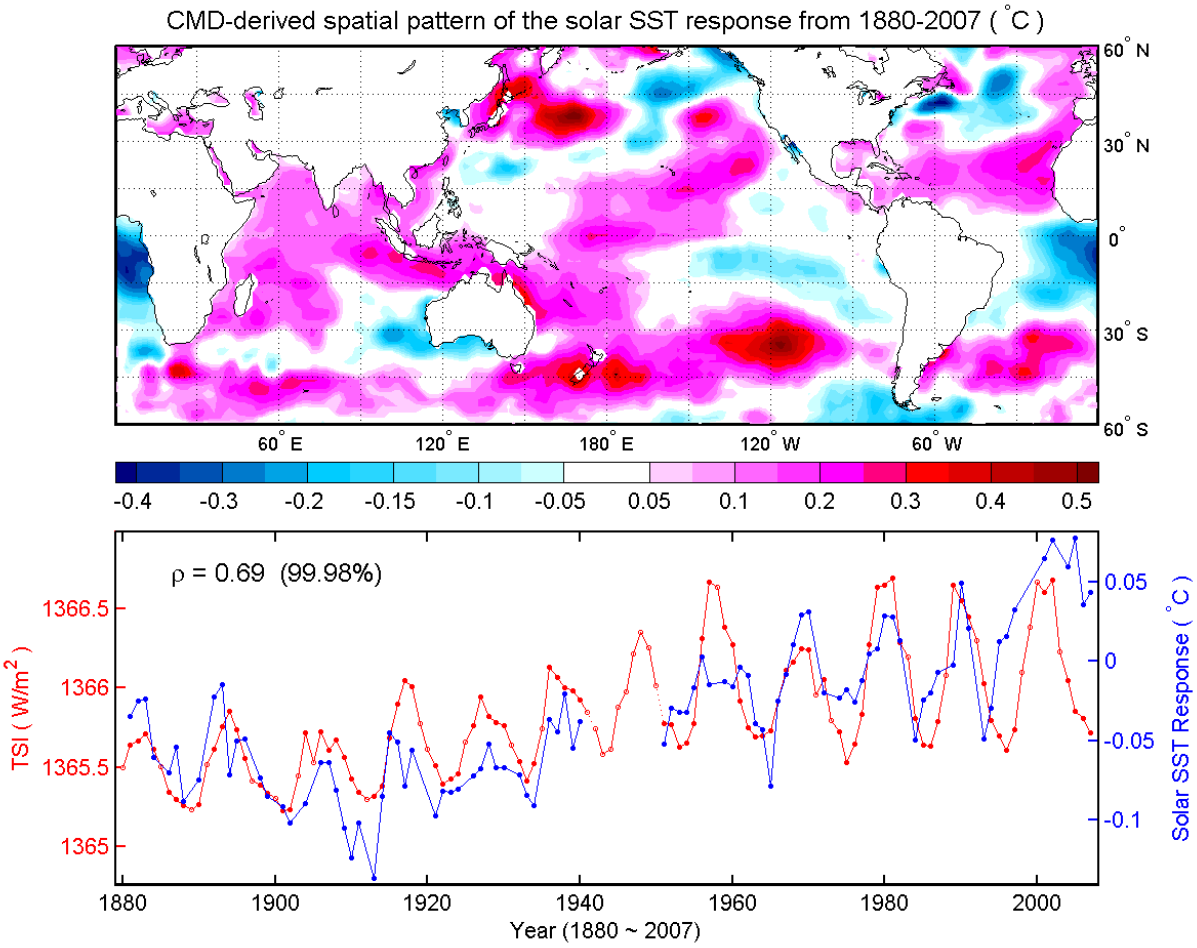


Figure A.2: Reproduction of Figure 3.3 using the same color scale as the one for the solar response in Figure 5.1. This CMD-derived spatial pattern is also normalized by the CMD value of the TSI time series between the solar max group and the solar min group.

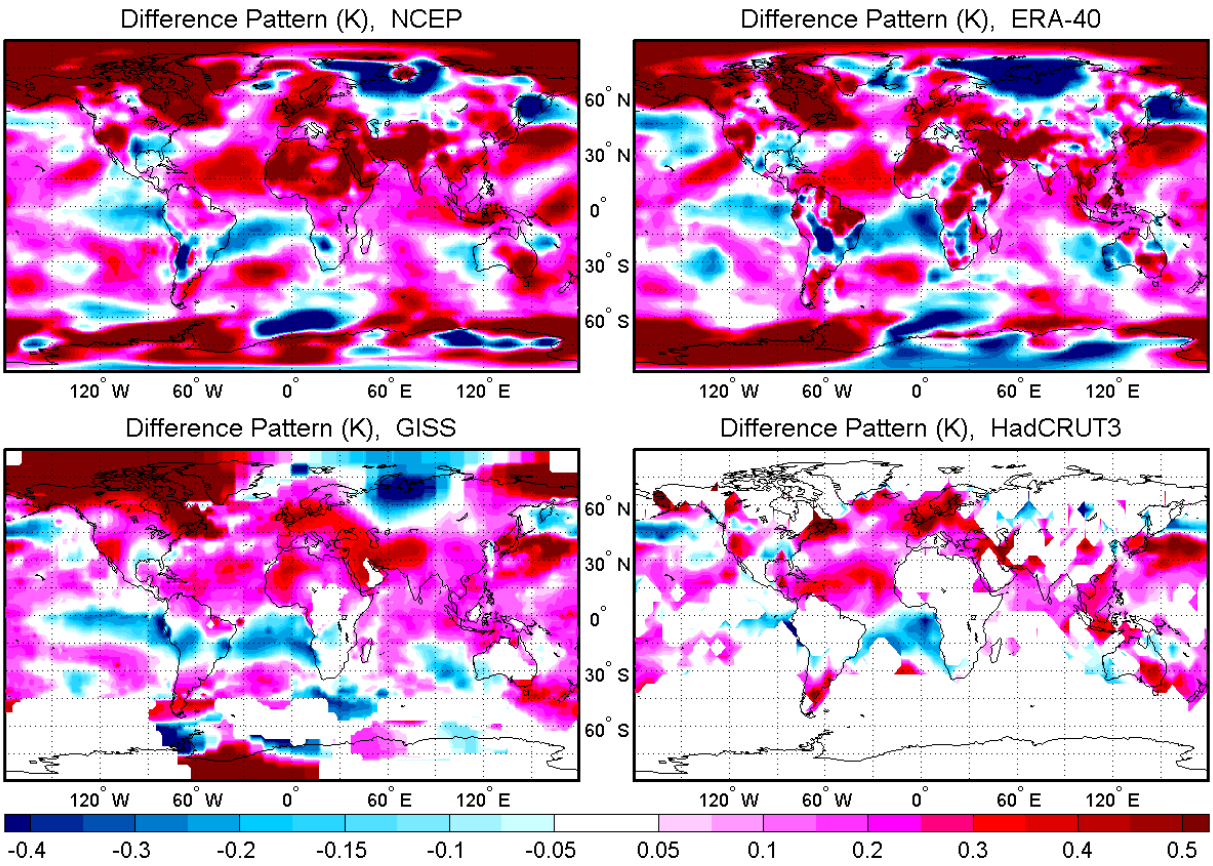


Figure A.3: Reproduction of Figure 3.7 using the same color scale as the one for the solar response in Figure 5.1. Thess CMD-derived spatial patterns are also normalized by the CMD value of the TSI time series between the solar max group and the solar min group.

# VITA

**Jiansong Zhou**

| | | | |
|---|---|---|---|
| • Ph.D. | Applied Mathematics | University of Washington | 2012 |
| • M.S. | Applied Mathematics | University of Washington | 2008 |
| • M.S. | Computational Mathematics | Peking University, Beijing, China | 2004 |
| • B.S. | Computational Mathematics | Peking University, Beijing, China | 2001 |