Identifying Anthropogenic Causes of the Observed Twentieth Century Surface Temperature Change: Frequentist and Bayesian Approach

by

Chun Kit (Terry) Lee
B.Sc., Simon Fraser University, 2001

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics and Statistics

We accept this thesis as conforming
to the required standard

[redacted]

Dr. M. Tsao, Supervisor (Department of Mathematics and Statistics)

[redacted]

Dr. F.W. Zwiers, Supervisor (Canadian Centre for Climate Modelling and Analysis)

[redacted]

Dr. A.J. Weaver, Outside Member (Department of Earth and Ocean Sciences)

[redacted]

Dr. V.V. Kharin, External Examiner (Canadian Centre for Climate Modelling and Analysis)

Supervisor: M. Tsao and F.W. Zwiers

ABSTRACT

It has been hypothesized that the recent warming trends in the climate system are caused by the increase in concentration of greenhouse gases and anthropogenic sulfate aerosols. The 'optimal fingerprinting' methodology is generally use to quantify the extent to which of forcing agents affect the climate. Recently, interest in applying a Bayesian methodology to the problem has emerged. Compared to the classical 'optimal fingerprinting' approach, the Bayesian approach allows the incorporation of prior beliefs of climate change into the analysis and allows direct probability assessments on the hypotheses of climate change detection and attribution. These probability assessments are more quantitative compared to the classical significance level and p-value. We analyze the output from CGCM1, CGCM2 and HadCM2 climate model with the HadCRUTv observational dataset using both the 'optimal fingerprinting' and Bayesian approaches, to identify which forcing agents are likely to have contributed to the temperature change in the twentieth century. The results from both approaches are very similar. With the Bayesian methodology, we find that the combined effect of greenhouse gases and sulfate aerosols may have contributed to the temperature change in the early twentieth century. Our analysis also suggests that the warming trends in the late twentieth century are highly unlikely to be due to natural climate variations alone. We also conduct robust analysis on the choice of prior used in Bayesian analysis to demonstrate that our results are insensitive to the choice of prior.

Examiners:

[redacted]

Dr. M. Tsao, Supervisor (Department of Mathematics and Statistics)

[redacted]

Dr. F.W. Zwiers, Supervisor (Canadian Centre for Climate Modelling and Analysis)

[redacted]

Dr. A.J. Weaver, Outside Member (Department of Earth and Ocean Sciences)

[redacted]

Dr. V.V. Kharin, External Examiner (Canadian Centre for Climate Modelling and Analysis)

Contents

List of Tables

List of Figures

1 Introduction

In recent years, people are more concerned with the global warming phenomenon. One of the hypothesized anthropogenic causes of global warming is the increase in concentration of greenhouse gases, such as CO_2 and CFCs, since the beginning of industrialization. Increase in concentration of greenhouse gases will reduce Earth's efficiency to radiate heat back to the space. Not only that, many greenhouse gases tend to stay in the atmosphere for hundreds of year after being emitted and therefore result in a long-term effect on the climate system. On the other hand, aerosols (microscopic airborne particles) could reflect incoming solar radiation back to space and hence result in a cooling effect for the Earth. Compared to greenhouse gases, tropospheric aerosols are relatively short-lived. Anthropogenic aerosols are largely derived from fossil fuel and biomass burning. Natural external forcing factors are also thought to play a role in explaining the global warming phenomenon. For example, the sun's energy output has varied by small amounts (about 0.1%) over the past two 11-year solar cycles (IPCC 2001). Since the fundamental source of all energy in the Earth's climate system is radiation from the sun, variation in solar output will thereby change the climate system. Also, volcanic activity can produce large amounts of sulphur containing gas, which is then transformed into sulphate aerosols. These aerosols could have a cooling effect that is detectable for a few years following major eruptions that inject large amounts of sulphate into the stratosphere. However, scientific knowledge on the effects due to aerosols and solar output variation remains highly uncertain.

Instrumental record of global surface temperature since 1861 indicated that the average global surface temperature has increased by 0.6 ± 0.2^{o}C since the late 19th century (IPCC 2001).

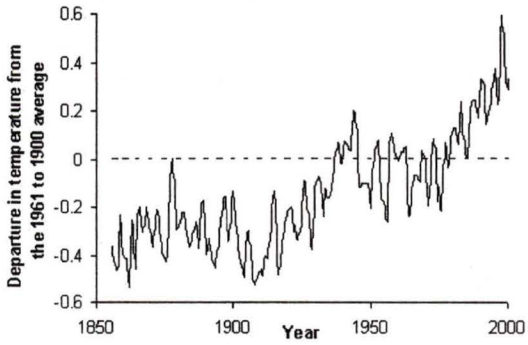

Figure 1: Global annual temperature anomalies from the TaveGL dataset

Figure 1 shows the annual global temperature anomalies relative to 1961 to 1990 for the TaveGL dataset (available on www.cru.uea.ac.uk). In general, anomalies can be calculated by subtracting the data from the average of the whole or part of the data. The anomalies shown in Figure 1 are calculated by subtracting the observed annual average from the average for 1961 to 1990.

As stated by the IPCC (2001), it is very likely that the 1990s was the warmest decade in the instrumental record since 1861. Researches on climate change have suggested that such an increase in temperature is unlikely to be due to natural variation of the climate system alone, but rather a combination of natural variation and external forcing to the climate system. A fair amount of research has been done to detect and attribute the causes of the changes in the climate system. The detection part of climate change refers to the process of demonstrating that the observed change is significantly different from what we can expect with natural variability alone. However, the detection of a change in the climate does not necessarily imply that its causes are well understood. The attribution aspect of the problem refers to the process of establishing cause and effect relation between the observed change and forcing agents. The attribution of climate change to external forcing involves statistical

analysis and the assessment of multiple lines of evidence to show that the observed changes are i) consistent with the anticipated responses to a combination of natural and external forcing and ii) inconsistent with alternative, physically-plausible explanations.

The ability to detect and attribute the historical variation in climate system will add confidence to projections of future climate change. Since the presence of natural climate variability in the climate system, the detection and attribution of climate change is a statistical signal-in-noise problem. The signal part comes from external forcing and the noise part comes from natural climate variation.

Detection and attribution studies require the use of sophisticated climate models. A climate model is a simplified mathematical representation of the Earth's climate system in three dimensions. Based on the law of physics, climate models have been developed to predict a number of environmental factors, for example, temperature at Earth's surface, circulation of ocean and wind currents and development of cloud cover. The models that will be used in this thesis are the first and second generation coupled ocean-atmosphere general circulation models (CGCMs) from the Canadian Centre for Climate Modelling and Analysis (CCCma) (CGCM1, Flato et al., 2000; CGCM2, Flato & Boer, 2001) and second generation CGCM from Hadley Center (HadCM2, Johns et al., 1997). More details about these models can be found on the CCCma website (www.cccma.bc.ec.gc.ca) and the Climate Research Unit website (www.cru.uea.ac.uk). These models are developed to simulate the full physical climate system and they contain four key components: atmosphere, ocean, sea ice and land surface. Controlled simulations with the model (also called control runs) are obtained by fixing the forcing agents at some pre-specify level. On the other hand, forced simulations (forced runs) are obtained by changing the level of the forcing agent over time. The forced simulations that I will consider in this thesis included changes in the concentration of the well-mixed green-

house gases as represented by an equivalent concentration of CO_2 (denoted as a G run of the climate model) and changes in anthropogenic sulfate aerosols in addition to greenhouse gases (GS run).

The observed data used in this thesis is the HadCRUTv dataset (Jones et al., 2001), which is also available at www.cru.uea.ac.uk. This dataset is based on a combination of monthly values (from 1870-1999) of land near surface air temperature anomalies and sea surface temperature anomalies relative to 1961-90 and is presented on a 5^o(latitude)$\times 5^o$(longitude) grid. This will lead to $36\times72=2592$ spatial grid points for the entire globe for each month. However, a lot of the grid values are missing due to the fact that no measurements of temperature are recorded for that grid box. Figure 2 shows the distribution of available grid values in the HadCRUTv dataset for January 1900, 1950 and 1995. In most studies, missing data are not filled in and are excluded from the analyses. Attention will be given to avoid bias from systematically missing data (Section 2.3). Depending on the researchers and the type of studies, the data can be aggregated into different sizes of grid boxes and decadal or annual temperature anomalies can be used. Other than surface air temperature, vertical patterns of temperature change (Allen & Tett, 1999), precipitation and surface pressure have also been used in detection and attribution studies.

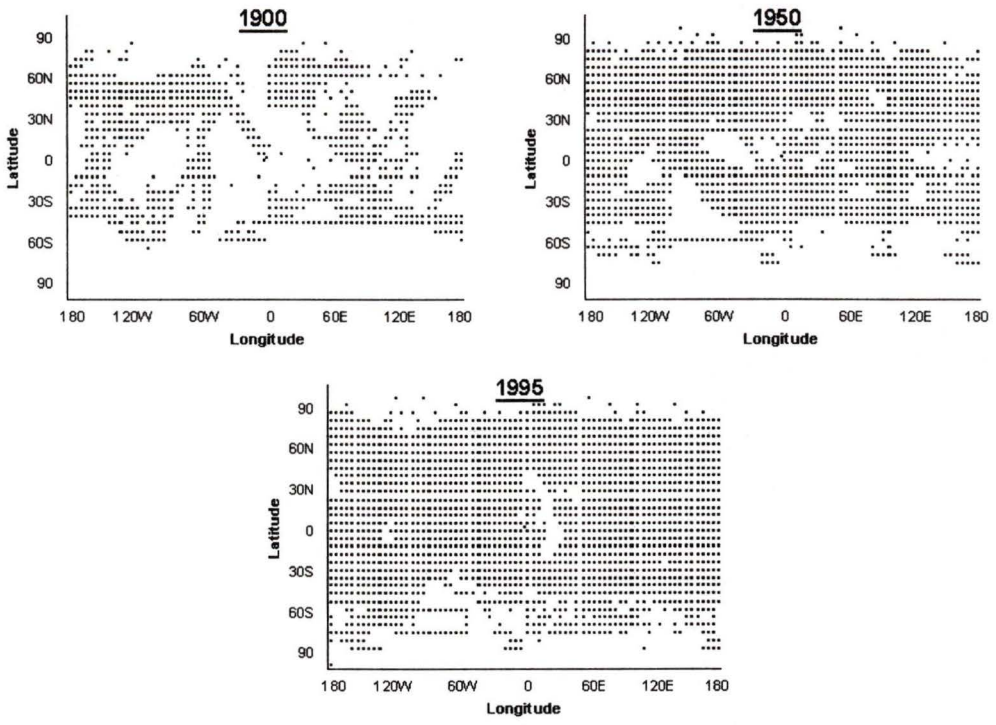

Figure 2: Available grid values for January 1900, 1950 and 1995 in the HadCRUTv dataset

The statistical tool that has been widely used in detection and attribution studies is 'optimal fingerprinting' (see, for example, Hasselmann, 1993, 1997; Hegerl et al., 1997; Allen & Tett, 1999). 'Optimal fingerprinting', in principle, is just general linear least-square regression, in which we estimate the amplitudes that best fit the climate model output to the observed data. In 'optimal fingerprinting', the detection and attribution problem will be cast with the following model

$$\mathbf{y} = \sum_{i=1}^{m} \underline{x}_i \beta_i + \boldsymbol{\varepsilon} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{y}$ is a $n \times 1$ random vector that consists of the observed data, $\mathbf{X} = (\underline{x}_1|...|\underline{x}_m)$ is a $n \times m$ matrix composed of the m signal patterns (also called fingerprints, which is the output obtained from forced runs of a climate model), $\boldsymbol{\beta} = (\beta_1, ..., \beta_m)^T$ is a $m \times 1$ vector composed of the unknown amplitudes and $\boldsymbol{\varepsilon}$ is a $n \times 1$ random vector of natural climate noise. The fields in

$\mathbf{y}$ usually contain anomalies of temperature observations over 30-50 years time scale, arrayed in space and time. The fields are masked so that they represent only those grid boxes with adequate data (this process is usually referred to as observation masking). The dimension of $\mathbf{y}$, n (the total number of observation contained in any one single realization of the field), is large.

Detection of climate change is associated with the rejection of the null hypothesis that the amplitudes ($\boldsymbol{\beta}$) are equal to zero. Next, assuming a detection, attribution is then associated with the failure to reject the null hypothesis that the amplitudes are equal to 1. For attribution, the failure to reject the null only implies that the data simply does not provide enough evidence to reject the null and it does not imply the acceptance of the null. However, the goal of attribution is to determine whether the amplitudes are equal to 1 (i.e. whether to accept the null or not). So, the classical hypothesis testing approach for attribution is not statistically well-posed. Also, the classical approach to hypothesis testing does not allow a direct probability assessment on the true amplitudes. The reason is that the true amplitudes are not considered as random variables in the classical viewpoint.

Aware of the problems with classical approach, Berliner et al. (2000) used a Bayesian approach in combination with linear regression to conduct their study. They obtained the estimated amplitude by using 'optimal fingerprinting' and then, make a detection claim if the posterior probability that the true amplitude lies in a neighborhood around zero (say $\mathcal{D}$) is small. Similarly, attribution is claimed if the posterior probability that the true amplitude lies in a neighborhood around 1 (say $\mathcal{A}$) is large. By defining the detection and attribution problem in this way, they allow direct probability assessment of detection and attribution and avoid the ill-posed attribution problem in the classical approach. Other forms of Bayesian analyses have also been discussed in climate change literature, for example, Hasselmann (1998) and

Leroy (1998). However, these papers only provide incomplete views of the potential of the Bayesian approach and so I will not discuss them in my thesis.

In this thesis, my Bayesian analysis will be based on Berliner et al.'s (2000) formulation with some extensions and modifications. One of these extensions is to include two forcing agents in the analysis. In addition, I will also present results by using Bayesian confidence sets. Moreover, the natural climate variability model will be specified differently from Berliner et al. (2000), in way that is more consistent with other recent optimal detection studies (such as Stott el al., 2001). In Berliner et al. (2000), the natural climate variability is estimated by assuming it to be space-time separable, which may not be a valid assumption. Instead, I will use the output from a climate model to estimate the space-time covariance structure of the natural climate variability (See Chapter 2). Other extensions, such as formulating a Bayesian regression model, will not be discussed in this thesis. I will simply focus on the detection and attribution aspect of the problem.

Other classical methods have also been used, such as centered and uncentered statistic (see, for example, Santer et al. (1995)). The centered statistic method correlates the observed and signal anomalies in space relative to their respective spatial means, whereas the uncentered statistic correlates these fields without removing their spatial means. However, these methods will not be discussed here. In Chapter 2, I will describe in detail on how 'optimal fingerprinting' is done and an example using CGCM1 and CGCM2 will be presented. Bayesian inference of one forcing agent with CGCM1 and CGCM2 will be described in Chapter 3. Chapter 4 will extend the Bayesian inference to two forcing agents with HadCM2. In Chapter 5, I will present some remarks and discussion about my analysis.

2 Classical Approach - 'Optimal fingerprinting'

2.1 Setting up the problem

In 'optimal fingerprinting', the detection and attribution problem will be cast with the following model

$$\mathbf{y} = \sum_{i=1}^{m} \underline{x}_i \beta_i + \boldsymbol{\varepsilon} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2}$$

Description of this model is given in Chapter 1. The assumptions that have been made in a lot of studies are that i) there are no measurement errors in $\mathbf{y}$, and ii) $\boldsymbol{\varepsilon}$ is MVN with mean $\underline{0}$ and covariance matrix $\mathbf{C}_N = E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T)$. Hegerl et al. (2000) considered the effects of measurement errors in $\mathbf{y}$ and they concluded that the effects are very small. So here, I will simply assume that there are no measurement errors in $\mathbf{y}$.

$\mathbf{C}_N$ is unknown and can be estimated from a control run of a climate model (Section 2.2). Generally, $\mathbf{C}_N \neq \sigma^2\mathbf{I}$. Due to this, ordinary least squares regression would give inefficient estimates of $\boldsymbol{\beta}$. The solution to this is to introduce a pre-whitening agent $\mathbf{P}$ such that

$$E(\mathbf{P}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\mathbf{P}^T) = \mathbf{P}\mathbf{C}_N\mathbf{P}^T = \mathbf{I} \tag{3}$$

and the elements in $\mathbf{P}\boldsymbol{\varepsilon}$ are i.i.d. N(0,1) random variables. From (3), we can see that $\mathbf{P}^T\mathbf{P} = \mathbf{C}_N^{-1}$. Now minimizing the regression sum of square $RSS(\boldsymbol{\beta})$,

$$\begin{aligned} RSS(\boldsymbol{\beta}) &= (\mathbf{P}(\mathbf{X}\boldsymbol{\beta} - \mathbf{y}))^T \mathbf{P}(\mathbf{X}\boldsymbol{\beta} - \mathbf{y}) \\ &= (\mathbf{P}\mathbf{X}\boldsymbol{\beta} - \mathbf{P}\mathbf{y})^T(\mathbf{P}\mathbf{X}\boldsymbol{\beta} - \mathbf{P}\mathbf{y}) \end{aligned} \tag{4}$$

with respect to $\boldsymbol{\beta}$. Assuming $\mathbf{X}$ is noise free, we have

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{P}^T\mathbf{P}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{P}^T\mathbf{P}\mathbf{y} = (\mathbf{X}^T\mathbf{C}_N^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{C}_N^{-1}\mathbf{y} = \mathbf{F}^T\mathbf{y} \tag{5}$$

Under the assumption that $\boldsymbol{\varepsilon}$ is MVN, $\hat{\boldsymbol{\beta}}$ will be normally distributed and will be BLUE by the Gauss-Markov theorem. So, $E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$. If we ignore the uncertainty in estimating $\mathbf{C}_N$ and assuming $\mathbf{X}$ is noise free, the variance of $\hat{\boldsymbol{\beta}}$ is given by

$$V(\hat{\boldsymbol{\beta}}) = V(\mathbf{F}^T\mathbf{y}) = \mathbf{F}^T V(\mathbf{y})\mathbf{F} = \mathbf{F}^T\mathbf{C}_N\mathbf{F} = (\mathbf{X}^T\mathbf{P}^T\mathbf{P}\mathbf{X})^{-1} = (\mathbf{X}^T\mathbf{C}_N^{-1}\mathbf{X})^{-1} \tag{6}$$

and a 100(1-α)% confidence ellipsoid for $\boldsymbol{\beta}$ is

$$(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T V(\hat{\boldsymbol{\beta}})^{-1}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \leq \chi^2_{m,1-\alpha} \tag{7}$$

which is distributed as chi-square with m degrees of freedom. Marginal confidence ellipsoids can be constructed for subsets of signals by reducing the number of degrees of freedom and removing the appropriate rows and columns from $\hat{\boldsymbol{\beta}}$ and $V(\hat{\boldsymbol{\beta}})$.

To avoid bias in calculating the variance of $\hat{\boldsymbol{\beta}}$ (bias comes from the estimation of $\mathbf{C}_N$), Allen & Tett (1999) and references therein suggest that different control runs should be used for calculating the value of $\hat{\boldsymbol{\beta}}$ and the variance of $\hat{\boldsymbol{\beta}}$. This will be discussed in Section 2.2. $\mathbf{X}$ extracted from CGCMs will contain noise because output from each individual forced run is a combination of signal and natural climate noise. To account for this, averages of signals (called ensemble averages) are used in the columns of $\mathbf{X}$. Even using ensemble averages, the response pattern $\mathbf{X}$ is still subject to some sampling uncertainty, i.e., a second ensemble would result in a slightly different $\mathbf{X}$. The variance of the response pattern from an M-member ensemble is approximately $1/M$ times the variance in $\boldsymbol{\varepsilon}$. So, as a first-order correction to resolve the sampling uncertainty, Allen & Tett (1999) suggested the variance of $\hat{\boldsymbol{\beta}}$ should be inflated by $1+1/M$, where M is the number of ensemble members. Using this inflation factor only makes the overall results more conservative, but the bias created by the noise in $\mathbf{X}$ remains. In

general, noise in $\mathbf{X}$ will tend to bias $\hat{\boldsymbol{\beta}}$ towards zero (Fuller, 1987, p.3).

Detection is performed by testing the null hypothesis that $\boldsymbol{\beta} = 0$. Detection of climate change is claimed, at significance level α, for all m signals jointly if the origin is not in confidence ellipsoid (7). Assuming detection, then we can perform the attribution test by testing the null hypothesis that $\boldsymbol{\beta} = \underline{1}$. If ellipsoid (7) contains $(1, 1, ...1)^T$, then we can claim that the observed trend is consistent with the fingerprints. Similarly, marginal detection and attribution tests on a subset of signals can be conducted by using the marginal confidence ellipsoid. As mentioned in Chapter 1, the hypothesis testing procedure for attribution is not well-posed.

As noted by Allen et al. (2000), the conclusion of detection may be uninteresting from a physical point of view since we may already know that a particular forcing agent must have some effect on the climate system. So, what we are really interested in is the range of $\hat{\boldsymbol{\beta}}$ that is consistent with the observed data. This will tell us the extent to which the climate model under- or over-estimates the influence of the forcing agents included in model (1). On the other hand, the claim of attribution is only meaningful if the corresponding range of uncertainty of $\hat{\boldsymbol{\beta}}$ is relatively small.

The regression approach described above has been used in a number of variants depending on how the researcher has treated the time evolution of the signals. In some studies, space-only signal vectors are used. In this type of study, researchers make the assumption that the spatial pattern of the signals does not change over the time period considered. In some studies, space-time signal vectors are used. Consideration of the time evolution of signal patterns is especially useful in multi-signal case. This is because spatial signal patterns tend to be closed to being co-linear when only spatial elements are considered. Taking temporal aspects of the signals into consideration aids in their differentiation by reducing their colineraity. In this

thesis, my signal patterns will consists of both spatial and temporal elements.

2.2 Estimating the natural climate noise covariance matrix

As we have seen in Section 2.1, $\mathbf{C}_N$ is crucial in solving the regression problem. However, $\mathbf{C}_N$ is unknown and has to be estimated. Instrumental observations recorded during the past 150 years cannot provide a reliable estimate of $\mathbf{C}_N$ due to the fact that natural climate noise is confounded with the effect of external forcing during the instrumental period. Also, the length of the observed record is not long enough to estimate $\mathbf{C}_N$ reliably. Thus $\mathbf{C}_N$ is generally estimated from a long control run in which concentrations of greenhouse gases and aerosols are fixed at a present or pre-industrial levels. Even though a growing number of long control run are becoming available, there remains limitations on the spatial scale (size of the grid boxes) in which detection and attribution studies can be conducted. In practice, the available control runs, which are limited to 1000-2000 years, are not long enough to simultaneously estimate the internal variability on a small spatial scale over a 30-50 year time scale. Consequently, detection and attribution studies are conducted in large spatial scales. In general, $\mathbf{C}_N$ can be estimated from a control run by:

$$\hat{\mathbf{C}}_N = \frac{1}{a}\boldsymbol{\varphi}\boldsymbol{\varphi}^T \tag{8}$$

where $\boldsymbol{\varphi} = (\varphi_1 \mid ... \mid \varphi_i \mid ... \mid \varphi_a)$ consists of a columns of $n \times 1$ vector extracted from the control run and each φ_i is a $\mathbf{y}$-like vector. Each column of $\boldsymbol{\varphi}$ will apply the same observation mask to account for missing data. Since $\mathbf{y}$ (observed data) typically consists of data over 30-50 years and control run are limited to 1000-2000 years, the number of independent vectors that can be extracted from the control is less than n, the number of elements in $\mathbf{y}$. Therefore, $\hat{\mathbf{C}}_N$ is non-invertible.

The solution that has been widely adopted in the climate research community is to use the

transformation $\mathbf{P}$ that satisfy (3) to estimate $\mathbf{C}_N$. If we assume that $\hat{\mathbf{C}}_N$ provides a reliable estimate of the noise covariance in the subspace spanned by the k gravest Empirical Orthogonal Functions (EOFs; eigenvectors of $\hat{\mathbf{C}}_N$) of a control run, then a natural transformation to use is $\mathbf{P}^{(k)}$, where the rows of $\mathbf{P}^{(k)}$ are the k^{th} eigenvectors of $\hat{\mathbf{C}}_N$ weighted by square root of their corresponding eigenvalues. So $\mathbf{P}^{(k)}\hat{\mathbf{C}}_N\mathbf{P}^{(k)T}$ is equal to $k \times k$ identity matrix by construction. Thus (5) becomes

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{P}^{(k)T}\mathbf{P}^{(k)}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{P}^{(k)T}\mathbf{P}^{(k)}\mathbf{y} = \mathbf{F}_1^T\mathbf{y} \tag{9}$$

To avoid bias in calculating the variance of $\hat{\boldsymbol{\beta}}$, we will replace (6) by the following estimate,

$$\tilde{V}(\hat{\boldsymbol{\beta}}) = \hat{V}(\mathbf{F}_1^T\omega_i) \tag{10}$$

where ω_i is a φ_i-like vector, but is come from a control run different from the one used in $\hat{\mathbf{C}}_N$. So ω_i and φ_i are statistically independent. From (10), we get

$$\begin{aligned}\tilde{V}(\hat{\boldsymbol{\beta}}) &= E(\mathbf{F}_1^T\omega_i\omega_i^T\mathbf{F}_1) - E(\mathbf{F}_1^T\omega_i)E(\mathbf{F}_1^T\omega_i) \\ &= \mathbf{F}_1^T E(\omega_i\omega_i^T)\mathbf{F}_1 \\ &\approx \mathbf{F}_1^T\tilde{\mathbf{C}}_N\mathbf{F}_1\end{aligned} \tag{11}$$

where $E(\omega_i\omega_i^T) \approx \tilde{\mathbf{C}}_N = \boldsymbol{\omega}\boldsymbol{\omega}^T/b$ with ν degree of freedom and $\boldsymbol{\omega} = (\omega_1 \mid ... \mid \omega_i \mid ... \mid \omega_b)$ consists of b columns of $n \times 1$ vectors that comes from a control run different from the one used in $\hat{\mathbf{C}}_N$. If all the ω_i are independent, then ν would equal to b, but this is never the case in practice. To account for noise in $\mathbf{X}$, $\tilde{V}(\hat{\boldsymbol{\beta}})$ will be inflated by $1 + 1/M$. Equations (10) and (11) hold because the regression is done on a linear scale, so the variance of $\hat{\boldsymbol{\beta}}$ does not depend on the actual amplitude of the signal in $\mathbf{y}$. Hence, an estimate of $V(\hat{\boldsymbol{\beta}})$ can be obtained by applying the same operator $\mathbf{F}_1$(which was used to extract $\hat{\boldsymbol{\beta}}$ from $\mathbf{y}$) to ω_i. Also, note that

$E(\mathbf{F}_1^T\omega_i) = 0$ because this is the expected amplitude of $\hat{\boldsymbol{\beta}}$ with pure noise regressed on the signal $\mathbf{X}$. After updating the variance and substituting it into (7), the confidence ellipsoid becomes

$$(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T \tilde{V}(\hat{\boldsymbol{\beta}})^{-1}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \leq m\mathbf{F}_{m,\nu} \tag{12}$$

which has a $\mathbf{F}$ distribution with m and ν degrees of freedom in the numerator and denominator respectively, taking into account the uncertainty in estimating $\hat{\boldsymbol{\beta}}$ and $\tilde{V}(\hat{\boldsymbol{\beta}})$.

The problem in using EOF analysis is that detection and attribution results depend critically on the choice of k. In general, the estimated variance of $\hat{\boldsymbol{\beta}}$ decreases close to monotonically as k increases (Allen & Tett, 1999). So, as k increases, the confidence ellipsoid shrinks. The reason is that increasing k introduces EOFs that have variances which are unrealistically low and these will in turn decrease the variance of $\hat{\boldsymbol{\beta}}$ by the optimization procedure. Thus a constraint on the choice of k is needed. Allen & Tett (1999) used a residual consistency test to determine the appropriate k. The residual given by the regression model is

$$\hat{\boldsymbol{\varepsilon}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} \tag{13}$$

The idea behind the consistency test is that if $\hat{\mathbf{C}}_N^{(k)}$ can adequately represents the variability in the real world in the truncated space, i.e., the subspace spanned by the first k EOFs of the control run do not contain patterns that have unrealistically low (or high) variances, then the transformed residuals $\mathbf{P}\hat{\boldsymbol{\varepsilon}}$ should behave like i.i.d. normally distributed random variables. So our null hypothesis would be that the transformed residuals are i.i.d. normally distributed random variables and the test statistics is:

$$r^2 = \hat{\boldsymbol{\varepsilon}}^T \mathbf{P}^{(k)T} \mathbf{P}^{(k)} \hat{\boldsymbol{\varepsilon}} = \hat{\boldsymbol{\varepsilon}}^T \hat{\mathbf{C}}_N^{(k)-1} \hat{\boldsymbol{\varepsilon}} \sim \chi^2_{k-m} \tag{14}$$

which would be distributed as chi-square with $k - m$ degrees of freedom under the null hypothesis. If an increase in k introduces EOFs that have unrealistically low variances in the control run, then r^2 will shift to an unreasonably high percentile of χ^2 and so we will reject the null, giving us some warning that estimates of uncertainty are unlikely to be reliable. If different control runs are used for optimization and testing, the statistics changes to:

$$r^2 = \hat{\boldsymbol{\varepsilon}}^T \tilde{\mathbf{C}}_N^{(k)-1} \hat{\boldsymbol{\varepsilon}} \sim (k-m)\mathbf{F}_{k-m,\nu} \tag{15}$$

The above statistic will converge to a χ^2_{k-m} as $\nu \to \infty$ and thus, given the significance level α, the χ^2_{k-m} quantile will always be smaller than the $(k-m)\mathbf{F}_{k-m,\nu}$ quantile. If we use a χ^2 quantile in (15), we will always be more likely to reject the null than using the $\mathbf{F}$ quantile. One should be careful in using such consistency test because if we allow k to vary, we introduce an additional degree of uncertainty into our analysis. This raises the question of which result to present and the danger of presenting the more 'favorable' results. (Allen et al., 2000)

2.3 Example

In this section, we will reproduce part of Zwiers & Zhang (2003)'s result as an example. They assess the extent to which the combined effect of greenhouse gases and sulfate aerosols may be detectable in the twentieth century in six spatial domains that decreases in size. In the following example, I will only present the results from one of the spatial domain, the entire globe, as an example. I first aggregate the $5^o \times 5^o$ grid values of the data into $30^o \times 40^o$ boxes. Separate regression analyses will be performed for each of six overlapping 5-decade periods (period 1: 1900-49, period 2: 1910-59, ..., period 6: 1950-99). The data vector $\mathbf{y}$ for any particular period is constructed from the 5 decade sequence of decadal averages of annual mean temperature anomalies relative to the twentieth century mean, calculated from

non missing years in the period. Annual means are treated as missing if even one month within the year is missing. Decadal means are treated as missing if less than 6 of the 10 years are present. Missing values are not filled in. The signal pattern $\mathbf{X}$ is obtained from the corresponding decades in the combined ensemble average of three CGCM1 GS runs and three CGCM2 GS runs, and are matched in space and time with the non-missing observations.

The reason of using decadal averages instead of annual averages is to account for: i) there can be unusually large natural climate variation in the observed data, and ii) signals obtained from climate model are noise contaminated. So, by using decadal averages, there is hope to filter out some of these noise. In most of the studies, 5-decade periods are used. It is because control runs are limited to 1000-2000 years and it is hard to use periods longer than 5 decade and still provide good estimates of $\mathbf{C}_N$. On the other hand, if the time periods used are too short, there will not be enough time evolution to enter into $\mathbf{X}$. Such evolution is crucial in detection and attribution studies because we are interest in how the signal pattern changes over time.

The covariance matrices $\hat{\mathbf{C}}_N$ and $\tilde{\mathbf{C}}_N$ are calculated from 1600 years of control simulation, composed of 600 years from CGCM1 and 1000 years from CGCM2 (the first 400 years from CGCM1 are not used due to model drift). The 1600 year control simulation is partitioned into two 800 year subsets: first 300yr of CGCM1 and 500yr of CGCM2 for $\hat{\mathbf{C}}_N$ and last 300yr of CGCM1 and 500yr of CGCM2 for $\tilde{\mathbf{C}}_N$. Each $\boldsymbol{\varphi}_i$ and $\boldsymbol{\omega}_i$ has been organized into overlapping 5-decade vectors (overlapped by 4 decade), and observation mask is applied. By using overlapped segments, we can extract more information and hence increase the degrees of freedom in $\tilde{\mathbf{C}}_N$. However, the vectors in $\boldsymbol{\omega}$ will be correlated and so the 'true' ν is unknown. Stott et al. (2001) and references therein suggest using ν equal to 1.5 times the number of non-overlapping segments in the data used to estimate $\tilde{\mathbf{C}}_N$, and I will follow this in my

calculations.

Figure 3 displays the residual consistency test statistics for all six periods. The dotted lines represents the upper and lower 5% critical values of χ^2. χ^2 is used instead of $\mathbf{F}$ because the χ^2 is more conservative. With approximately $k > 15$, the model over-simulates the internal variability. At small k (< 5), the model tends to under-simulate the internal variability. Hence, the number of retained EOFs at which we should trust the analysis would be around 10.

Figures 4 and 5 show the estimated amplitude $\hat{\beta}$ and its 95% confidence interval versus the number of EOFs retained in the truncation (k) for periods 1 and 6. The confidence intervals are calculated from (12) with $m = 1$ and $\nu = 24$. For period 1, with k around 10, the confidence interval does not contain zero, which suggest that $H_0 : \beta = 0$ can be rejected. In addition, the confidence interval contains 1 for k around 10, which indicates that $H_0 : \beta = 1$ cannot be rejected. So the combined influence of greenhouse gas and sulfate aerosols is detectable and attributable. However, we should be careful in interpreting the attribution result because the uncertainty in $\hat{\beta}$ is fairly large. Analogously, detection and attribution tests are done on the other five periods. Detection and attribution of climate change can only be found in periods 2 and 6. However, the uncertainty of $\hat{\beta}$ for both periods is large. In particular, the confidence interval is about (0.1, 1.7) for period 2 and about (0.15, 1.2) for period 6. In both of these two periods, $\hat{\beta}$ is less than 1. This indicates the combined CGCM1 and CGCM2 signal may have over-predicted the combined response to greenhouse gases and sulfate aerosols. That is, the climate model output has to be scaled down to reproduce the observed record. For periods 3, 4 and 5, the confidence intervals contain zero over all level of k. So $H_0 : \beta = 0$ cannot be rejected.

The analysis above suggests that the combined effect of greenhouse gases and sulfate

aerosols are detectable in 1900-49, 1910-59 and 1950-99. Also, attribution claims can be made for these periods. However, we should be cautious about the results because we have not included other forcing agents into the analysis. So, we may run into the danger of detecting and attributing the wrong signal.

	1900-49	1910-59	1920-69	1930-79	1940-89	1950-99
Detection	Yes	Yes	No	No	No	Yes
Attribution	Yes	Yes				Yes

Table 1: Summary of detection and attribution result with 'optimal fingerprinting' for the combined effect of greenhouse gases and sulfate aerosol forcing

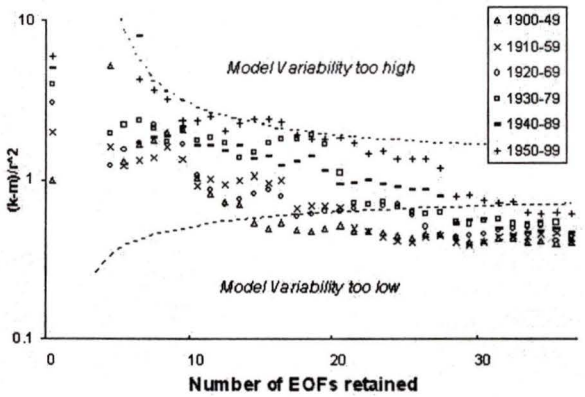

Figure 3: Residual consistency test statistics for different 50 years periods as a function of the number of EOFs retained – after Zwiers & Zhang (2003)

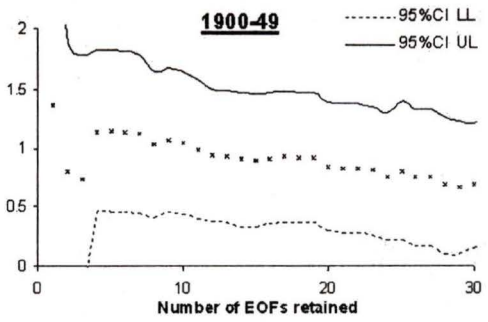

Figure 4: 95% confidence interval of the estimated amplitude $\hat{\beta}$ for 1900-49

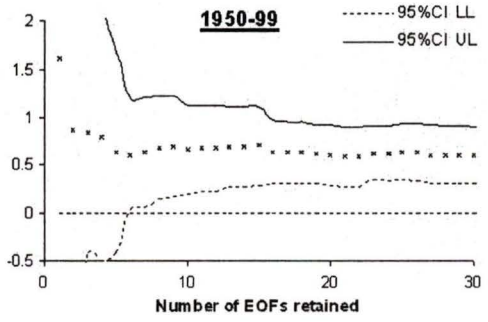

Figure 5: 95% confidence interval of the estimated amplitude $\hat{\beta}$ for 1950-99

3 Bayesian Approach - One-signal case

The Bayesian approach in climate change detection and attribution, proposed by Berliner et al. (2000), builds on the regression analysis result from Chapter 2. In the classical approach, inferences about the parameter $\boldsymbol{\beta}$ are made using a confidence ellipsoid, in which our results are based on significance level α or p-values. On the other hand, the Bayesian approach will make statistical inferences about $\boldsymbol{\beta}$ in terms of direct probability statements. These probability statements are conditional on the observed value of $\hat{\boldsymbol{\beta}}$. While the classical approach is associated with probability statements about how β behave across repeated realization, Bayesian inference aims instead at making probability statements about the true state of β given a particular realization of the data.

In the Bayesian approach, we would like to make probability statements about $\boldsymbol{\beta}$ given $\hat{\boldsymbol{\beta}}$. Thus, we needed the conditional distribution of $\boldsymbol{\beta}$ given $\hat{\boldsymbol{\beta}}$. Namely,

$$\pi(\boldsymbol{\beta} \mid \hat{\boldsymbol{\beta}}) = \frac{f(\hat{\boldsymbol{\beta}}, \boldsymbol{\beta})}{f(\hat{\boldsymbol{\beta}})} = \frac{f(\hat{\boldsymbol{\beta}} \mid \boldsymbol{\beta})\pi(\boldsymbol{\beta})}{\int f(\hat{\boldsymbol{\beta}} \mid \boldsymbol{\beta})\pi(\boldsymbol{\beta})d\boldsymbol{\beta}} \tag{16}$$

Equation (16) is usually referred to as the Bayes' theorem. In this expression, $\pi(\boldsymbol{\beta})$, is called the prior distribution of $\boldsymbol{\beta}$. It expresses our beliefs about $\boldsymbol{\beta}$ (the true value of amplitudes), in terms of a probability density function, before observing the 'data' ($\hat{\boldsymbol{\beta}}$). On the other hand, $f(\hat{\boldsymbol{\beta}} \mid \boldsymbol{\beta})$ reflects the distribution of $\hat{\boldsymbol{\beta}}$ given the true amplitude $\boldsymbol{\beta}$. $\pi(\boldsymbol{\beta} \mid \hat{\boldsymbol{\beta}})$, which reflects the updated beliefs of $\boldsymbol{\beta}$ after observing the 'data', is called the posterior distribution of $\boldsymbol{\beta}$ given $\hat{\boldsymbol{\beta}}$. In another words, the posterior distribution combines the prior beliefs about $\boldsymbol{\beta}$ with the information about $\boldsymbol{\beta}$ contained in $\hat{\boldsymbol{\beta}}$, to give a composite picture of our revised beliefs about $\boldsymbol{\beta}$.

Inference problems concerning $\boldsymbol{\beta}$ can easily be dealt with from the Bayesian viewpoint.

Since the posterior distribution contains all of the available information about $\boldsymbol{\beta}$, any inferences concerning $\boldsymbol{\beta}$ should be drawn from this distribution. One approach to inference is hypothesis testing on $\boldsymbol{\beta}$, in which we calculate the posterior probabilities $p_0 = Pr(H_0 \mid \hat{\boldsymbol{\beta}})$ and $p_1 = Pr(H_1 \mid \hat{\boldsymbol{\beta}})$ and decides between H_0 and H_1 according to the values of p_0 and p_1. The advantage of Bayesian hypothesis testing procedure is that p_0 and p_1 are actual probabilities of the hypotheses given the data and prior beliefs as opposed to indirect p-values and 'statistical significance' in classical hypothesis testing. Another approach to inference is to present a confidence set for $\boldsymbol{\beta}$. In the Bayesian approach, a HPD (highest posterior density) region of $\boldsymbol{\beta}$ could be constructed for this purpose. A HPD region is the Bayesian analog of a classical confidence set, but treating β as random rather than fixed. The 100(1-α)% HPD region for $\boldsymbol{\beta}$ is the subset R that has the smallest possible volume in the $\mathbb{R}^m$ such that:

$$Pr(\boldsymbol{\beta} \in R \mid \hat{\boldsymbol{\beta}}) = \int_R \pi(\boldsymbol{\beta} \mid \hat{\boldsymbol{\beta}}) = 1 - \alpha \tag{17}$$

To make an inference using HPD region, one could apply a similar approach as in interpreting a classical confidence ellipsoid. That is, if the HPD region does not contain $(0, 0, ...0)^T$, then we claim a detection. Furthermore, if the HPD region contains $(1, 1, ...1)^T$, then we could cautiously claim an attribution. The attribution claim should be made with caution because even though when a HPD region contains $(1, 1, ...1)^T$, it does not necessary imply the probability of attribution is large. It only implies that $\boldsymbol{\beta}$ would have a high probability to be in the region defined by the HPD region . So, if the HPD region only consists of a 'small' region around 1, then we can confidently make an attribution claim. Otherwise, we should be cautious when making such a claim.

3.1 Setting up the problem

In this chapter, I will focus on analysis with one forcing agent (GS forcing). From the results of 'optimal fingerprinting', assuming $\mathbf{X}$ is noise free, the distribution of $\hat{\beta} \mid \beta$ is

$$f(\hat{\beta} \mid \beta) = \phi(\beta, \hat{\sigma}^2) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left\{-0.5(\hat{\beta} - \beta)^2/\hat{\sigma}^2\right\} \tag{18}$$

where $\hat{\sigma}^2 = \tilde{V}(\hat{\beta}) = \mathbf{F}_1^T \tilde{\mathbf{C}}_N \mathbf{F}_1$. Note that the above is only approximately true since we do not know the true $\mathbf{C}_N$ to calculate the variance of $\hat{\beta}$. Instead, I used $\tilde{\mathbf{C}}_N$ to estimate the variance of $\hat{\beta}$. If ensemble average of signals is used to account for noise in $\mathbf{X}$, $\hat{\sigma}^2$ will be inflated by $1+1/M$. The detection and attribution framework motivates the prior to be a combination of two components (Berliner et al., 2000). Following Berliner et al. (2000), the prior distribution $\pi(\beta)$ I used is a mixture of two normal distributions given by

$$\begin{aligned} \pi(\beta) &= p\phi(0, \tau^2) + (1-p)\phi(\mu_A, \tau_A^2) \\ &= \frac{p}{\sqrt{2\pi\tau^2}} \exp\left\{-0.5\beta^2/\tau^2\right\} + \frac{1-p}{\sqrt{2\pi\tau_A^2}} \exp\left\{-0.5(\beta - \mu_A)^2/\tau_A^2\right\} \end{aligned} \tag{19}$$

where τ^2 and τ_A^2 are unknown and can be estimated.

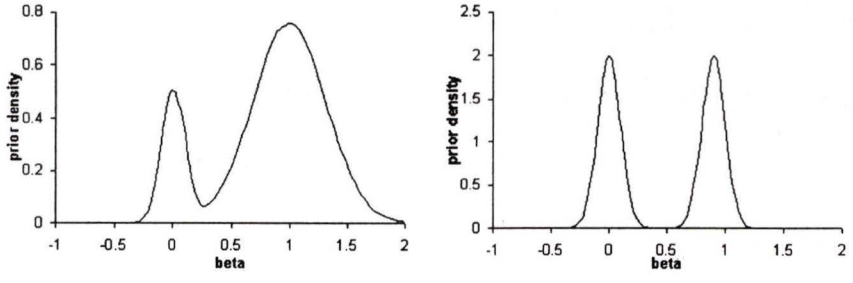

Figure 6: Examples of the shape of the prior given by equation (19) (Left: p=0.4, $\tau^2 = \tau_A^2 = 0.1$, $\mu_A = 1$; Right: p=0.5, $\tau^2 = \tau_A^2 = 0.01$, $\mu_A = 0.9$)

The motivation behind prior (19) is to model the following two possibilities: i) no signal is detected in the observations with probability p and ii) the observed climate change is consistent with the signal pattern with probability $1-p$. That is, the first part of the prior distribution is to model information about β if there is no climate change detected in the fashion reflected by the signal pattern. The choice of zero as the prior mean seems quite natural as we would expect β to be zero if no climate change is detected. The second part of the prior represents when the observed climate change is consistent with what is showing by the signal pattern. In the climate model runs I used, if the observed climate change is attributable to the particular forcing agent, I will expect β to be 1. So, I will use μ_A=1 in my analysis. The τ^2 and τ_A^2 reflect the variability in β when projecting the observed data onto the signal pattern. Specifically, τ^2 and τ_A^2 represent the variability of β when the observed change in climate is not significantly different than what we can expected with natural climate variability ($\beta = 0$) and when the signal is consistent with the observations ($\beta = \mu_A$) respectively. In the Bayesian viewpoint, one could use his/her subjective judgment to set the values of τ and τ_A. In my analysis, I do not have any prior information for a reasonable value of τ and τ_A, so I will estimate them instead.

One way to estimate τ and τ_A is to obtain a sample of βs. To estimate τ^2, we could use segments of the control run as $\mathbf{y}$ and regress it on the fingerprint $\mathbf{X}$ (forced run) to get a sample of βs (using regression method outlined in Chapter 2). Both of $\mathbf{X}$ and the control run segment used in $\mathbf{y}$ have to be matched with the non-missing observations. Then the variance from the resulting sample can be used as an estimate of τ^2. Analogously, τ_A^2 can be estimated in a similar manner but using segments of different forced runs as $\mathbf{y}$ and $\mathbf{X}$. Due to the fact that there are trends in the forced runs, the segments used in both $\mathbf{y}$ and $\mathbf{X}$ would have to be matched in both time and space. Details will be discussed in the next section.

One potential problem with the prior in (19) is that we are restricting ourselves to only two possibilities, the presence and absence of a signal. This ignores the possibility that the climate model may strongly under- or over-estimate the response to the external forcing. That is, we may not give enough weight to β that is between 0 and μ_A and to β that is bigger than μ_A if τ^2 and τ_A^2 are small. Thus, I also introduce a prior, that is a subset of (19), which gives almost equal weight to β between 0 and μ_A and smaller weight to values outside [0, μ_A]. Such a prior can be obtained by setting $p = 0.5$ and $\tau = \tau_A = \mu_A/2$. That is:

$$\pi(\beta) = 0.5\phi(0, (\mu_A/2)^2) + 0.5\phi(\mu_A, (\mu_A/2)^2) \tag{20}$$

In my analysis, μ_A in (20) is set to 1 because we will expect β to be 1 if the observations are consistent with the signal pattern. However, before the analysis, if we know that the climate model tends to (over-) under-estimate the response to external forcing, we should (decrease) increase μ_A accordingly. The effect of using different μ_A will be discussed in Section 3.3.

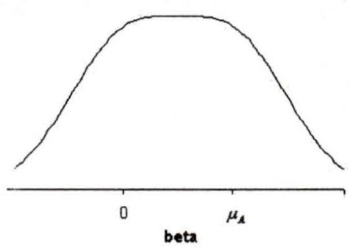

Figure 7: Prior distribution given by (20): this prior give equal weight to β values between 0 and μ_A

Figure 7 shows the prior distribution given by (20). The reason I give smaller weight to values outside [0,1] is because i) it is unlikely that increase in greenhouse gases concentration will have a cooling effect on temperature (this corresponds to $\beta < 0$), ii) I believe that

the climate models used in my analysis are less likely to under-simulate the response to the external forcing (this corresponds to $\beta > 1$) and iii) it is very unlikely that the climate model would under-estimate the response to the external forcing by more than a factor of two (this corresponds to $\beta > 2$).

One should note that in a Bayesian framework, the choice of prior is subjective. So, if one have strong prior belief in attribution, prior (19) can be used with p set to a very small value. On the other hand, if one does not have any prior knowledge on the distribution of β, a noninformative prior can also be used. Such a prior gives equal weight to all β values between -∞ and ∞.

After I defined prior (19), I can obtain the posterior distribution of β. The posterior distribution that is obtained is given by (See Appendix A.1 for derivations)

$$\pi(\beta \mid \hat{\beta}) = q\phi(\tilde{\mu}, \tilde{\tau}^2) + (1-q)\phi(\tilde{\mu}_A, \tilde{\tau}_A^2) \tag{21}$$

where

$$\begin{aligned} \tilde{\mu} &= \frac{\tau^2\hat{\beta}}{\tau^2+\hat{\sigma}^2}, \qquad \tilde{\tau}^2 = \frac{\tau^2\hat{\sigma}^2}{\tau^2+\hat{\sigma}^2} \\ \tilde{\mu}_A &= \frac{\tau_A^2\hat{\beta}+\hat{\sigma}^2\mu_A}{\tau_A^2+\hat{\sigma}^2}, \qquad \tilde{\tau}_A^2 = \frac{\tau_A^2\hat{\sigma}^2}{\tau_A^2+\hat{\sigma}^2} \\ q &= \left(1+\frac{1-p}{p}\sqrt{\frac{\tau^2+\hat{\sigma}^2}{\tau_A^2+\hat{\sigma}^2}}\ \exp\left\{-0.5\left[\frac{(\hat{\beta}-\mu_A)^2}{\tau_A^2+\hat{\sigma}^2}-\frac{\hat{\beta}^2}{\tau^2+\hat{\sigma}^2}\right]\right\}\right)^{-1} \end{aligned} \tag{22}$$

Again, the above is only approximately true because we do not know the true variance of $\hat{\beta}$.

I next formalize the criteria for detection and attribution. Following Berliner et al. (2000), I define detection as occurring if the posterior probability

$$Pr(\beta \in \mathcal{D} \mid \hat{\beta}) \quad \text{is small} \tag{23}$$

where $\mathcal{D}$ is some neighbourhood of zero and we define attribution as occurring if

$$Pr(\beta \in \mathcal{A} \mid \hat{\beta}) \quad \text{is large} \tag{24}$$

where $\mathcal{A}$ is some neighbourhood of 1. For attribution, in addition to looking at the probability in (24), one should also examine the posterior distribution before making an attribution claim. Consider the following example: suppose we have a posterior like the one shown in Figure 8 and we define $\mathcal{A}$ to be (0.9, 1.1), then β will have almost equal probability to be in $\mathcal{A}$ and in (0.5, 0.7). So, even though $Pr(\beta \in \mathcal{A} \mid \hat{\beta})$ is around 0.5, we still should not make an attribution claim. That is, we will have to be careful when interpreting the attribution result if $Pr(\beta \in \mathcal{A} \mid \hat{\beta})$ is not very large (near 1) or if q is neither small nor large.

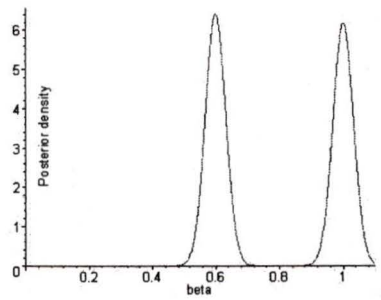

Figure 8: An example of a problematic posterior distribution

As suggested by Berliner et al. (2000), two important questions arise. First, how should $\mathcal{D}$ and $\mathcal{A}$ be defined? Second, what do we mean by small and large probabilities in (23) and (24)? In a Bayesian framework, these decisions are subjective. Subjective decision also occurs in the frequentist approach. One example is choosing the cutoff significance level between whether to reject the null hypothesis of no detection and the null hypothesis of attribution.

3.2 Climate change assessment with different priors

In the following three subsections, separate analyses with the three priors discussed in the last section will be described. For illustrative purposes, I will present inferences using both HPD regions and criteria (23) and (24). The data setup will be the same as in Section 2.3, with a separate analysis for each of the six time periods. Again, I am interested in showing the observed temperature change is detectable and attributable to GS forcing. In my analysis, I defined the non-detection region as $\mathcal{D} = (-\infty, 0.1)$. The interval used in my non-detection region is slightly different from Berliner et al. (2000). They used $[0-\epsilon, 0+\epsilon]$ to be their non-detection region, where ϵ is a 'small' number. In my analysis, all negative values are included in the interval. The inclusion of all negative values is to reflect the fact that detection is associated with β greater than zero. When β is less than zero, it implies that the model projects the observations with the wrong sign. When the amplitude is of the wrong sign, no detection should be claimed even the amplitude may be quite different from zero.

I also defined three attribution regions denoted as $\mathcal{A}_1 = (0.9, 1.1)$, $\mathcal{A}_2 = (0.8, 1.2)$ and $\mathcal{A}_3 = (0.7, 1.3)$ to be used in the analysis. Which attribution region is to be used to interpret the results of the analysis depends on the level of β the researcher is comfortable with when making an attribution claim. In the following analysis, a decision on whether to make an attribution claim will be based on $Pr(\beta \in \mathcal{A}_2 \mid \hat{\beta})$. The other two attribution regions will be used as comparisons. In addition, I will claim to have detected climate change if $Pr(\beta \in \mathcal{D} \mid \hat{\beta}) < 0.1$. The cutoff point 0.1 is chosen so that the probability of wrongly claiming a detection is at most 0.1, which would be parallel with the 10% significance level in the classical approach. For attribution, I will consider $Pr(\beta \in \mathcal{A}_2 \mid \hat{\beta})$ to be large if it is greater than 0.6. This implies that the probability of wrongly making an attribution claim is at most 0.4. Even though this error rate is quite high, I think that it is still acceptable. It is because

the large uncertainty in $\hat{\beta}$ would make the posterior fairly spread out and makes us impossible to obtain a near-certain attribution claim (that is to have a attribution probability that is close to 1 for a very small region around β=1). So if I set the cutoff to be too high (say 0.9), my Bayesian analysis will not be able to make an attribution claim even if such claim is 'true'. Table 2 listed my rules for quantifying the evidence of detection and attribution. Note that these rules are chosen subjectively and are merely chosen to increase the readability of the results. Tables are provided at the end of this section to summarize the results of the three analyses.

Evidence of detection	Very Strong	Strong	No
$Pr(\beta \in \mathcal{D} \mid \hat{\beta})$	< 0.01	0.01 - 0.1	> 0.1

Evidence of attribution	No	Small	Mild	Strong
$Pr(\beta \in \mathcal{A}_2 \mid \hat{\beta})$	< 0.4	0.4 - 0.5	0.5 - 0.6	> 0.6

Table 2: Rules for quantifying posterior probabilities for detection and attribution

Prior A

The following analysis is done using the prior given by (19) with μ_A=1. Since I do not have any information suggesting whether there is a climate change, I will set p=0.5. For each of the six time period, I estimate the value of τ and τ_A. The reason I do not use the same value of τ and τ_A for all six periods is because data coverage is different for each time period and I believe this may have an effect on the value of τ and τ_A. To estimate τ^2, I partition the 1600 year control run into 152 50-year overlapping segments (56 from CGCM1 and 96 from CGCM2), which will result in 152 βs at each EOF truncation level for each time period.

Each of these segments consists of 5 decades matched with the non-missing observations. The resulting variance from the sample will be my estimate of τ^2. To estimate τ_A^2, a more complicated procedure is applied. Since there are six CGCM forced runs available, I will use segments from one forced run as $\mathbf{y}$ and segments of the ensemble of the other five runs as $\mathbf{X}$ to obtain a sample of β. Each forced run has 200 years and each segment is 5 decades, so this will result in 16 overlapped segments. The segments used in $\mathbf{y}$ and $\mathbf{X}$ are matched in time to account for the trends in the signal. To account for the difference in data coverage in each time period, observation mask is applied to both of the forced run segments used in $\mathbf{y}$ and $\mathbf{X}$. By permuting the forced run used in $\mathbf{y}$ and $\mathbf{X}$, we will have $16 \times 6 = 96$ βs in total for each time period and the variances from the sample will be the estimate of τ_A^2.

Figure 9 shows the prior distribution obtained in this way. It seems that the priors do not give a lot of weight for β near 0.5. Figure 10 shows the estimated τ and τ_A for periods 1, 3 and 6. We can see that τ is less than τ_A for periods 1 and 6. For period 2, the same trend is observed. For periods 3, 4 and 5 (periods 4 and 5 not shown), τ is greater than τ_A at most EOF truncation. Nevertheless, τ and τ_A are only differ by 0.1 to 0.05 in all six periods. Therefore, it may be sensible to set τ equal to τ_A. However, since I have already estimated τ and τ_A, I will not set them equal to each other in my analysis to avoid losing any information.

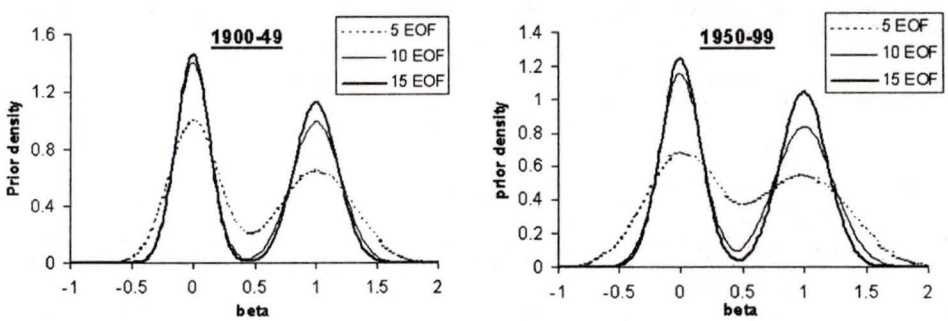

Figure 9: Prior A - Prior distribution for 1900-49 and 1950-99 with 5, 10 and 15 EOFs retained

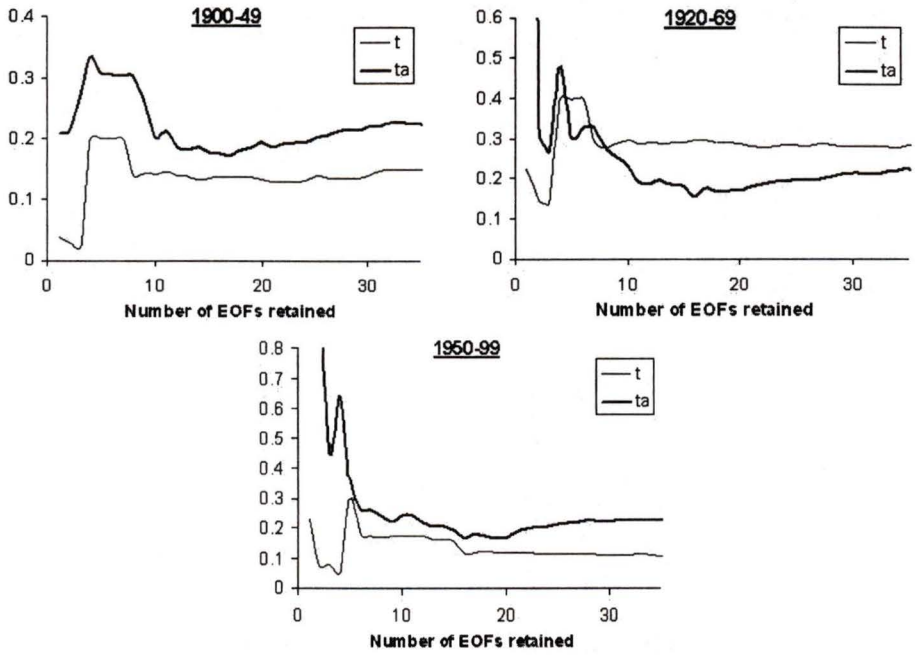

Figure 10: Prior A - Estimates of the value of prior parameters (τ and τ_A) as a function of the number of EOFs retained

Figure 11 shows the posterior probabilities in $\mathcal{D}$ and $\mathcal{A}$ and the posterior parameter q. For period 1 (1900-49), $Pr(\beta \in \mathcal{D} \mid \hat{\beta})$ is close to zero and hence very strong evidence for detection. The attribution likelihood, $Pr(\beta \in \mathcal{A}_2 \mid \hat{\beta})$, increases until k=15 (k is the number of EOFs retained) and then declines for k >15. The residual consistency test from Chapter 2 suggests that my analysis should be conducted with k around 10. For k around 10, $Pr(\beta \in \mathcal{A}_2 \mid \hat{\beta}) \approx$ 0.7, which suggests strong evidence for attribution. In addition, q is very small, which means the weight on the posterior component centered on $\tilde{\mu}$ is small. Figure 12 shows the posterior distribution for period 1 at k=5, 10 and 15. Unlike the posterior distribution in Figure 8, the distribution here is well behaved. Therefore, I claim an attribution in period 1 with strong evidence.

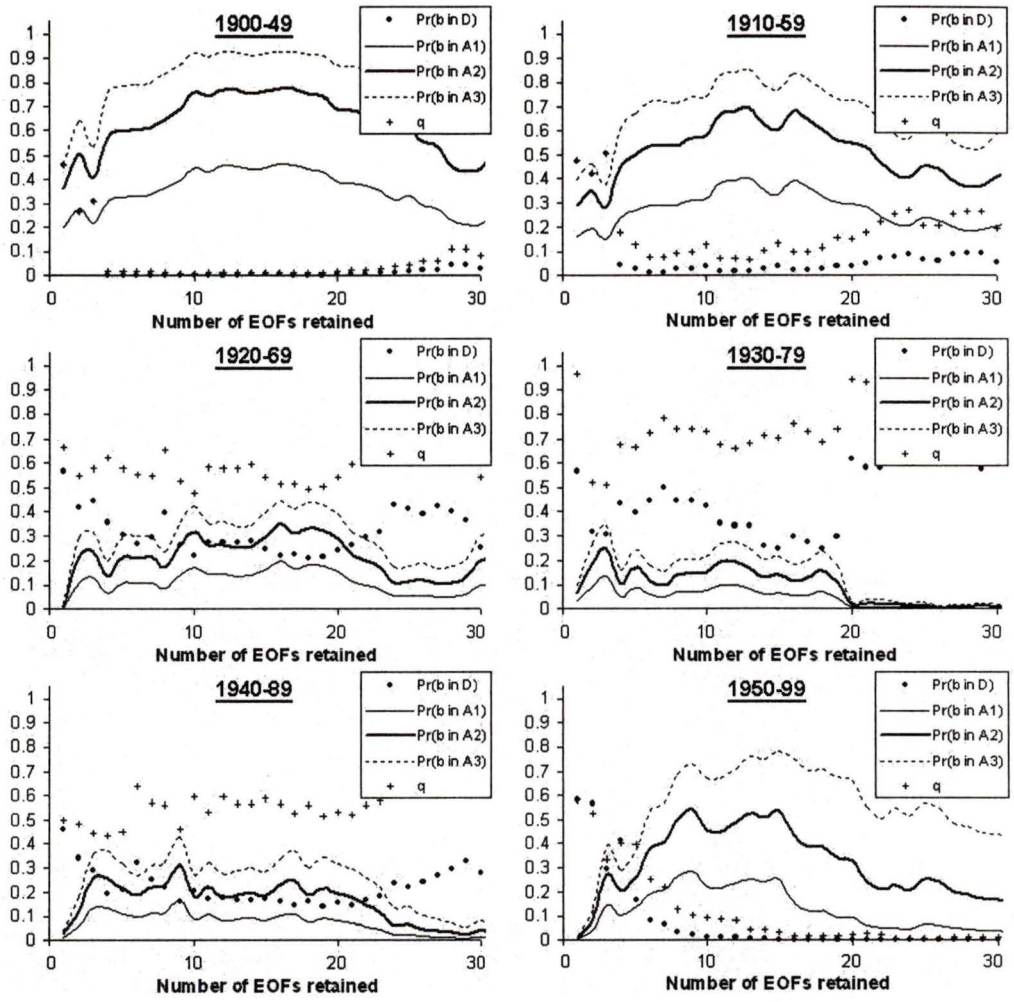

Figure 11: Prior A - Posterior parameter and probabilities of no detection and attribution

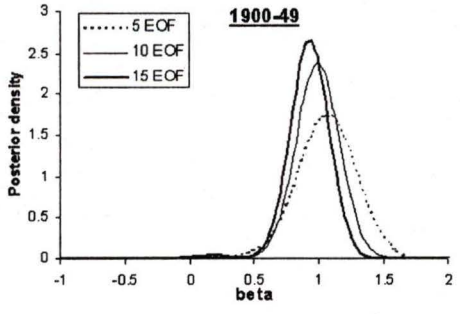

Figure 12: Prior A - Posterior distribution for 1900-49 with 5, 10 and 15 EOFs retained

For period 2, $Pr(\beta \in \mathcal{D} \mid \hat{\beta}) \approx 0.03$ with k around 10, which suggests there is also strong evidence for detection in this period. At the same time, $Pr(\beta \in \mathcal{A}_2 \mid \hat{\beta}) \approx 0.55$ for k around 10, and q is on the order of 0.1 which suggest the posterior is not problematic. Therefore, there is mild evidence in which to claim an attribution. In period 6, $Pr(\beta \in \mathcal{D} \mid \hat{\beta})$ is essentially zero for k around 10. So there is very strong evidence for detection in period 6. Also, $Pr(\beta \in \mathcal{A}_2 \mid \hat{\beta}) \approx 0.45$ for k around 10. Figure 13 shows the posterior distribution of period 6 and the posterior is well behaved for k =10 and 15. For $k = 5$, the posterior distribution are almost equally weighted on the two posterior components. Since I am only interested in k around 10, I conclude there is some evidence in which to claim an attribution in period 6. However, a careful look at the posterior indicates that most of the posterior probabilities are allocated to $\beta \in (0.6, 0.9)$, indicating that the amplitude of the signal may have over-simulated by CGCM1 and CGCM2 combined. For periods 3, 4 and 5, $Pr(\beta \in \mathcal{D} \mid \hat{\beta})$ > 0.2 for k around 10 and thus there is insufficient evidence to claim a detection of the GS signal in the observations of these periods.

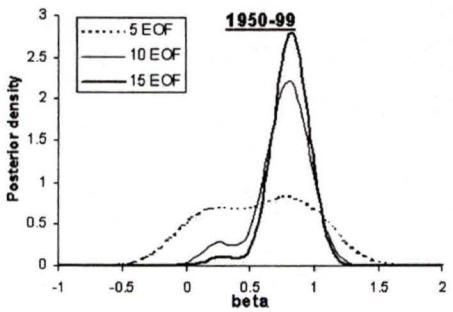

Figure 13: Prior A - Posterior distribution for 1950-99 with 5, 10 and 15 EOFs retained

In the frequentist approach, it may be useful to construct a confidence interval on $\hat{\beta}$. In the Bayesian viewpoint, we could construct a HPD region of β for this purpose. To construct the 100(1-α)% HPD region for β, it is necessary to find the subset R with the smallest possible

length that satisfies (17). Since my posterior is a mixture of two normal distributions, some work is required to obtain such a subset. For simplicity, instead of finding the minimum length subset, I will find the interval R that has 100(1-α/2)% tail probabilities (denoted IR interval). The length of the IR interval will be close to that of R when q is near 0 or 1.

Figure 14 shows the 95% IR intervals for β in periods 1 and 6. In period 1, for k around 10, the 95% IR interval is about (0.6, 1.4), which is slightly narrower than the confidence interval from Section 2.3. For period 2 (not shown) and period 6, the 95% IR intervals for k around 10 are about (0.1, 1.2). Since all of these IR intervals contain one but not zero, detection and attribution for GS signal is cautiously claimed in periods 1, 2 and 6. However, these IR intervals are very wide, especially for periods 2 and 6, which indicates that the evidence for attribution is not strong. For periods 3, 4 and 5 (not shown), the IR intervals are about (-0.4, 1.2) for k around 10 and hence detection of the GS signal is not claimed in these periods.

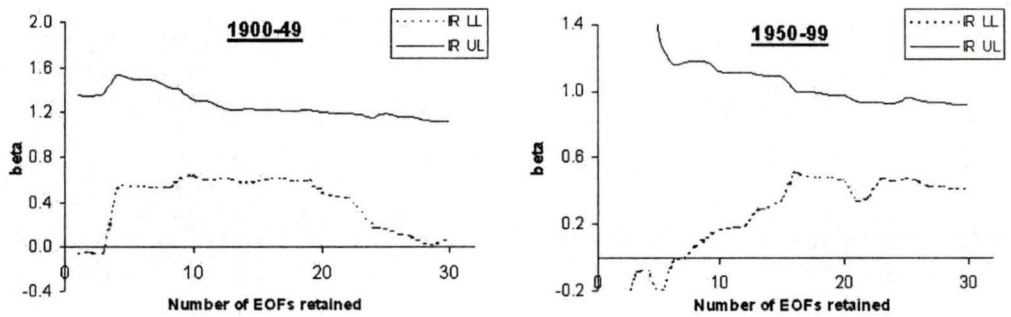

Figure 14: Prior A - 95% IR intervals of the true amplitude, β, in periods 1 and 6

Prior B

The analysis described above is now repeated with prior (20) with $\mu_A = 1$. The posterior parameter q and probabilities in $\mathcal{D}$ and $\mathcal{A}$ for all six periods are shown in Figure 15. Again, we are only interest in results for k around 10.

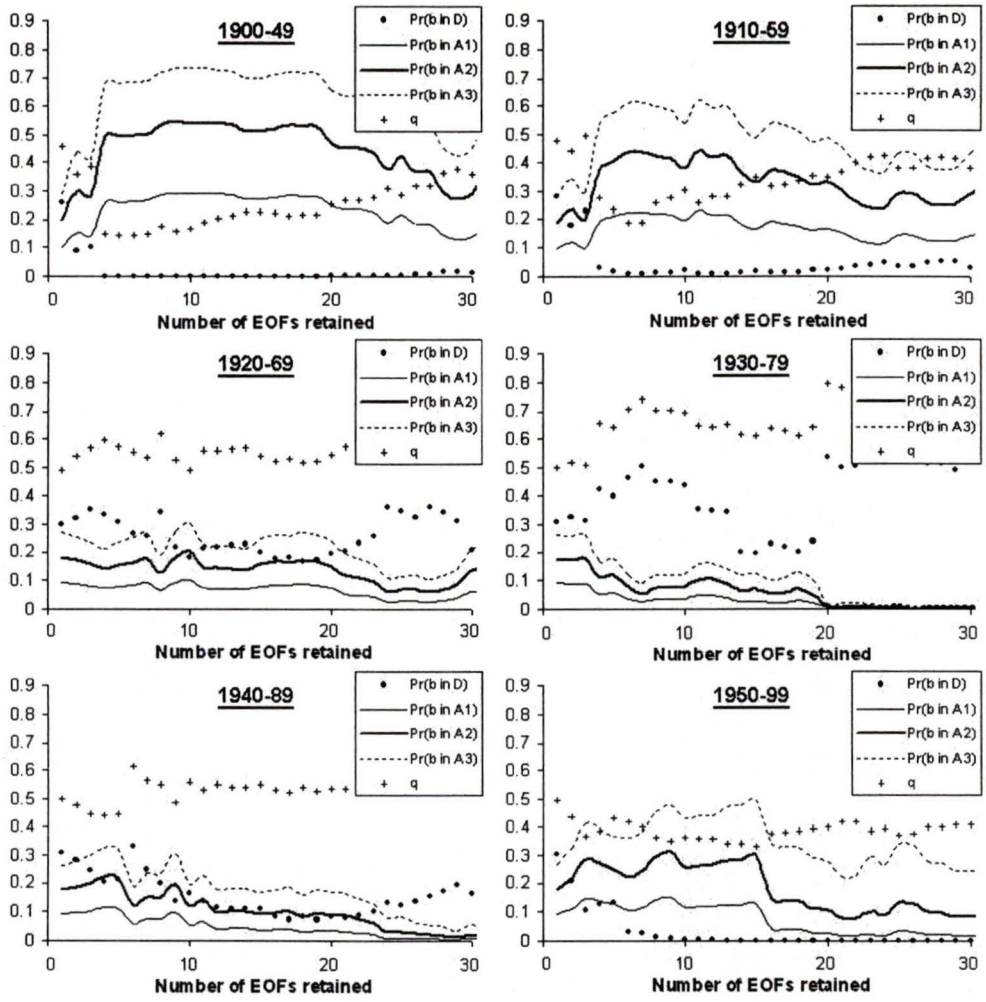

Figure 15: Prior B - Posterior parameter and probabilities of no detection and attribution

For period 1, there is very strong evidence for detection and mild evidence for attribution. Since q is small, we do not have a problematic posterior (See Figure 16) and thus, I cautiously claim attribution for period 1.

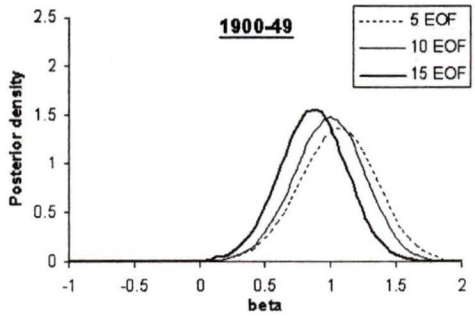

Figure 16: Prior B - Posterior distribution for 1900-49 with 5, 10 and 15 EOFs retained

There is also strong evidence for detection in periods 2 and 6. In addition, $Pr(\beta \in \mathcal{A}_2 \mid \hat{\beta})$ ≈ 0.4 for period 2 and is approximately 0.25 for period 6 and thus, I very cautiously claim attribution in period 2. In period 6, most of the posterior probability are allocated to $\beta \in (0.4, 0.8)$, which suggests the climate model may over-estimate the response to the GS forcing. On the other hand, $Pr(\beta \in \mathcal{D} \mid \hat{\beta}) > 0.1$ for periods 3, 4 and 5 and hence I do not make a detection claim in these periods.

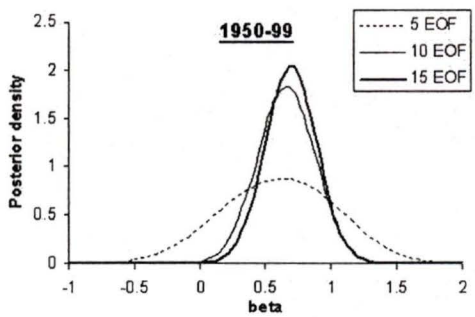

Figure 17: Prior B - Posterior distribution for 1950-99 with 5, 10 and 15 EOFs retained

Figure 18 shows the 95% IR intervals for β in periods 1 and 6. We can see that the 95% IR interval for period 1 is about (0.5, 1.5) and that it is about (0.2, 1.1) for period 6 for $k \approx 10$. Not displayed here, the 95% IR interval for k around 10 is about (0.2, 1.4) for period 2. So,

I cautiously make a detection and attribution claim in these three periods. For periods 3, 4 and 5, the IR intervals are about (-0.4, 1.2) and therefore, as above, no detection claims are made in these periods.

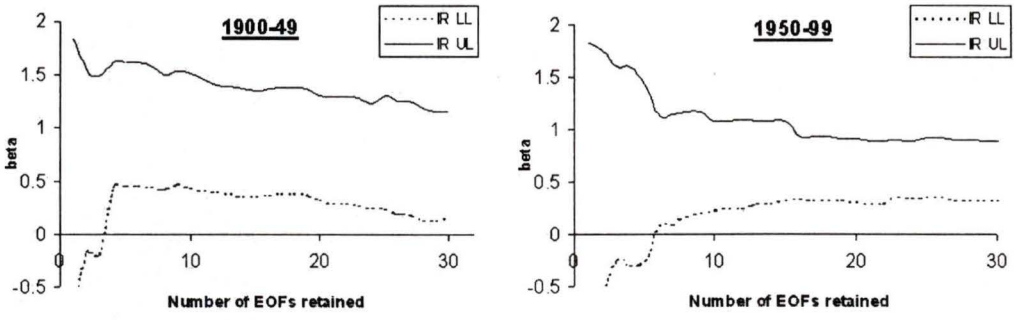

Figure 18: Prior B - 95% IR intervals of the true amplitude, β, in periods 1 and 6

Prior C

The analyses described above were repeated once more with a noninformative prior which gives equal weight to all β between $-\infty$ to ∞. A noninformative prior is used when we do not have any knowledge on the distribution of β and so we give equal weight to any real values of β. The noninformative prior is defined as:

$$\pi(\beta) \propto 1 \tag{25}$$

and the posterior will be (obtained by Bayes theorem):

$$\pi(\beta \mid \hat{\beta}) = \frac{f(\hat{\beta} \mid \beta)}{\int f(\hat{\beta} \mid \beta) d\beta} = f(\hat{\beta} \mid \beta) = \phi(\hat{\beta}, \hat{\sigma}^2) \tag{26}$$

Posterior probabilities of no detection and attribution for all six periods are shown in Figure 19. Figure 20 shows the posterior distributions for periods 1 and 6 at different k. From these figures, detection is claimed for periods 1 and 6 with very strong evidence and

attribution is claimed cautiously for period 1. For period 2, detection is claimed with strong evidence.

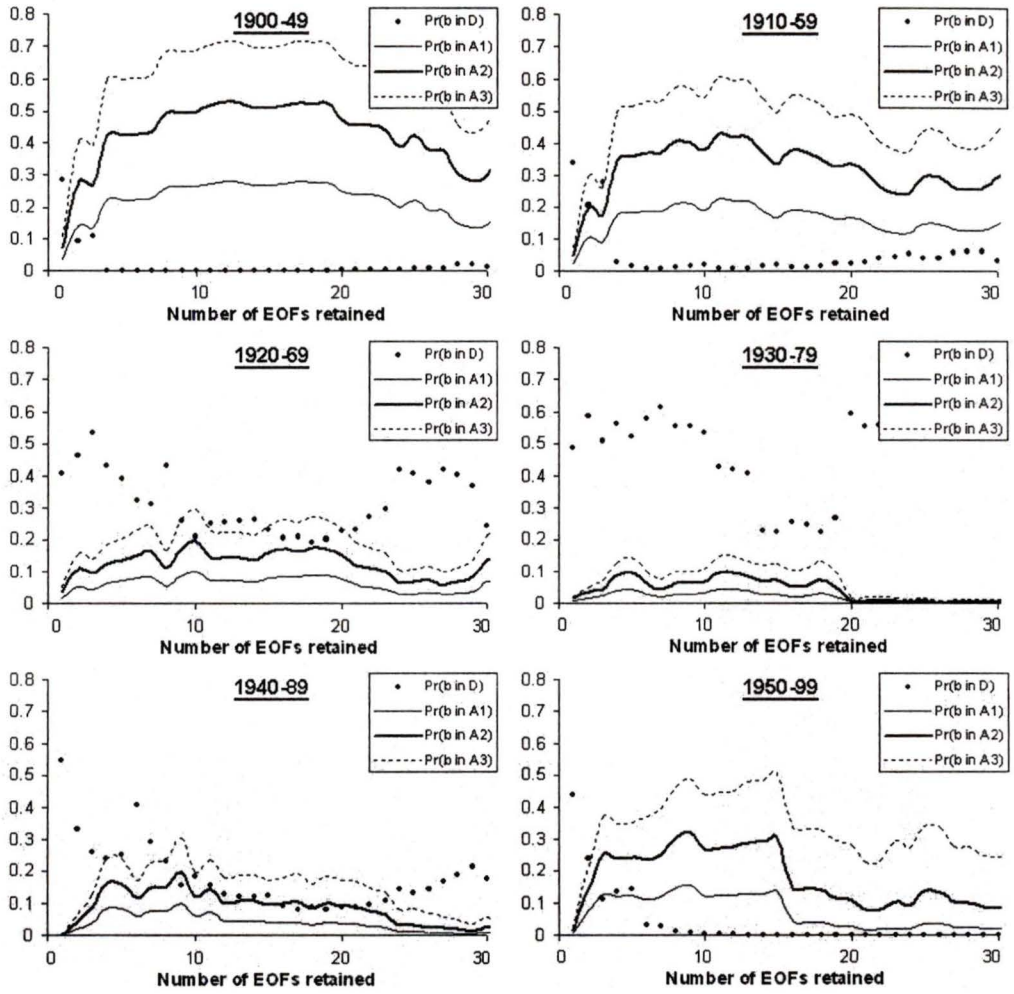

Figure 19: Prior C - Posterior parameter and probabilities of no detection and attribution

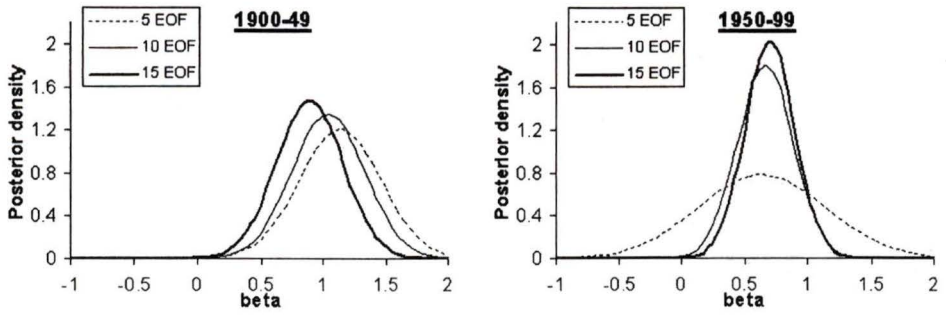

Figure 20: Prior C - Posterior distributions for 1900-49 and 1950-99 with 5, 10 and 15 EOFs retained

Since our posterior only consists of one normal distribution, the HPD region will be the same as IR interval. The 95% IR intervals for periods 1 and 6 are shown in Figure 21. These intervals are similar to the confidence intervals obtained in Section 2.3 due to the choice of a noninformative prior. All we need to do is to substitute the **F** distribution in (12) to a chi-square distribution. So, the IR intervals are slightly narrower than the confidence intervals reported in Section 2.3. For periods 1, 2 and 6, the IR intervals contain one but not zero for $k \approx 10$. For periods 3, 4 and 5, the IR intervals contain zero for most value of the k.

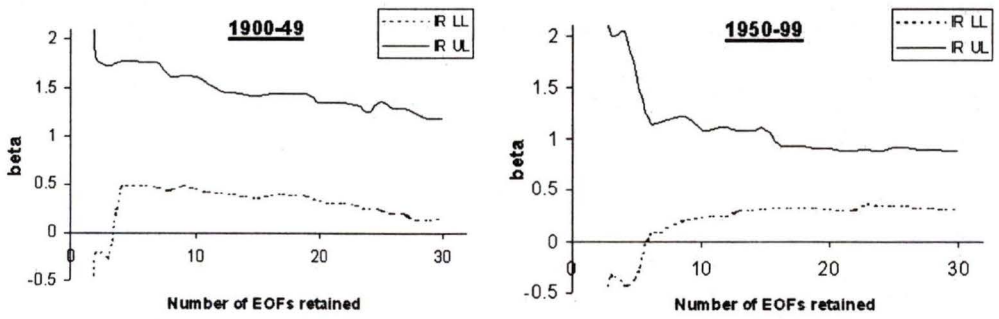

Figure 21: Prior C - 95% IR intervals of the true amplitude, β, for period 1 and 6

Summary of three analyses

Tables 3 and 4 summarize the results from using criteria (23) and (24). Detection results are the same among the three priors and attribution results are fairly robust within each attribution region. When using IR intervals for inference, I claimed detection and attribution (claimed cautiously) in periods 1, 2 and 6 for all three priors. Hence, my analysis appears to be robust with respect to the choice of prior when using the IR interval for inference. Even though I claim attribution for these three periods, the IR intervals are fairly wide. Comparing the Bayesian result to those obtained in Chapter 2, we can see that they are very similar.

From the three analyses, the results indicate that the observed changes in temperature during 1900-49, 1910-59 and 1950-99 are unlikely to be due to natural climate variability alone. For 1900-49, with criteria (23) and (24), my analysis suggests that the observed change might be attributable to the combined effect of greenhouse gas and sulfate aerosols; but such a statement cannot be claimed with confidence for the other two time periods. We should be cautious when interpreting this result, because we have not investigated the influences of other potential forcing agents or considered the fingerprints from other climate models. In all three analyses, most of the posterior probability is allocated to $0.5 < \beta < 0.9$ in 1950-99. Thus, it is likely that the combined CGCM1 and CGCM2 signal has over-estimated the response to GS forcing.

Detection results

	1900-49	1910-59	1920-69	1930-79	1940-89	1950-99
Prior A	V.Strong	Strong	No	No	No	V.Strong
Prior B	V.Strong	Strong	No	No	No	V.Strong
Prior C	V.Strong	Strong	No	No	No	V.Strong

Table 3: Evidence of detection across different priors obtained by calculating the posterior probability of no detection

Attribution results with $\mathcal{A}_1$

	1900-49	1910-59	1920-69	1930-79	1940-89	1950-99
Prior A	No	No	No	No	No	No
Prior B	No	No	No	No	No	No
Prior C	No	No	No	No	No	No

Attribution results with $\mathcal{A}_2$

	1900-49	1910-59	1920-69	1930-79	1940-89	1950-99
Prior A	Strong	Mild	No	No	No	Small
Prior B	Mild	Small	No	No	No	No
Prior C	Small	No	No	No	No	No

Attribution results with $\mathcal{A}_3$

	1900-49	1910-59	1920-69	1930-79	1940-89	1950-99
Prior A	Strong	Strong	No	No	No	Strong
Prior B	Strong	Mild	No	No	No	Small
Prior C	Strong	Mild	No	No	No	Small

Table 4: Evidence of attribution across different priors obtained by calculating the posterior probability of attribution in three different regions

3.3 Robustness of the one-signal analysis with respect to prior specification

In the previous section, analysis was conducted with three very different priors. These different priors yield very similar conclusions for all six periods and so my analysis appears to be robust to the choice of prior. To further illustrate the robustness of my analysis, I will present results for a plausible class of priors that are contained in (20).

In the previous analyses, I assume that μ_A =1, but in fact I may be uncertain about μ_A because there is a possibility that the climate model could over or under-estimate the response to external forcing. If the model tends to over-estimate the response to external forcing, I should set $\mu_A < 1$ accordingly. Conversely, if the model tends to under-estimate the response to external forcing, I should use $\mu_A > 1$ in my prior. So, I define a plausible class of priors that are contained in (20) to be:

$$\Gamma = \{\pi(\beta) : 0.5 \leq \mu_A \leq 2\} \tag{27}$$

The cutoff is chosen to be 0.5 and 2 is because it is very unlikely that the climate model will over-estimate or under-estimate the response to external forcing by more than a factor of two.

To conduct the robustness analysis, I computed the bounds for $Pr(\beta \in \mathcal{D} \mid \hat{\beta})$ and $Pr(\beta \in \mathcal{A}_2 \mid \hat{\beta})$ over Γ. Figures 22 and 23 show the upper and lower bounds of $Pr(\beta \in \mathcal{D} \mid \hat{\beta})$ and $Pr(\beta \in \mathcal{A}_2 \mid \hat{\beta})$ for periods 1, 2 and 6. Notice that the ranges of the probabilities are tighter for detection. Specifically, for period 1 (period 6), the width of the bound for $Pr(\beta \in \mathcal{D} \mid \hat{\beta})$ is about 0.003 (0.002) and is about 0.2 (0.16) for $Pr(\beta \in \mathcal{A}_2 \mid \hat{\beta})$ for $k \approx 10$. Similar behavior with the probability bound can be found in period 2. Not displayed here, for periods 3, 4 and 5, the width of the bounds for $Pr(\beta \in \mathcal{D} \mid \hat{\beta})$ is around 0.02 and around 0.1 for $Pr(\beta \in \mathcal{A}_2 \mid \hat{\beta})$. This suggests that detection probability is robust and attribution probability is less robust with respect to the priors within Γ for all six time periods.

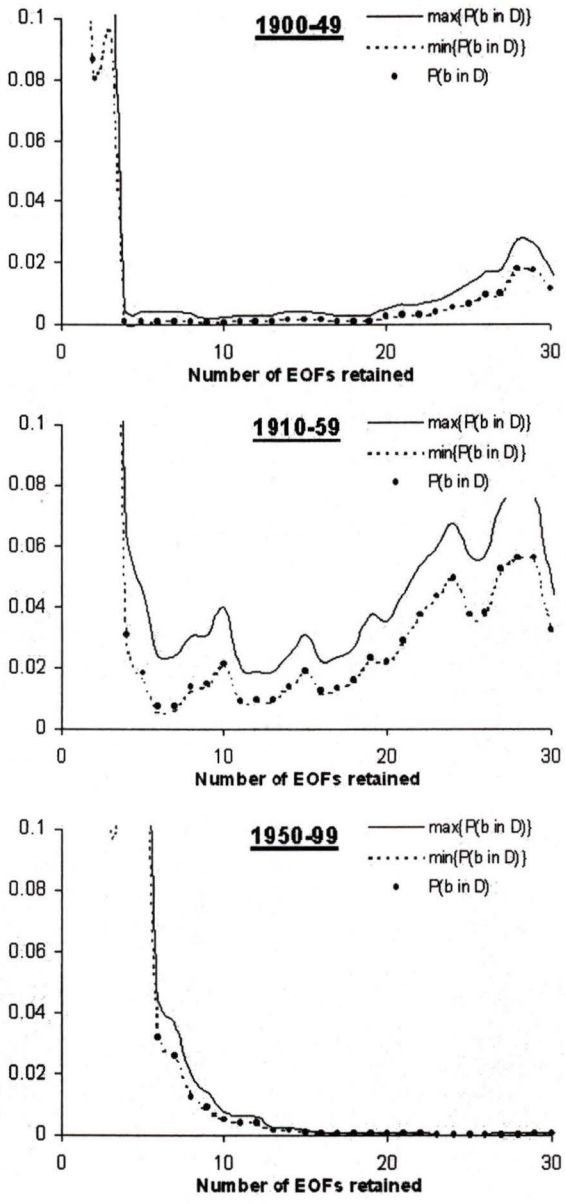

Figure 22: Posterior probability bounds for detection, $Pr(\beta \in \mathcal{D} \mid \hat{\beta})$, within the class of prior Γ

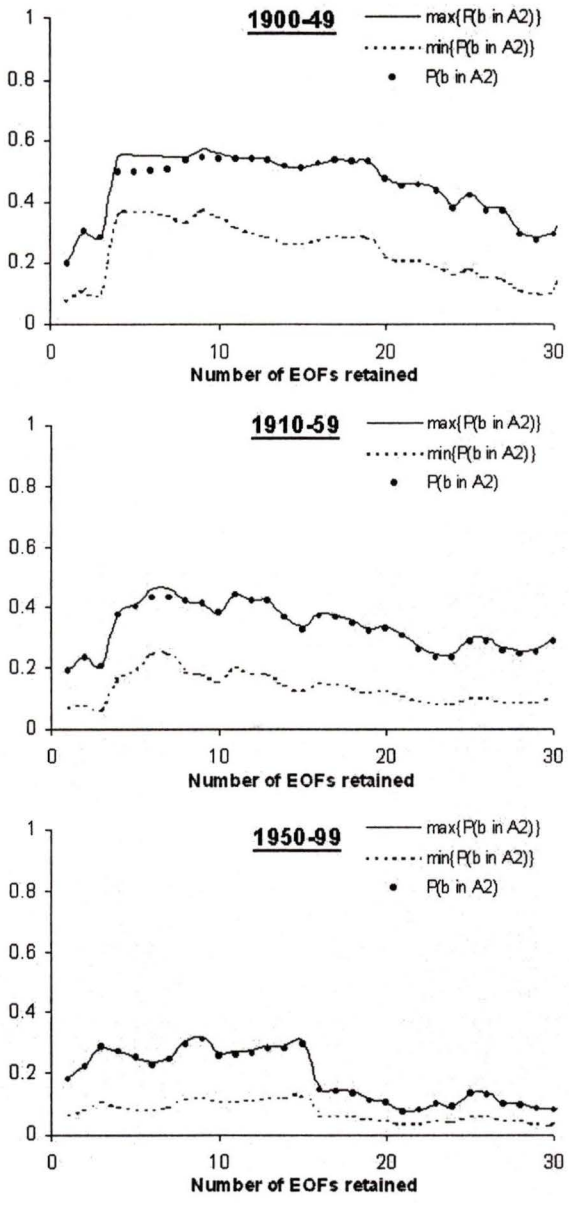

Figure 23: Posterior probability bounds for attribution, $Pr(\beta \in \mathcal{A}_2 \mid \hat{\beta})$, within the class of prior Γ

4 Bayesian Approach - Two-signal case

In this chapter, I will extend the Bayesian analysis to two signals. In the pervious two chapters, I considered the combined effect of greenhouse gases and sulfate aerosols. In this chapter, I will consider these effects separately. For completeness, I will first present the two-signal analysis using 'optimal fingerprinting'. I will then proceed into the Bayesian analysis.

The signals used in the following analysis are from the HadCM2 coupled ocean-atmosphere general circulation model (Johns et al., 1997). In particular, ensemble averages of nine GS runs and ensemble averages of four G runs will be used in the two columns of $\mathbf{X}$. As in the pervious chapters, the HadCRUTv observational dataset will be used. Again, grid values will be aggregated into $30^o \times 40^o$ boxes. Then separate analyses will be preformed for each of the six overlapping 5-decade periods (period 1: 1896-1946, period 2: 1906-56, ..., period 6: 1946-96). Missing data will be treated as in Section 2.3 and the decadal temperature anomalies used are calculated relative to 1896-1996. Since the HadCM2 data available to me is organized into decadal averages starting at years ending with six, so the time periods used in this chapter are different from the previous two chapters. To construct the covariance matrices $\hat{\mathbf{C}}_N$ and $\tilde{\mathbf{C}}_N$, a 1600 year HadCM2 control run is used. The 1600 year control run is partitioned into two 800 year subsets which are used to compute $\hat{\mathbf{C}}_N$ and $\tilde{\mathbf{C}}_N$ respectively. Again, the residual consistency test will be performed to determine the appropriate truncation level (k).

Since both of the G and GS runs contain greenhouse gases forcing, and it is of interest to estimate the total greenhouse gas effect, I will extract such information from the regression model. Our regression model for the two-signal case is,

$$\mathbf{y} = \beta_G \mathbf{X}_G + \beta_{GS} \mathbf{X}_{GS} + \boldsymbol{\varepsilon} \tag{28}$$

where $\mathbf{X}_G$ and $\mathbf{X}_{GS}$ are the G and GS signals respectively, and β_G and β_{GS} are the corresponding amplitudes.

Following Allen & Tett (1999), I will estimate the total effect of greenhouse gases by first denote $\hat{\beta}_G$ and $\hat{\beta}_{GS}$ to be the amplitudes estimated from the regression model. Then, assuming the GS signal is the sum of pure greenhouse gases signal ($\mathbf{X}_G = \mathbf{X}_{GHG}$) and pure sulfate aerosols signal ($\mathbf{X}_{SUL}$), we have

$$\begin{aligned}
\hat{y} &= \hat{\beta}_G \mathbf{X}_G + \hat{\beta}_{GS} \mathbf{X}_{GS} \\
&= \hat{\beta}_G \mathbf{X}_G + \hat{\beta}_{GS} (\mathbf{X}_G + \mathbf{X}_{SUL}) \\
&= \hat{\beta}_G \mathbf{X}_{GHG} + \hat{\beta}_{GS} (\mathbf{X}_{GHG} + \mathbf{X}_{SUL}) \\
&= (\hat{\beta}_G + \hat{\beta}_{GS}) \mathbf{X}_{GHG} + \hat{\beta}_{GS} \mathbf{X}_{SUL} \\
&= \hat{\beta}_{GHG} \mathbf{X}_{GHG} + \hat{\beta}_{SUL} \mathbf{X}_{SUL}
\end{aligned} \tag{29}$$

Hence, $\hat{\beta}_{GHG} = \hat{\beta}_G + \hat{\beta}_{GS}$ gives an estimate of the total greenhouse effect and $\hat{\beta}_{SUL} = \hat{\beta}_{GS}$ gives an estimate of the pure sulfate aerosols effect. Because the regression is done on a linear scale, precisely the same result could be obtained by subtracting $\mathbf{X}_G$ from $\mathbf{X}_{GS}$ initially to give a $\mathbf{X}_{SUL}$ signal to input into the regression. However, this signal will tend to be noisier than $\mathbf{X}_{GS}$ because both of the G and GS signals contain noise.

Next, I define $\hat{\boldsymbol{\beta}}^* = (\hat{\beta}_G, \hat{\beta}_{GS})^T$ and $\boldsymbol{\beta}^* = (\beta_G, \beta_{GS})^T$. From Chapter 2, the distribution of $\hat{\boldsymbol{\beta}}^* \mid \boldsymbol{\beta}^*$ is

$$f(\hat{\boldsymbol{\beta}}^* \mid \boldsymbol{\beta}^*) = \phi(\boldsymbol{\beta}^*, \hat{\boldsymbol{\Sigma}}^*) \tag{30}$$

where $\hat{\boldsymbol{\Sigma}}^* = \tilde{V}(\hat{\boldsymbol{\beta}}^*) = \mathbf{F}_1^T \tilde{\mathbf{C}}_N \mathbf{F}_1$. The above is only approximately true because $\mathbf{C}_N$ is unknown, which is needed to compute the true variance of $\hat{\boldsymbol{\beta}}^*$. Since I am interested in making inferences on β_{GHG} and β_{SUL}, I will have to extract the corresponding information from $\hat{\boldsymbol{\beta}}^*$

and $\hat{\boldsymbol{\Sigma}}^*$. In order to do so, I would need to perform a linear transformation on $\hat{\boldsymbol{\beta}}^*$ and $\boldsymbol{\beta}^*$. If I define

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad \boldsymbol{\beta} = \begin{bmatrix} \beta_{GHG} \\ \beta_{SUL} \end{bmatrix}, \qquad \hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_{GHG} \\ \hat{\beta}_{SUL} \end{bmatrix} \tag{31}$$

then we have,

$$\mathbf{A}\boldsymbol{\beta}^* = \begin{bmatrix} \beta_G + \beta_{GS} \\ \beta_{GS} \end{bmatrix} = \begin{bmatrix} \beta_{GHG} \\ \beta_{SUL} \end{bmatrix} = \boldsymbol{\beta}, \qquad \mathbf{A}\hat{\boldsymbol{\beta}}^* = \begin{bmatrix} \hat{\beta}_{GHG} \\ \hat{\beta}_{SUL} \end{bmatrix} = \hat{\boldsymbol{\beta}}, \tag{32}$$

$$V(\hat{\boldsymbol{\beta}}) = V(\mathbf{A}\hat{\boldsymbol{\beta}}^*) = \mathbf{A}\hat{\boldsymbol{\Sigma}}^*\mathbf{A}^T$$

So the distribution of $\hat{\boldsymbol{\beta}} \mid \boldsymbol{\beta}$ is,

$$f(\hat{\boldsymbol{\beta}} \mid \boldsymbol{\beta}) = \phi(\mathbf{A}\boldsymbol{\beta}^*, \mathbf{A}\hat{\boldsymbol{\Sigma}}^*\mathbf{A}^T) = \phi(\boldsymbol{\beta}, \hat{\boldsymbol{\Sigma}}) \tag{33}$$

where

$$\hat{\boldsymbol{\Sigma}} = \begin{bmatrix} \tilde{V}(\hat{\beta}_G) + \tilde{V}(\hat{\beta}_{GS}) + 2C\tilde{O}V(\hat{\beta}_{GS}, \hat{\beta}_G) & \tilde{V}(\hat{\beta}_{GS}) + C\tilde{O}V(\hat{\beta}_{GS}, \hat{\beta}_G) \\ \tilde{V}(\hat{\beta}_{GS}) + C\tilde{O}V(\hat{\beta}_{GS}, \hat{\beta}_G) & \tilde{V}(\hat{\beta}_{GS}) \end{bmatrix} \tag{34}$$

To account for the noise in signals, $\hat{\boldsymbol{\Sigma}}$ will be inflated by $1 + 1/M$. In the following analysis, M is equal to 4.

4.1 'Optimal fingerprinting'

The residual consistency test statistics for all six periods are shown in Figure 24. The dotted line represents the χ^2 bounds at the 5% significance level. With approximately $k > 15$, the model under-simulates the internal variability. So the maximum number of EOFs retained at which I should trust my analysis would be around 15.

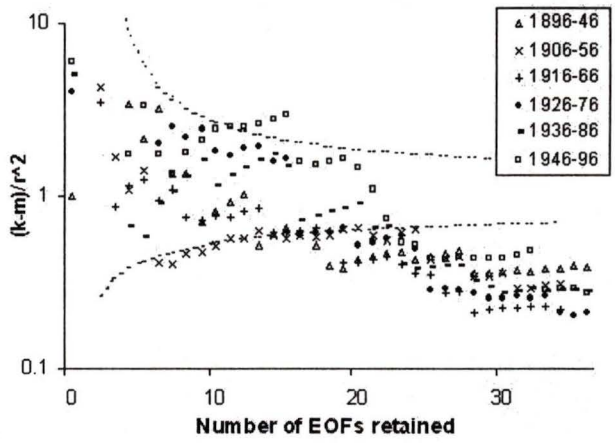

Figure 24: Residual consistency test statistics for different 50 years periods as a function of the number of EOFs retained

Figure 25 shows the 95% confidence ellipsoids for periods 1 and 6 at 15 EOFs retained. The lines indicate the marginal 95% confidence intervals for β_{GHG} and β_{SUL}. The point were the two lines cross indicate the estimate of $\boldsymbol{\beta}$ that is obtained when 15 EOFs are retained.

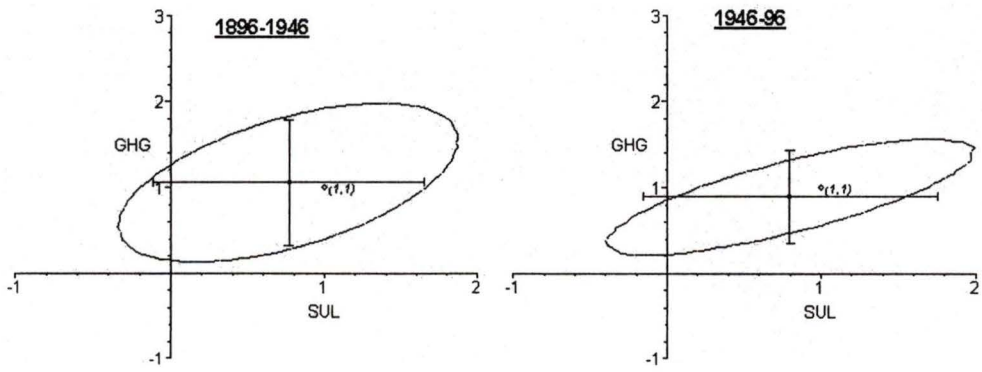

Figure 25: 95% confidence ellipsoids for 1896-1946 and 1946-96 at 15 EOFs retained

Point [0,0] lies outside the ellipsoids for both periods 1 and 6 and so the hypothesis of β_{GHG} and β_{SUL} are both equal to zero can be rejected. Also, the point [1,1] lies within the ellipsoids and hence an attribution claim can be made for the joint greenhouse gases and

sulfate aerosols effects. However, both ellipsoids are fairly large and this makes the attribution result less meaningful. Marginal tests can also be performed for β_{GHG} and β_{SUL}. From Figure 25, the hypothesis $\beta_{GHG} = 0$ can be rejected for both periods 1 and 6 because the marginal confidence intervals for β_{GHG} do not contain zero. The intervals also contain 1 and thus the hypothesis $\beta_{GHG} = 1$ cannot be rejected. On the other hand, the hypothesis that $\beta_{SUL} = 0$ cannot be rejected in either period 1 or 6 as part of the confidence interval lies in the negative range.

When considering the signals jointly, detection is claimed for periods 1 and 6; but when looking at the signals separately, I cannot make an attribution claim for the SUL signal. This happens because i) different **F** critical values are used to construct the confidence region in the joint and the marginal test, and ii) the joint detection test searches for evidence against the hypothesis that both of the amplitudes are zero. So, if one of the amplitude is significantly different from zero, then regardless of what the other amplitude is, I will always make a joint detection claim.

For periods 2 to 5 (not shown), the joint GHG and SUL signals are not detectable. Marginally, GHG is detectable and attributable for period 2, but SUL is not detectable. For periods 3, 4 and 5, the hypothesis of $\beta_{GHG} = 0$ and the hypothesis of $\beta_{SUL} = 0$ cannot be rejected.

Table 5 summarizes my detection and attribution results. The results I obtained are fairly consistent with Tett et al. (1999) (They did not analyze the 1896-49 period). In addition to what I get, they also make a attribution claimed for the marginal SUL signal in period 6. The cause of such discrepancy may possibly due to more EOFs are retained in my analysis.

	GHG and SUL	GHG	SUL
Detection	1, 6	1, 2, 6	
Attribution	1, 6	1, 2, 6	

Table 5: Periods at which detection and attribution is claimed for the corresponding variable(s) at 15 EOFs retained with 'optimal fingerprinting' approach

4.2 Bayesian inference

In the two-signal case, I will introduce a prior analogous to the one used in the one-signal case, in which the prior will give almost equal weight to $\beta_{GHG} \in [0,1]$ and $\beta_{SUL} \in [0,1]$. Mathematically, the prior is defined as a mixture of four bivariate normal distributions,

$$\pi(\boldsymbol{\beta}) = p_1\phi(\boldsymbol{\mu}_1, \boldsymbol{\Lambda}_1) + p_2\phi(\boldsymbol{\mu}_2, \boldsymbol{\Lambda}_2) + p_3\phi(\boldsymbol{\mu}_3, \boldsymbol{\Lambda}_3) + p_4\phi(\boldsymbol{\mu}_4, \boldsymbol{\Lambda}_4) \tag{35}$$

where $\phi(\ ,\ )$ is a bivariate normal pdf and

$$\begin{aligned}
\boldsymbol{\beta} &= \begin{bmatrix} \beta_{GHG} \\ \beta_{SUL} \end{bmatrix}, \quad p_1 = p_2 = p_3 = p_4 = 0.25 \\
\boldsymbol{\mu}_1 &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{\mu}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \boldsymbol{\mu}_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \boldsymbol{\mu}_4 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
\boldsymbol{\Lambda}_1 = \boldsymbol{\Lambda}_2 = \boldsymbol{\Lambda}_3 = \boldsymbol{\Lambda}_4 &= \begin{bmatrix} 0.25 & 0.1 \\ 0.1 & 0.25 \end{bmatrix}
\end{aligned} \tag{36}$$

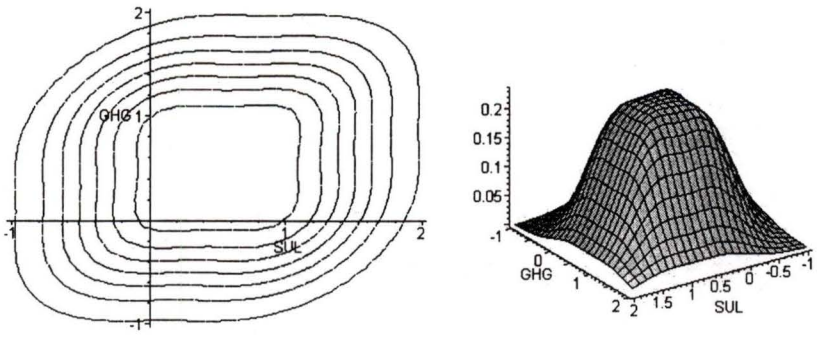

Figure 26: Contour plot and 3D plot of the prior used in two signal analysis

The covariance between β_{GHG} and β_{SUL} is set to 0.1 instead of zero because I believe that they are positively correlated. The reason is, for a particular set of observations, if I get a large amplitude for the GHG signal, then I will expect the amplitude of SUL signal to also be large to offset the GHG signal. From the beginning of this chapter, the distribution of $\hat{\boldsymbol{\beta}} \mid \boldsymbol{\beta}$ is

$$f(\hat{\boldsymbol{\beta}} \mid \boldsymbol{\beta}) = \phi(\boldsymbol{\beta}, \hat{\boldsymbol{\Sigma}}) \tag{37}$$

Using (33), (35) and Bayes' theorem, I obtained the posterior distribution,

$$\pi(\boldsymbol{\beta} \mid \hat{\boldsymbol{\beta}}) = r_1\phi(\boldsymbol{\Gamma}_1\boldsymbol{\omega}_1, \boldsymbol{\Gamma}_1) + r_2\phi(\boldsymbol{\Gamma}_2\boldsymbol{\omega}_2, \boldsymbol{\Gamma}_2) + r_3\phi(\boldsymbol{\Gamma}_3\boldsymbol{\omega}_3, \boldsymbol{\Gamma}_3) + r_4\phi(\boldsymbol{\Gamma}_4\boldsymbol{\omega}_4, \boldsymbol{\Gamma}_4) \tag{38}$$

where for i=1, ..., 4, j=1, ..., 4,

$$\begin{aligned}
&\boldsymbol{\Gamma}_i = (\boldsymbol{\Lambda}_i^{-1} + \hat{\boldsymbol{\Sigma}}^{-1})^{-1} \qquad \boldsymbol{\omega}_i = (\boldsymbol{\Lambda}_i^{-1}\boldsymbol{\mu}_i + \hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\beta}}) \\
&r_i = \left[1 + \sum_{j \neq i} \frac{q_j}{q_i}\right]^{-1} \\
&q_i = \frac{p_i \mid \boldsymbol{\Gamma}_i \mid^{0.5}}{2\pi \mid \boldsymbol{\Lambda}_i \mid^{0.5} \mid \hat{\boldsymbol{\Sigma}} \mid^{0.5}} \exp\left\{-0.5\left[\boldsymbol{\mu}_i^T\boldsymbol{\Lambda}_i^{-1}\boldsymbol{\mu}_i + \hat{\boldsymbol{\beta}}^T\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\beta}} - \boldsymbol{\omega}_i^T\boldsymbol{\Gamma}_i\boldsymbol{\omega}_i\right]\right\}
\end{aligned} \tag{39}$$

Details of the derivation of this posterior are given in Appendix A.2.

I next formalize the criteria for detection and attribution in two-signal case. As in Chapter 3, our criteria for detection is based on a small posterior probability that $\boldsymbol{\beta}$ lies in $\mathcal{D}$ where $\mathcal{D}$ is some neighbourhood of zero, that is,

$$Pr(\boldsymbol{\beta} \in \mathcal{D} \mid \hat{\boldsymbol{\beta}}) \quad \text{is small} \tag{40}$$

and my criteria for attribution is based on a large posterior probability that $\boldsymbol{\beta}$ lies in $\mathcal{A}$ where $\mathcal{A}$ is some neighbourhood of one, that is,

$$Pr(\boldsymbol{\beta} \in \mathcal{A} \mid \hat{\boldsymbol{\beta}}) \quad \text{is large} \tag{41}$$

where the regions $\mathcal{D}$ and $\mathcal{A}$ depend on the variables of interest. Figure 27 shows the non-detection and attribution regions for different signals. Again, Table 2 will be used to quantify the evidence of detection and attribution. Analogous to the one signal case, before making an attribution claim, it is important to examine the posterior to determine whether the posterior has a problematic distribution (for example, a bimodal posterior with peaks that are close in height). I defined three attribution regions $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ for comparison purpose. Again, my decision about whether to make an attribution claim will be based on $\mathcal{A}_2$.

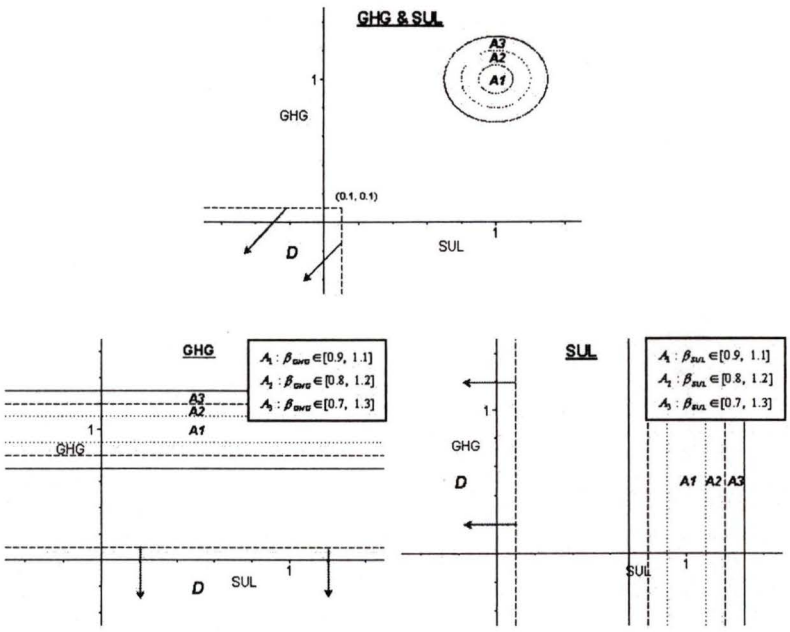

Figure 27: Non-detection and attribution region for different signals

For the joint GHG and SUL signals, I defined the attribution region to be circular rather than rectangular so that the points on the boundary will be the same distance from the point [1,1]. In particular, the radius of $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ are 0.1, 0.2 and 0.3 respectively. The regions for the marginal case are defined to be rectangular because I am only interested in one signal at a time. This means that the 'non-interested' signal can take any value in $\mathbb{R}$. The idea becomes trivial if we look at the following integral for the attribution region of (0.8, 1.2) for the GHG signal.

$$\begin{aligned} & Pr(attribution\ for\ the\ GHG\ signal) \\ &= Pr(\beta_{GHG} \in [0.8, 1.2]) \\ &= \int_{0.8}^{1.2} \int_{-\infty}^{\infty} \pi(\boldsymbol{\beta} \mid \hat{\boldsymbol{\beta}}) d\beta_{SUL} d\beta_{GHG} \end{aligned} \tag{42}$$

Figures 28 and 29 show the posterior distributions for periods 1 and 6 for k=5, 10 and 15. By looking at these figures, we can see that the posterior distributions are well behaved.

In addition, the posterior distributions tend to be more spread out when only a few EOFs are retained ($k = 5$). This is due to the fact that if I retained too few EOFs, the variability in the data cannot be sufficiently explained and hence the variance components in $\hat{\Sigma}$ will be large. This will then cause the posterior distribution to be more spread out. Also, at k=15, the posterior distribution for period 6 tends to lie in the area where both β_{SUL} and β_{GHG} is less than 1. So, it could be the case that the HadCM2 model over-estimates the response to both greenhouse gases and sulfate aerosols.

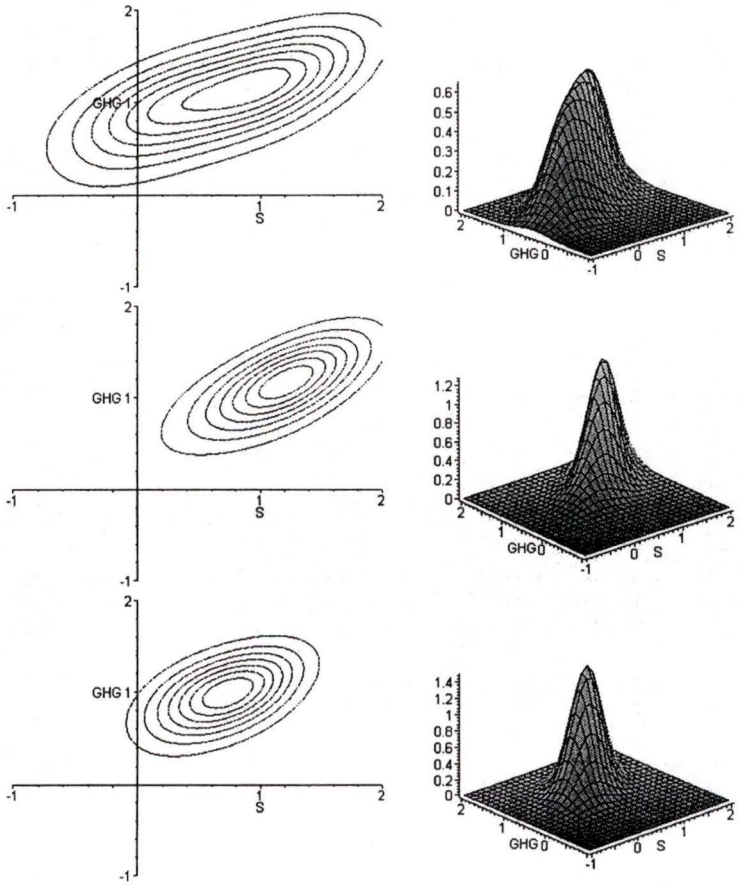

Figure 28: Posterior distribution for period 1 in two-signal case (Top to bottom: 5, 10, 15 EOFs)

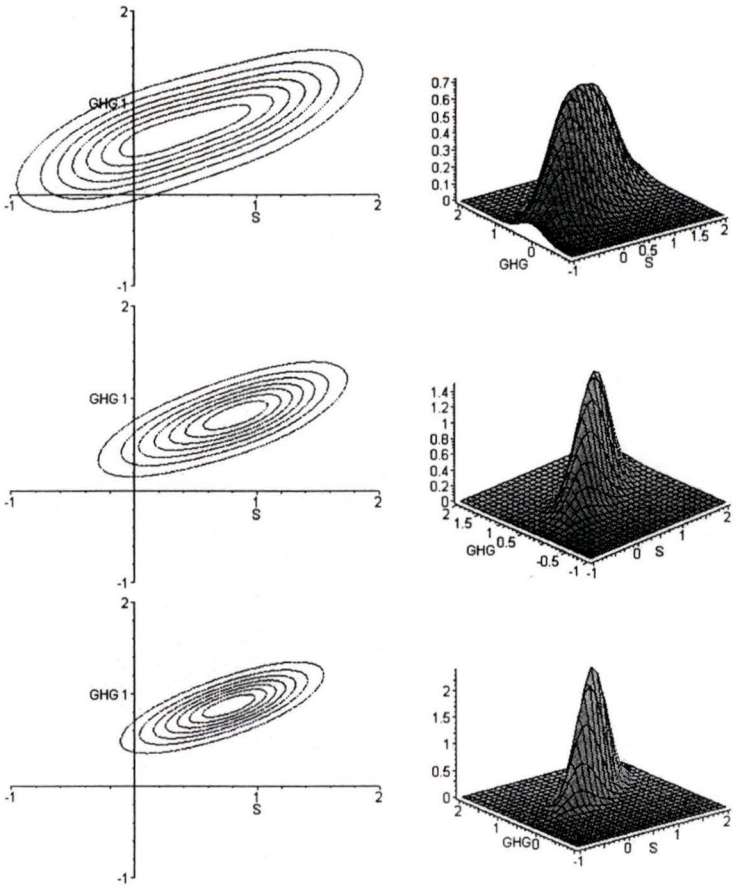

Figure 29: Posterior distribution for period 6 in two-signal case (Top to bottom: 5, 10, 15 EOFs)

Figure 30 shows the posterior probabilities in $\mathcal{D}$, $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ as a function of the number of EOFs retained (k) for all six periods for the joint GHG and SUL signals. In all periods except 3 and 4, the $Pr(\boldsymbol{\beta} \in \mathcal{D} \mid \hat{\boldsymbol{\beta}}) < 0.05$ for k around 15, showing that the likelihood of the GHG and SUL signals have been jointly detected is very large. However, $Pr(\boldsymbol{\beta} \in \mathcal{A}_2 \mid \hat{\beta}) < 0.2$ for all periods, indicating that it would be inappropriate to make a joint attribution claim.

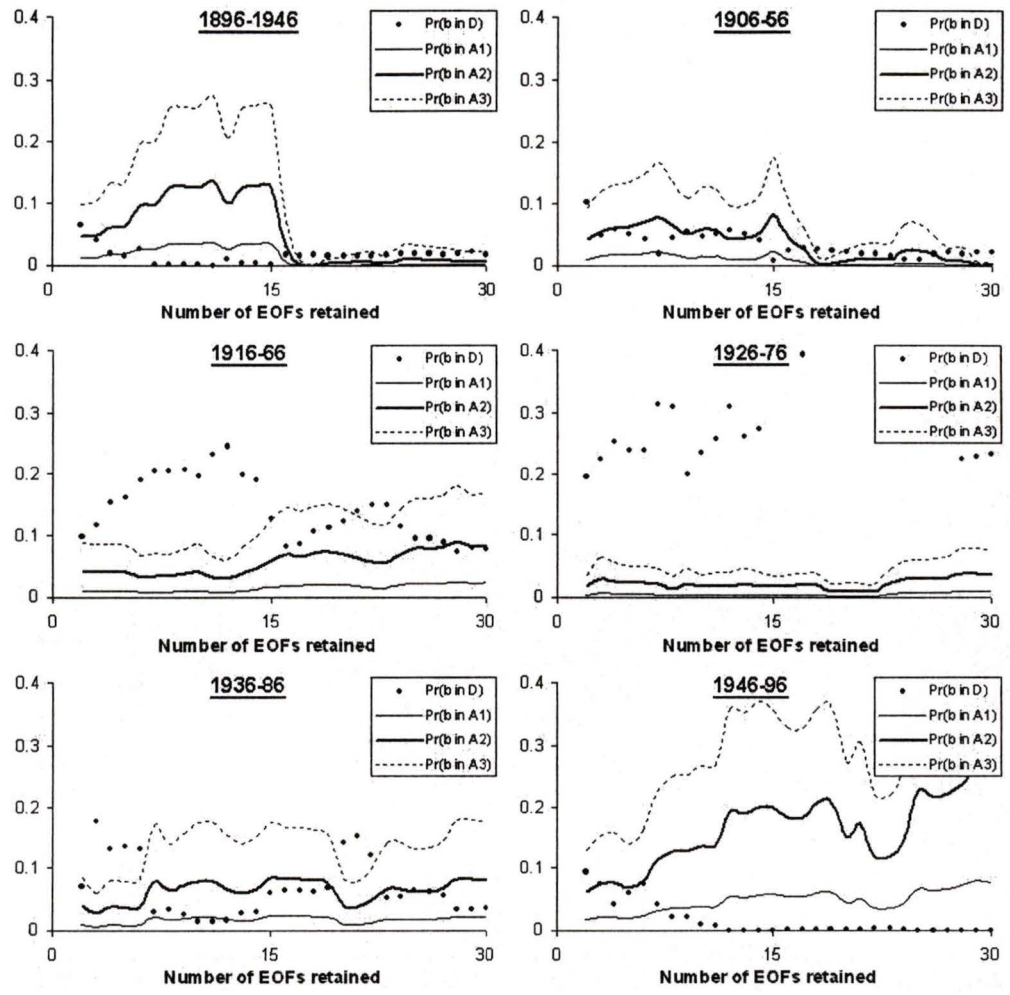

Figure 30: Posterior probabilities for no detection and attribution for the joint GHG and SUL signals

Marginal posterior probabilities for GHG signal and SUL signal are shown in Figures 31 and 32. Table 6 summarizes the detection and attribution results (using $\mathcal{A}_2$) for different signals across the six time periods. Marginally, GHG signal is detectable for all periods except 3 and 4 and SUL signal is detectable in periods 1 and 6. At the same time, the influence of GHG is cautiously attributable in periods 1 and 6, but the influence of SUL is not attributable in any periods.

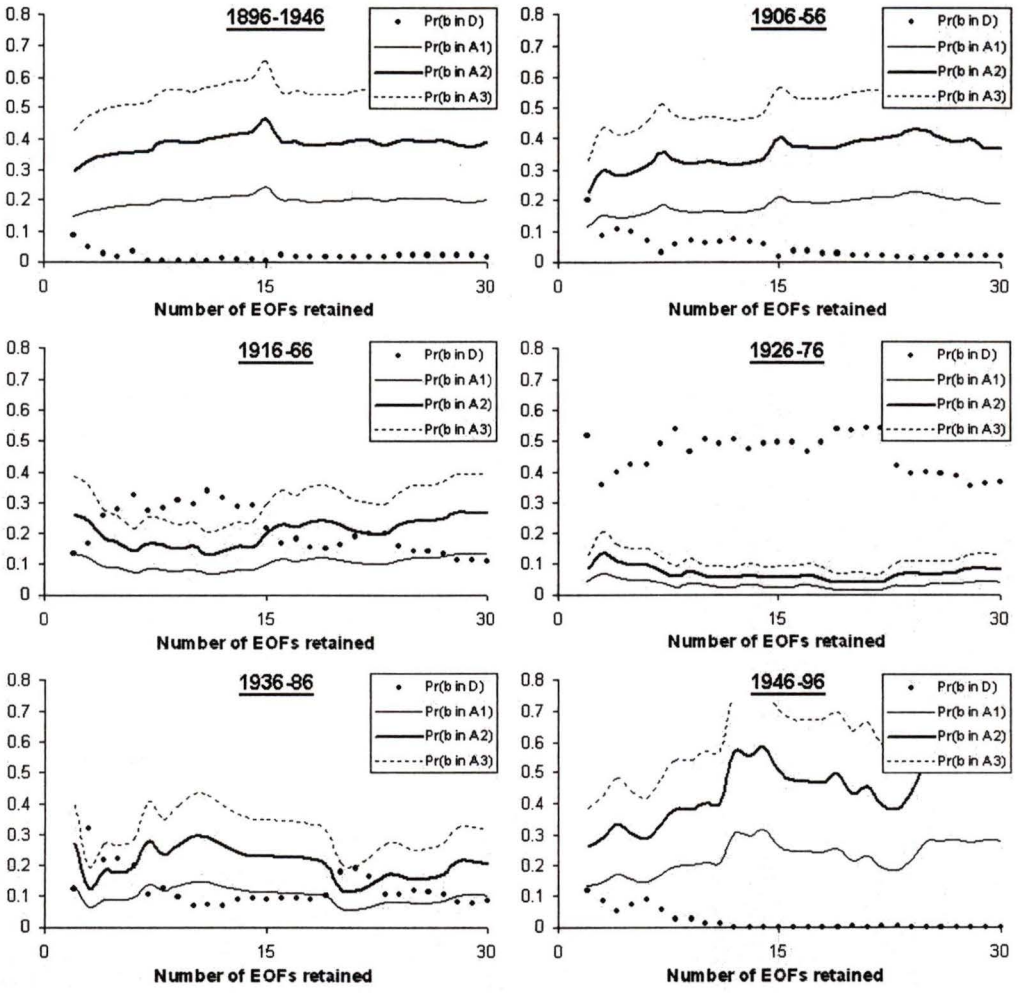

Figure 31: Posterior probabilities for no detection and attribution for GHG signal

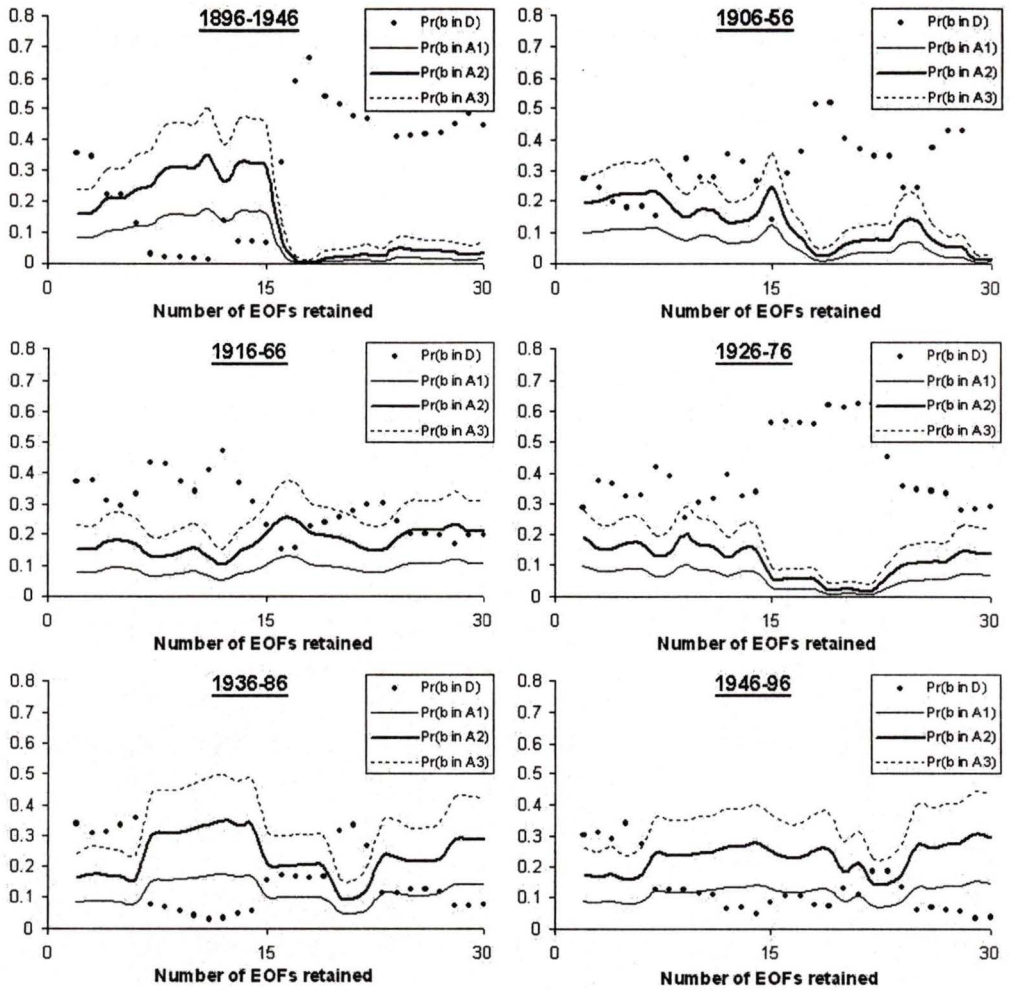

Figure 32: Posterior probabilities for no detection and attribution for SUL signal

Signal	1896-46	1906-56	1916-66	1926-76	1936-86	1946-96
GHG & SUL						
Detection	V.Strong	V.Strong	No	No	Strong	V.Strong
Attribution	No	No			No	No
GHG						
Detection	V.Strong	V.Strong	No	No	Strong	V.Strong
Attribution	Small	No			No	Small
SUL						
Detection	Strong	No	No	No	No	Strong
Attribution	No					No

Table 6: Evidence of detection and attribution across different signals at each of the six time periods

Table 7 shows the attribution results across the different attribution regions. Periods 3 and 4 are omitted from the table because there is no evidence of detection for any of the signals in these two time periods. When using $\mathcal{A}_1$ as the attribution region, there is no evidence for making an attribution claim for any of the signals. On the contrary, with $\mathcal{A}_3$ as the attribution region, we found strong evidence for making attribution claim for GHG signal in some cases. For the joint GHG and SUL signals, no attribution can be made with any of the attribution region.

Region	1896-46	1906-56	1936-86	1946-96
$\mathcal{A}_1$				
GHG & S	No	No	No	No
GHG	No	No		No
S	No			No
$\mathcal{A}_2$				
GHG & S	No	No	No	No
GHG	Small	No		Small
S	No			No
$\mathcal{A}_3$				
GHG & S	No	No	No	No
GHG	Strong	Mild		Strong
S	Small			No

Table 7: Evidence of attribution across the three attribution region for different signals: $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$

In the two-signal case, it is difficult to construct a HPD region for $\boldsymbol{\beta}$ because the posterior is a mixture of four bivariate normal distributions. As a consequence, the region R that has the smallest possible volume in the parameter space could possibly be a region of irregular shape. The volume for such a region may be hard to compute because we will need to evaluate integrals at some possibly complicated equations. So, I will not discuss inference with HPD region in the two-signal case.

From the two-signal Bayesian analysis, the results indicate that the observed change in temperature in periods 1, 2, 5 and 6 are unlikely to be entirely the results of natural climate variations. However, no 'strong' attribution claims are made when considering the signals jointly or marginally.

4.3 Robustness of the two-signal analysis with respect to prior specification

In the one-signal case, the analysis results are robust with respect to the three priors I used and over the class of prior within $\boldsymbol{\Gamma}$. In the two-signal case, due to the number of parameters

involved, I will simply illustrate the robustness of my analysis by comparing the detection and attribution probabilities for several different priors. In general, these priors will be of the same form as (35) and (36) with the following changes.

Prior A	Same as (35) and (36)
Prior B	$p_1 = p_4 = 0.45,\ p_2 = p_3 = 0.05$
Prior C	$p_1 = 0.7,\ p_2 = p_3 = p_4 = 0.1$
Prior D	$p_1 = p_2 = p_3 = 0.1,\ p_4 = 0.7$

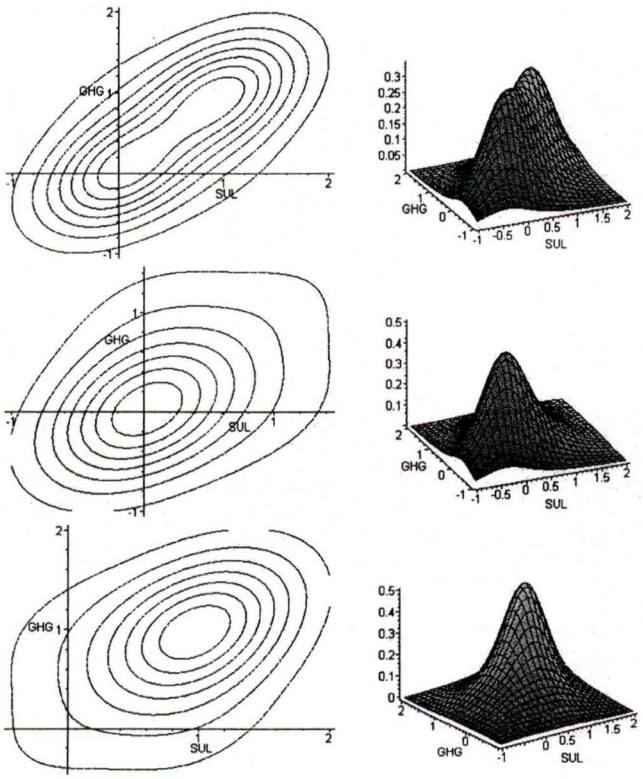

Figure 33: Contour and 3D plot of prior B, C and D (top to bottom)

The selection of the three other priors reflects a range of beliefs regarding the possibilities for detection and attribution. In particular, prior B gives equal weight to the no detection and attributable hypothesis. On the contrary, prior C and D give 70% of their weight to the

no detection hypothesis and to the attributable hypothesis respectively.

For each prior, I computed $Pr(\boldsymbol{\beta} \in \mathcal{D} \mid \hat{\boldsymbol{\beta}})$ and $Pr(\boldsymbol{\beta} \in \mathcal{A}_2 \mid \hat{\boldsymbol{\beta}})$ and recorded the maximum and minimum of these posterior probabilities among the four priors at each time period. Figures 34 and 35 show the results for periods 1 and 6. The dots indicate the posterior probabilities obtained previously with Prior A. At low EOF truncations, the bounds of the detection probability tend to be wide for both periods 1 and 6. Similar results are obtained in the other four periods (not shown). On the other hand, the bounds on the attribution probability are better behaved in all six periods. In general, for $k \approx 15$, the width of the detection probability range is less than 0.1 and it is less than 0.2 for the attribution probability. Thus, my detection and attribution results would change only slightly across the four priors. This suggests my analysis is quite robust with respect to the choice of prior.

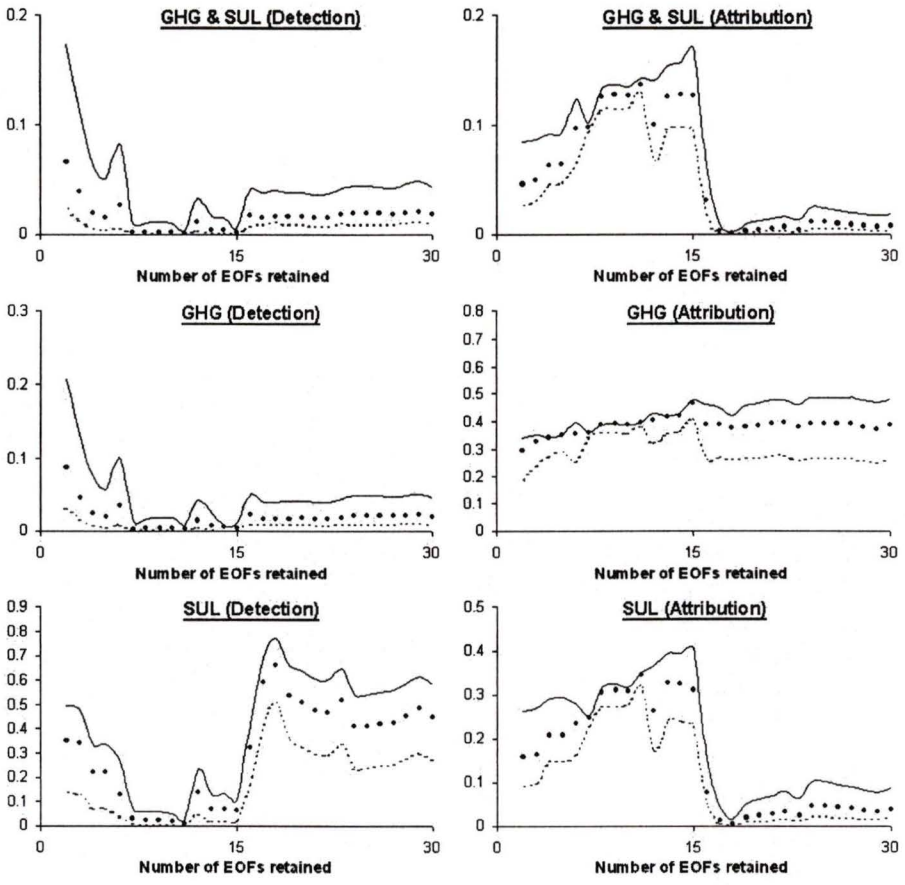

Figure 34: Bounds of posterior no detection and attribution probabilities for period 1

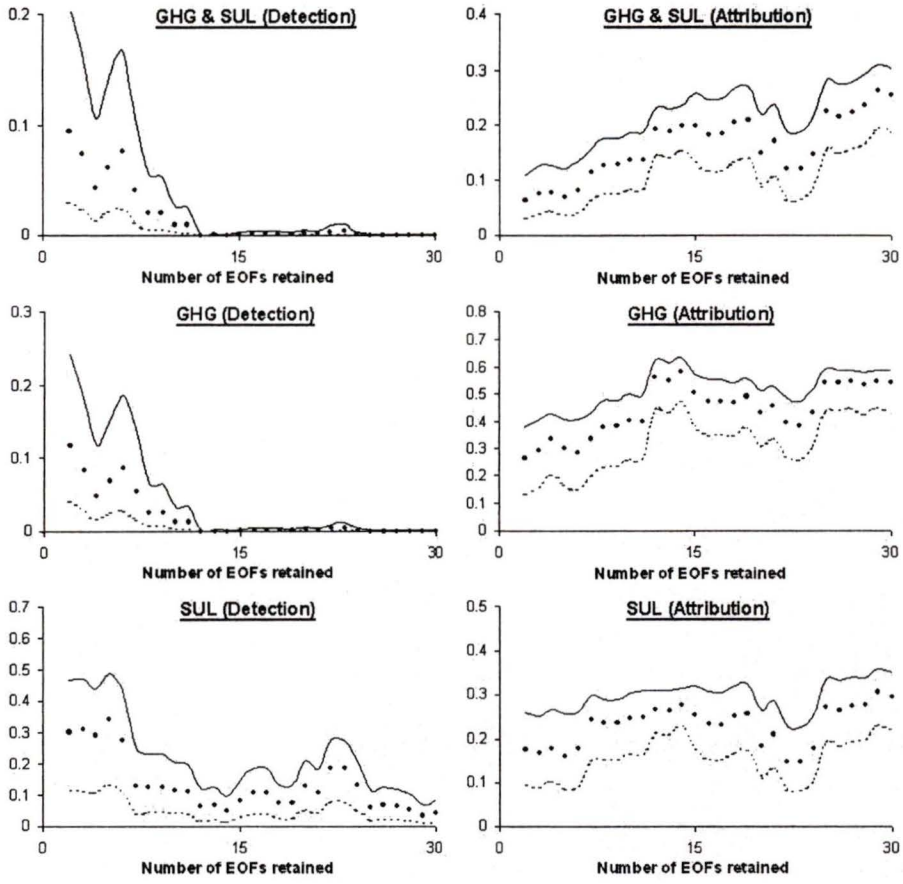

Figure 35: Bounds of posterior no detection and attribution probabilities for period 6

5 Conclusion

I have presented both classical and Bayesian solutions to the climate change detection and attribution problem. The results obtained are very similar across the two approaches. In summary, when considering the combined effect of greenhouse gases and sulfate aerosols, my Bayesian analysis suggests that there is strong evidence for detection in the 1900-49, 1910-59 and 1950-99 periods and there is mild evidence on which to base an attribution claim in 1900-49. When considered jointly, there is strong evidence to claim detection except in 1916-66 and 1926-76. Marginally, SUL signal is detected in 1896-1946 and 1946-96 and GHG signal is detected in 1896-1946, 1906-56 and 1946-96. In the two-signal analysis, no attribution claim (joint or marginal) is made with the Bayesian approach.

Most of the Bayesian detection results presented here support the conclusion of Tett et al. (1999) and other climate change studies. However, my attribution results are quite different from the others; for instants, most of the studies have claimed an attribution for the GS signal in 1950-99. As stated by the IPCC (2001), the warming trend in the first half of the twentieth century could be due to some combination of solar irradiance, increase in greenhouse gases and highly unusual natural variation. The warming in the second half of the century is most likely due to substantial warming from increase in greenhouse gases concentration and partially offset by the cooling from sulfate aerosols. My Bayesian analysis seems to support the IPCC assessment of cause for the second half of the century, in which we have large evidence of detecting the GS signal. On the other hand, for the first half of the century, my Bayesian analysis suggests an attribution claim for the GS signal. However, since I have not look at the influence of other forcing agents, it is questionable that my claim is actually valid.

If we assume greenhouse gases and sulfate aerosols are the only forcing agents that affected the climate, my Bayesian analysis indicates that the combined CGCM1 and CGCM2 model

tend to over-estimate the combined effect of greenhouse gases and sulfate aerosols. It is because most of the posterior probability in 1950-99 is allocated to $0.5 < \beta < 0.9$. My analysis also suggests that the HadCM2 model tends to over-estimate both greenhouse gases and sulfate aerosols effect because most the posterior distribution in 1946-96 is lay in an area where β_{GHG} and β_{SUL} is both less than unity.

In my analysis, I computed the posterior probability for four regions: $\mathcal{D}, \mathcal{A}_1, \mathcal{A}_2$ and $\mathcal{A}_3$. In fact, it may be useful to compute posterior probabilities for regions that correspond to the case when the model under- or over-estimates the response to external forcing. These probabilities will help us to assess the extent to which the climate model may over- or under-simulate the external forcing.

In the Bayesian results I presented, most of the attribution claim is made with caution. Near-certain attribution claim is not possible because the posterior is fairly spread out in most cases due to the large uncertainty in $\hat{\boldsymbol{\beta}}$. Even if the posterior is somewhat centered at $\underline{1}$, it would still not be sensible to make an attribution claim due to the spread of the posterior. So before making an attribution claim, it is necessary to examine the posterior distribution carefully rather than just looking at simple summaries of the posterior distribution.

The two main advantages of using a Bayesian analysis are the incorporation of prior beliefs in the analysis and direct probability assessment of hypotheses. But at the same time, one may argue that the inclusion of prior beliefs adds subjectivity to the analysis. However, I have demonstrated that results are insensitive to the choice of prior when the prior is chosen in a reasonable manner. That is, the posterior density is somewhat dependent on the distribution of the data and it is not very sensitive to my prior beliefs. In other words, my results are mainly driven by the estimated amplitudes obtained at the regression stage and are less affected by my subjective beliefs on what the true amplitudes are.

In the Bayesian analysis we presented, calculations of the posterior parameters and probabilities are not complicated. Some computationally intensive procedures, like finding the IR interval in one-signal case and evaluating the posterior probabilities in two-signal case, are handled numerically with the Netwon-Raphson scheme and numerical integration respectively.

I would like to stress that this thesis is concentrated on illustrating a Bayesian solution to the detection and attribution problem. It is important to look at the influence of other potential forcing agents, such as solar output variations, before one can confidently identify which forcing agents caused the observed changes in the twentieth century temperature.

As in the classical approach, the findings from Bayesian analysis suffer from large uncertainty in the estimate of $\boldsymbol{\beta}$. The uncertainty is large because i) there are a lot of missing data and ii) the control run used in the analysis is limited in length. The large uncertainty will cause the attribution claim less meaningful and makes near-certain attribution claim impossible. A possible solution to this would be to construct a fully Bayesian regression model, which might reduce the uncertainty when making an attribution claim. In this type of model, it involves formulating a prior distribution for the noise structure of $\mathbf{y}$ to be inputted into the regression. By giving the noise some probability structure, there is hope to reduce the uncertainty in $\boldsymbol{\beta}$. However, due to large dimensionality of the problem (dimension of $\mathbf{y}$ is large), it will be a daunting task to come up with such a probability model. Even if one could come up with such a probability model, it is still questionable that the model can 'significantly' represent the nature climate variation in the real world.

Throughout the analysis, I have inflated the variance of $\hat{\beta}$ by $1 + 1/M$ to account for the noise contained in the signals. Nevertheless, this inflation factor only makes the overall analysis more conservative and does not eliminate the bias caused by the noise contained in the signals. An algorithm that accounts for noise in both the observed data and signals is

readily available. One algorithm suggested is total least squares (see for example, Allen, 1999; Allen et al, 2000), in which the noise structure in the signal is assumed to be the same as that in the observations. Then, the same pre-whitening operator, $\mathbf{P}$ can be applied to both of the observations and the signals to estimate $\hat{\boldsymbol{\beta}}$. Other possible solutions, such as the use of error-in-variable models (See for example, Fuller, 1987), may be applicable in the climate change detection setting. In an error-in-variable models, both $\mathbf{y}$ and $\mathbf{X}$ are considered to have additive noise and a variety of regression techniques are applied to estimate β. Different from total least squares, error-in-variable model allows the noise structure in $\mathbf{y}$ and $\mathbf{X}$ to be different. More work will have to be done before we can bring these techniques into the climate change detection setting.

In short, the Bayesian analyses presented here are preliminary in nature. Further investigation should be done with other climate models and with other forcing agents. Other extensions, such as incorporating three or more signals into the analysis at the same time, might be challenging, as we would need to define a prior that has four or more dimensions. In the Bayesian literature, a hierarchical model can be used to reduce the dimensions of the prior. A possible hierarchical model that can be used in detection and attribution studies is

$$\pi(\beta_1, \beta_2, ..., \beta_m \mid \hat{\boldsymbol{\beta}}) \propto f(\hat{\beta} \mid \beta_1, \beta_2, ..., \beta_m)\pi(\beta_1 \mid \beta_2, \beta_3, ..., \beta_m)\pi(\beta_2 \mid \beta_3, \beta_4, ..., \beta_m)...\pi(\beta_m) \tag{43}$$

where β_i, $i = 1, ..., m$ is the true amplitude for the m^{th} signal and $\pi()$ distribution on the right hand side of the equation are prior distribution for the corresponding β_i given the other β_is. To make inferences for this model, one could still apply criteria (23) and (24) to make detection and attribution claims. But before making any inferences, it is necessary to obtain the posterior distribution $\pi(\beta_1, \beta_2, ..., \beta_m \mid \hat{\boldsymbol{\beta}})$. In order to do so, one has to define all the distributions on the right hand side of (43) and this could be challenging if m is large.

Nevertheless, if we do not have any knowledge on these priors, we can simply use priors that are noninformative.

6 References

Allen, M.R. (1999): Algorithms for climate change detection, *Techn Report*, Rutherford Appleton Laboratory, OX11 0QX, U.K.

Allen, M.R., N.P. Gillett, G. Hegerl, R. Schnur, P.A. Stott, G. Boer, T.L. Delworth, G.S. Jones, J.F.B. Mitchell and T.P. Barnett (2000): Quantifying anthropogenic influence on recent near-surface temperature change, *Surveys in Geophysics*, to appear. 2002

Allen, M.R., and S.F.B. Tett (1999): Checking for model consistency in optimal fingerprinting. *Climate Dynamics*, **15**, 419-434.

Berger, J.O. (1985): *Statistical decision theory and Bayesian analysis*, Springer-Verlag.

Berliner, L.M., R.A. Levine and D.J. Shea (2000): Bayesian climate change assessment. *Journal of Climate*, **13**, 3805-3820.

Flato, G.M. and G.J. Boer (2001): Warming asymmetry in climate change simulations. *Geophysical Research Letter*, **28**, 195-198.

Flato, G.M., G.J. Boer, W.G. Lee, N.A. McFarlane, D. Ramsden, M.C. Reader and A.J. Weaver (2000): The Canadian Centre for Climate Modelling and Analysis Global Coupled Model and its climate. *Climate Dynamics*, **16**, 427-450.

Fuller, W.A. (1987): *Measurement error models*, Wiley.

Hasselmann, K. (1993):Optimal fingerprinting for the detection and attribution of climate change. *Journal of Climate*, **6**, 1957-1971.

Hasselmann, K. (1997): On multifingerprint detection and attribution of anthropogenic climate change. *Climate Dynamics*, **13**, 601-611.

Hasselmann, K. (1998): Conventional and Bayesian approach to climate- change detection and attribution. *Quarterly Journal of the Meteorological Society*, **124**, 2541-2565.

Hegerl, G., K. Hasselmann, U. Cubasch, J.F.B. Mitchell, E. Roeckner, R. Voss and J. Waszkewitz (1997): On multi-fingerprint detection and attribution of greenhouse gas and aerosol forced climate change. *Climate Dynamics*, **13**, 613-634.

Hegerl, G., P.H. Jones and T.P. Barnett (2000): Effect of observational sampling error on the detection of anthropogenic climate change. *Journal of Climate*, **14**, 198-207.

IPCC (2001):*Climate Change 2001: The Scientific Basis.* J.T. Houghton, Y. Ding, D.J. Griggs, M. Noguer, P.J. can der Linden, X. Dai, K. Maskerr and C.A. Johnson(eds.), Cambridge University Press, Cambridge, United Kingdom and New York

Johns, T.C., R.E. Carnell, J.F.C rossley, J.M. Gregory, J.F.B. Mitchell, C.A. Senior, S.F.B. Tett, R.A. Wood (1997): The second Hadley Centre coupled ocean-atmosphere GCM: model description, spinup and validation. *Climate Dynamics*, **13**, 103-134.

Jones, P.H., T.J. Osborn, K.R. Briffa, C.K. Folland, E.B. Horton, L.V. Alexander, D.E. Parker and N.A. Rayer (2001): Adjusting for sampling density in grid box land and ocean surface temperature time series. *Journal of Geophysical Research*, **106**, 3371-3380.

Leroy, S.S. (1998): Detecting climate signals: Some Bayesian aspects. *Journal of Climate*, **11**, 640-651.

Santer,B.D., K.E. Taylor, J.E. Penner, T.M.L. Wigley, U. Cubasch and P.D. Jones (1995): Towards the detection and attribution of an anthropogenic effect on climate. *Climate Dynamics*, **12**, 77-100.

Stott, P.A., S.F.B. Tett, G.S. Jones, M.R. Allen, W.J. Ingram, and J.F.B. Mitchell (2001): Attribution of twentieth century temperature change to natural and anthropogenic causes. *Climate Dynamics*, **17**, 1-21.

Tett S.F.B., P.A. Stott, M.R. Allen, W.J. Ingram, and J.F.B. Mitchell (1999): Causes of twentieth century temperature change near the earth's surface. *Nature*, **399**, 569-572.

Zwiers, F.W. and X. Zhang (2003): Towards regional scale climate change detection. *Journal of Climate*, **16**, 793-797.

A Appendix

A.1 Derivation of the posterior for one-signal analysis

The following derivation is expanded from the steps layout in Berger (1985, p127, 128, 206). We have defined $\pi(\beta) = p\phi(0,\tau^2)+(1-p)\phi(\mu_A,\tau_A^2)$ and $f(\hat{\beta} \mid \beta) = \phi(\beta,\hat{\sigma}^2)$. So, the posterior distribution is

$$\begin{aligned} &\pi(\beta \mid \hat{\beta}) \\ &= \frac{f(\hat{\beta} \mid \beta)\pi(\beta)}{\int f(\hat{\beta} \mid \beta)\pi(\beta)d\beta} \\ &= \frac{p\phi(0,\tau^2)\phi(\beta,\hat{\sigma}^2) + (1-p)\phi(\mu_A,\tau_A^2)\phi(\beta,\hat{\sigma}^2)}{\int f(\hat{\beta} \mid \beta)\pi(\beta)d\beta} \end{aligned} \tag{44}$$

Now, $(1-p)\phi(\mu_A,\tau_A^2)\phi(\beta,\hat{\sigma}^2)$

$$\begin{aligned} &= \frac{1-p}{2\pi\hat{\sigma}\tau_A} \exp\left\{-0.5\left[\frac{(\hat{\beta}-\beta)^2}{\hat{\sigma}^2} + \frac{(\beta-\mu_A)^2}{\tau_A^2}\right]\right\} \\ &= \frac{1-p}{2\pi\hat{\sigma}\tau_A} \exp\left\{-0.5\left[\frac{\hat{\beta}^2}{\hat{\sigma}^2} + \frac{\mu_A^2}{\tau_A^2}\right]\right\} \exp\left\{-0.5\left[\beta^2\left(\frac{1}{\hat{\sigma}^2}+\frac{1}{\tau_A^2}\right) - 2\beta\left(\frac{\hat{\beta}}{\hat{\sigma}^2}+\frac{\mu_A}{\tau_A^2}\right)\right]\right\} \end{aligned} \tag{45}$$

Then we let $\rho = \frac{1}{\hat{\sigma}^2} + \frac{1}{\tau_A^2} = \frac{\tau_A^2+\hat{\sigma}^2}{\tau_A^2\hat{\sigma}^2}$, we have

$$\begin{aligned} &= \frac{1-p}{2\pi\hat{\sigma}\tau_A} \exp\left\{-0.5\left[\frac{\hat{\beta}^2}{\hat{\sigma}^2} + \frac{\mu_A^2}{\tau_A^2}\right]\right\} \exp\left\{-0.5\rho\left[\beta^2 - 2\frac{\beta}{\rho}\left(\frac{\hat{\beta}}{\hat{\sigma}^2}+\frac{\mu_A}{\tau_A^2}\right)\right]\right\} \\ &= \frac{1-p}{\sqrt{(2\pi\rho)}\hat{\sigma}\tau_A} \exp\left\{-0.5\left[\frac{\hat{\beta}^2}{\hat{\sigma}^2} + \frac{\mu_A^2}{\tau_A^2} - \frac{1}{\rho}\left(\frac{\hat{\beta}}{\hat{\sigma}^2}+\frac{\mu_A}{\tau_A^2}\right)^2\right]\right\} \phi\left(\frac{\hat{\beta}\tau_A^2+\mu_A\hat{\sigma}^2}{\tau_A^2+\hat{\sigma}^2}, \rho^{-1}\right) \end{aligned} \tag{46}$$

Expand the terms inside the exponential and cancel out the common terms, we have

$$= \frac{1-p}{\sqrt{2\pi(\tau_A^2+\hat{\sigma}^2)}} \exp\left\{-0.5\left[\frac{(\hat{\beta}-\mu_A)^2}{(\tau_A^2+\hat{\sigma}^2)^2}\right]\right\} \phi\left(\frac{\hat{\beta}\tau_A^2+\mu_A\hat{\sigma}^2}{\tau_A^2+\hat{\sigma}^2}, \frac{\tau_A^2\hat{\sigma}^2}{\tau_A^2+\hat{\sigma}^2}\right) \tag{47}$$

Similarly, the first term in the numerator of the posterior is

$$\frac{p}{\sqrt{2\pi(\tau^2+\hat{\sigma}^2)}}\exp\left\{-0.5\left[\frac{\hat{\beta}^2}{(\tau^2+\hat{\sigma}^2)^2}\right]\right\}\phi\left(\frac{\hat{\beta}\tau^2}{\tau^2+\hat{\sigma}^2},\frac{\tau^2\hat{\sigma}^2}{\tau^2+\hat{\sigma}^2}\right) \tag{48}$$

Now we let

$$\begin{aligned} r_1 &= \frac{p}{\sqrt{2\pi(\tau^2+\hat{\sigma}^2)}}\exp\left\{-0.5\left[\frac{\hat{\beta}^2}{(\tau^2+\hat{\sigma}^2)^2}\right]\right\} \\ r_2 &= \frac{1-p}{\sqrt{2\pi(\tau_A^2+\hat{\sigma}^2)}}\exp\left\{-0.5\left[\frac{(\hat{\beta}-\mu_A)^2}{(\tau_A^2+\hat{\sigma}^2)^2}\right]\right\} \end{aligned} \tag{49}$$

Then the posterior is

$$\begin{aligned} &\pi(\beta \mid \hat{\beta}) \\ &= \frac{r_1\phi\left(\frac{\hat{\beta}\tau^2}{\tau^2+\hat{\sigma}^2},\frac{\tau^2\hat{\sigma}^2}{\tau^2+\hat{\sigma}^2}\right)+r_2\phi\left(\frac{\hat{\beta}\tau^2}{\tau^2+\hat{\sigma}^2},\frac{\tau^2\hat{\sigma}^2}{\tau^2+\hat{\sigma}^2}\right)}{r_1+r_2} \\ &= q\phi\left(\frac{\hat{\beta}\tau^2}{\tau^2+\hat{\sigma}^2},\frac{\tau^2\hat{\sigma}^2}{\tau^2+\hat{\sigma}^2}\right)+(1-q)\phi\left(\frac{\hat{\beta}\tau^2}{\tau^2+\hat{\sigma}^2},\frac{\tau^2\hat{\sigma}^2}{\tau^2+\hat{\sigma}^2}\right) \end{aligned} \tag{50}$$

where

$$\begin{aligned} q &= \frac{r_1}{r_1+r_2} = \left(\frac{r_1+r_2}{r_1}\right)^{-1} \\ &= \left(1+\frac{r_2}{r_1}\right)^{-1} \\ &= \left(1+\frac{1-p}{p}\sqrt{\frac{\tau^2+\hat{\sigma}^2}{\tau_A^2+\hat{\sigma}^2}}\quad\exp\left\{-0.5\left[\frac{(\hat{\beta}-\mu_A)^2}{\tau_A^2+\hat{\sigma}^2}-\frac{\hat{\beta}^2}{\tau^2+\hat{\sigma}^2}\right]\right\}\right)^{-1} \end{aligned} \tag{51}$$

A.2 Derivation of the posterior for two-signal analysis

We defined

$$\pi(\boldsymbol{\beta}) = p_1\phi(\boldsymbol{\mu}_1, \boldsymbol{\Lambda}_1) + p_2\phi(\boldsymbol{\mu}_2, \boldsymbol{\Lambda}_2) + p_3\phi(\boldsymbol{\mu}_3, \boldsymbol{\Lambda}_3) + p_4\phi(\boldsymbol{\mu}_4, \boldsymbol{\Lambda}_4) \tag{52}$$

$$f(\hat{\boldsymbol{\beta}} \mid \boldsymbol{\beta}) = \phi(\boldsymbol{\beta}, \hat{\boldsymbol{\Sigma}}) \tag{53}$$

where $\phi(,)$ is a multivariate normal pdf. And

$$\begin{aligned}
&\pi(\boldsymbol{\beta} \mid \hat{\boldsymbol{\beta}}) \\
&= \frac{f(\hat{\boldsymbol{\beta}} \mid \boldsymbol{\beta})\pi(\boldsymbol{\beta})}{\int f(\hat{\boldsymbol{\beta}} \mid \boldsymbol{\beta})\pi(\boldsymbol{\beta})d\boldsymbol{\beta}} \\
&= \frac{p_1\phi(\boldsymbol{\mu}_1, \boldsymbol{\Lambda}_1)\phi(\boldsymbol{\beta}, \hat{\boldsymbol{\Sigma}}) + p_2\phi(\boldsymbol{\mu}_2, \boldsymbol{\Lambda}_2)\phi(\boldsymbol{\beta}, \hat{\boldsymbol{\Sigma}}) + p_3\phi(\boldsymbol{\mu}_3, \boldsymbol{\Lambda}_3)\phi(\boldsymbol{\beta}, \hat{\boldsymbol{\Sigma}}) + p_4\phi(\boldsymbol{\mu}_4, \boldsymbol{\Lambda}_4)\phi(\boldsymbol{\beta}, \hat{\boldsymbol{\Sigma}})}{\int f(\hat{\boldsymbol{\beta}} \mid \boldsymbol{\beta})\pi(\boldsymbol{\beta})d\boldsymbol{\beta}}
\end{aligned} \tag{54}$$

Now,

$$\begin{aligned}
&p_1\phi(\boldsymbol{\mu}_1, \boldsymbol{\Lambda}_1)\phi(\boldsymbol{\beta}, \hat{\boldsymbol{\Sigma}}) \\
&= \frac{p_1}{4\pi^2 \mid \boldsymbol{\Lambda}_1 \mid^{0.5} \mid \hat{\boldsymbol{\Sigma}} \mid^{0.5}} \exp\left\{-0.5\left[(\boldsymbol{\beta} - \boldsymbol{\mu}_1)^T\boldsymbol{\Lambda}_1^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu}_1) + (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T\hat{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\right]\right\} \\
&= \frac{p_1}{4\pi^2 \mid \boldsymbol{\Lambda}_1 \mid^{0.5} \mid \hat{\boldsymbol{\Sigma}} \mid^{0.5}} \exp\left\{-0.5\left[\boldsymbol{\mu}_1^T\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\mu}_1 + \hat{\boldsymbol{\beta}}^T\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\beta}}\right]\right\} \\
&\quad \times \exp\left\{-0.5\left[\boldsymbol{\beta}^T\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\beta} - \boldsymbol{\mu}_1^T\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\beta} - \boldsymbol{\beta}^T\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\mu}_1 - \hat{\boldsymbol{\beta}}^T\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\beta} - \boldsymbol{\beta}^T\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\beta}} + \boldsymbol{\beta}^T\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\beta}\right]\right\} \\
&= \frac{p_1}{4\pi^2 \mid \boldsymbol{\Lambda}_1 \mid^{0.5} \mid \hat{\boldsymbol{\Sigma}} \mid^{0.5}} \exp\left\{-0.5\left[\boldsymbol{\mu}_1^T\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\mu}_1 + \hat{\boldsymbol{\beta}}^T\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\beta}} - \boldsymbol{\omega}_1^T\boldsymbol{\Gamma}_1\boldsymbol{\omega}_1\right]\right\} \\
&\quad \times \exp\left\{-0.5(\boldsymbol{\beta} - \boldsymbol{\Gamma}_1\boldsymbol{\omega}_1)^T\boldsymbol{\Gamma}_1^{-1}(\boldsymbol{\beta} - \boldsymbol{\Gamma}_1\boldsymbol{\omega}_1)\right\} \\
&= \frac{p_1 \mid \boldsymbol{\Gamma}_1 \mid^{0.5}}{2\pi \mid \boldsymbol{\Lambda}_1 \mid^{0.5} \mid \hat{\boldsymbol{\Sigma}} \mid^{0.5}} \exp\left\{-0.5\left[\boldsymbol{\mu}_1^T\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\mu}_1 + \hat{\boldsymbol{\beta}}^T\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\beta}} - \boldsymbol{\omega}_1^T\boldsymbol{\Gamma}_1\boldsymbol{\omega}_1\right]\right\}\phi(\boldsymbol{\Gamma}_1\boldsymbol{\omega}_1, \boldsymbol{\Gamma}_1) \\
&= q_1\phi(\boldsymbol{\Gamma}_1\boldsymbol{\omega}_1, \boldsymbol{\Gamma}_1)
\end{aligned} \tag{55}$$

where $\boldsymbol{\Gamma}_1 = (\boldsymbol{\Lambda}_1^{-1} + \hat{\boldsymbol{\Sigma}}^{-1})^{-1}$ and $\boldsymbol{\omega}_1 = (\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\mu}_1 + \hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\beta}})$. The other three components in the numerator are calculated in a similar manner. So,

$$\begin{aligned}
&\pi(\boldsymbol{\beta} \mid \hat{\boldsymbol{\beta}}) \\
&= \frac{q_1\phi(\boldsymbol{\Gamma}_1\boldsymbol{\omega}_1, \boldsymbol{\Gamma}_1) + q_2\phi(\boldsymbol{\Gamma}_2\boldsymbol{\omega}_2, \boldsymbol{\Gamma}_2) + q_3\phi(\boldsymbol{\Gamma}_3\boldsymbol{\omega}_3, \boldsymbol{\Gamma}_3) + q_4\phi(\boldsymbol{\Gamma}_4\boldsymbol{\omega}_4, \boldsymbol{\Gamma}_4)}{q_1 + q_2 + q_3 + q_4} \\
&= r_1\phi(\boldsymbol{\Gamma}_1\boldsymbol{\omega}_1, \boldsymbol{\Gamma}_1) + r_2\phi(\boldsymbol{\Gamma}_2\boldsymbol{\omega}_2, \boldsymbol{\Gamma}_2) + r_2\phi(\boldsymbol{\Gamma}_3\boldsymbol{\omega}_3, \boldsymbol{\Gamma}_3) + r_4\phi(\boldsymbol{\Gamma}_4\boldsymbol{\omega}_4, \boldsymbol{\Gamma}_4)
\end{aligned} \tag{56}$$

where

$$\begin{aligned}
q_i &= \frac{p_i \mid \boldsymbol{\Gamma}_i \mid^{0.5}}{2\pi \mid \boldsymbol{\Lambda}_i \mid^{0.5} \mid \hat{\boldsymbol{\Sigma}} \mid^{0.5}} \exp\left\{-0.5\left[\boldsymbol{\mu}_i^T\boldsymbol{\Lambda}_i^{-1}\boldsymbol{\mu}_i + \hat{\boldsymbol{\beta}}^T\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\beta}} - \boldsymbol{\omega}_i^T\boldsymbol{\Gamma}_i\boldsymbol{\omega}_i\right]\right\} \\
r_i &= \left[1 + \sum_{j \neq i} \frac{q_j}{q_i}\right]^{-1}
\end{aligned} \tag{57}$$

VITA

Surname: Lee Given Names: Chun Kit (Terry)

Place of Birth: Hong Kong, China

Education Institutions Attended:

University of Victoria	2001-2003
Simon Fraser University	1998-2001
Capilano College	1997-1998

Degree Awarded:

B.Sc (Double Major in Statistics and Economics)
Simon Fraser University, 2001

Honours and Awards:

Simon Fraser University Honor Roll
Spring and Summer 2001
Graduate Fellowship
Department of Mathematics and Statistics, University of Victoria, 2001

Title of Thesis:

Identifying Anthropogenic Causes of the Observed Twentieth Century Surface Temperature Change: Frequentist and Bayesian Approach

Author:

[redacted]

Chun Kit (Terry) Lee
April 10, 2003