

A SYSTEMATIC FRAMEWORK FOR IMPROVING ESTIMATES OF  
ANTHROPOGENIC AEROSOL COOLING

by

Xiaoqin Yan  
A Dissertation  
Submitted to the  
Graduate Faculty  
of  
George Mason University  
In Partial fulfillment of  
The Requirements for the Degree  
of  
Doctor of Philosophy  
Climate Dynamics

Committee:

_____	Dr. Timothy DelSole, Dissertation Director
_____	Dr. Edwin K. Schneider, Committee Member
_____	Dr. Barry A. Klinger, Committee Member
_____	Dr. James Gentle, Committee Member
_____	Dr. Edwin K. Schneider, Department Chair
_____	Dr. Donna M. Fox, Associate Dean, Office of Student Affairs & Special Programs, College of Science
_____	Dr. Peggy Agouris, Dean College of Science
Date: _____	Spring Semester 2016 George Mason University Fairfax, VA

A Systematic Framework for Improving Estimates of  
Anthropogenic Aerosol Cooling

A dissertation submitted in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy at George Mason University

By

Xiaoqin Yan  
Master of Science  
University of Northern British Columbia, 2011  
Bachelor of Science  
Ocean University of China, 2007

Director: Dr. Timothy DelSole, Professor  
Department of Climate Dynamics

Spring Semester 2016  
George Mason University  
Fairfax, VA

Copyright © 2016 by Xiaoqin Yan  
All Rights Reserved

## Dedication

I dedicate this dissertation to my parents who support my study overseas unconditionally.

## Acknowledgments

I would like to thank the following people who made this thesis possible.

First and foremost, I am very thankful for my supervisor, Dr. Timothy DelSole, whose kindness and patience inspired me to join this program. He then provided me invaluable guidance and inspirational encouragement throughout the whole period of my study. His profound knowledge helped me to develop a solid understanding of this subject. And his optimism supported me through times when I might have given up. Without him, this thesis would not have been completed.

I am very grateful to my committee members Dr. Edwin Schneider, Dr. Barry Klinger and Dr. James Gentle for their constructive suggestions and support.

I am very grateful to Dr. Michael Tippett, for valuable suggestions and discussions.

I would like to acknowledge professors in climate dynamics for their good instructions in course teaching.

I would like to acknowledge Jennifer Adams and Thomas Wakefield for their technical support, and many research scientists in COLA for advice and help on research,

I would like to acknowledge my fellow students for their friendliness in climate dynamics program.

Lastly, I offer my regards to all people who supported me in any respect during the completion of my study.

# Table of Contents

	Page
List of Tables . . . . .	vii
List of Figures . . . . .	viii
Abstract . . . . .	0
1 Introduction . . . . .	1
2 Improving Detection and Attribution by Filtering Noise and Relaxing Assumptions	16
2.1 Traditional Fingerprinting . . . . .	16
2.2 Fingerprinting Based on Spatial Response Patterns . . . . .	18
2.3 Reducing Noise Using Discriminant Analysis . . . . .	21
2.4 Data . . . . .	22
2.5 Regularizing Optimal Fingerprinting . . . . .	23
2.6 Results . . . . .	25
2.7 Summary . . . . .	30
3 Improving Detection and Attribution by Augmenting Additional Variables . . .	32
3.1 Potential Detectability for Traditional Fingerprinting . . . . .	34
3.2 Potential Detectability for Responses that Are Separable in Space and Time	35
3.3 Potential Detectability Using only Spatial Information . . . . .	37
4 What Surface Observations are Important for Separating the Influences of Anthropogenic Aerosols From Other Forcings? . . . . .	40
4.1 Detection and Attribution Based on Temperature and Precipitation . . . .	40
4.2 Detectability Under Different Combinations of Variables, Spatial Structures, and Seasons . . . . .	49
4.3 Summary and Discussion . . . . .	58
5 Inferring Anthropogenic Aerosol Cooling From Land Observations . . . . .	61
5.1 Data . . . . .	61
5.2 Potential Detectability Over Land . . . . .	63
5.3 Signal-to-Noise Ratio and Multicollinearity . . . . .	66
5.4 Spatial Patterns . . . . .	68
5.5 Confidence Intervals for the Polynomial Coefficients . . . . .	70

5.5.1	Joint Temperature-Precipitation Fingerprints . . . . .	70
5.5.2	Temperature Fingerprints . . . . .	72
5.6	Time Series . . . . .	73
5.7	Forced Responses During 1948-2013 . . . . .	76
5.8	Summary and Discussion . . . . .	80
6	Potential Detectability of Other Variables and Domains . . . . .	82
7	Summary . . . . .	87
A	Generalization of Optimal Fingerprinting . . . . .	90
	Bibliography . . . . .	95

## List of Tables

Table		Page
6.1	The variables, domains, and time averaging windows used to define forced response vectors, and the associated shorthand nomenclature for each. . . .	83

## List of Figures

Figure	Page
<p>1.1 Overview of interactions between aerosols and solar radiation and their impact on climate. The left panels show the instantaneous radiative effects of aerosols (direct effects), while the right panels show their overall impact after the climate system has responded to their radiative effects (semi-direct effects). From FAQ 7.2, Figure 1 of IPCC AR5 WG1. . . . .</p>	3
<p>1.2 Overview of aerosol-cloud interactions and their impact on climate (indirect effect). Panels (a) and (b) represent a clean and a polluted low-level cloud, respectively. From FAQ 7.2, Figure 2 of IPCC AR5 WG1. . . . .</p>	4
<p>1.3 Estimated probability distribution functions for the radiative forcing of total anthropogenic (block filled black curve), long-lived greenhouse gases and ozone (block filled red curve), and aerosol direct and cloud albedo (blue dashed). From fig. 8.16 of IPCC AR5 WG1. . . . .</p>	5
<p>1.4 Climate response patterns estimated from the leading EOF patterns of multi-model mean CMIP5 simulations for: SST in aerosol runs (a) and GHG runs (b); rainfall in aerosol runs (c) and GHG runs (d). The variances explained by these patterns are indicated in the upper right of each panel. From fig. 2 of Xie et al. (2013). . . . .</p>	8
<p>1.5 Scaling factors (top) and derived trends (bottom) over the twentieth century for the GHG (red), anthropogenic aerosol (green), and natural forcings (blue) contributions for HadCM3, PCM, and GFDL models. Solid lines denote analyses including both global mean and subglobal information, dash lines denote analyses including just global mean data, and dotteddashed lines denote analyses including indices. The solid black line in the bottom panel shows observed temperature trend over the century. Diamonds show temperature changes from models unscaled by comparison with observations. From fig. 3 of Stott et al. (2006). . . . .</p>	10

1.6	(a) Estimated contributions of greenhouse gas (GHG, green), anthropogenic aerosols (yellow) and natural (blue) forcing components to observed global mean surface temperature changes over the 1951-2010 period. (b) Corresponding scaling factors by which simulated responses to GHG (green), other anthropogenic (yellow) and natural forcings (blue) must be multiplied to obtain the best fit to Hadley Centre/Climatic Research Unit gridded surface temperature data set 4 (HadCRUT4; Morice et al., 2012) observations based on multiple regressions using response patterns from nine climate models individually and multi-model averages (multi). Results are shown based on an analysis over the 1901-2010 period (squares, Ribes et al. (2013)), an analysis over the 1861-2010 period (triangles, Gillett et al. (2013)) and an analysis over the 1951-2010 period (diamonds, Jones et al. (2013)). Coloured bars show best estimates of the attributable trends (a) and 5 to 95% confidence ranges of scaling factors (c). Vertical dashed lines show the best estimate HadCRUT4 observed trend over the period concerned. From fig. 10.4 of Bindoff and et al. (2013). . . . .	12
2.1	Temporal evolution of Legendre polynomials 1-5 for periods of 1950-2004. .	20
2.2	Eigenvectors 2-5 of the Laplace operator over the HadCRUT4 observational grid. . . . .	24
2.3	95% confidence intervals of global-mean temperature changes associated with the linear trend due to AA (blue) and noAA (red) forcing in CSIRO (upper) an IPSL (lower), as estimated from HadCRUT4 (thick error bars with crosses), historical simulations (thick error bars with dots) and from single-forcing simulations (thin error bars correspond to different ensemble members), for three different calculation methods: “traditional” means that the forced response vectors are derived from the ensemble means of single forcing simulations; “S/N optimized” means the forced response vectors are derived from the first discriminants of the forced response; and “space-only” means the forced response vectors include only the forced response patterns of single forcing simulations. Signs of changes due to AA have being flipped to be positive. . . . .	26

2.4	Results of maximizing detectability of anthropogenic aerosols (left) and all forcings but anthropogenic aerosols (right) in the CSIRO model for annual mean temperature. The top row shows maximized total-to-noise variance ratios (blue line with dots), and the 95% confidence interval under the hypothesis of equal variances (red shading), as a function of the dimensions of laplacian eigenfunctions. The second row shows time series of the most detectable component in a 55-year segment of the control runs (black) and in the twentieth century (blue and red lines), when the component is represented by 20 laplacian eigenfunctions. The blue curve shows the ensemble member used in the maximization while the red curve shows the component in an independent ensemble member. The third row shows the temperature patterns, respectively, associated with the most detectable anomaly using 20 laplacian eigenfunctions. The temperature pattern has units Kelvin per unit time series. . . . .	28
2.5	Similar to fig. 2.4, but for IPSL . . . . .	29
3.1	Schematic illustration of total, forced and internal variability. Internal variability is simulated by a control simulation (black curve). Total variability is simulated by a historical simulation that contains both internal and forced variability (red curves). Multiple red curves indicate independent ensemble members of the historical simulation. . . . .	33
4.1	Spatial patterns of Spherical Harmonics 2, 3, 4, 5 . . . . .	41
4.2	Temporal evolution of Legendre polynomials 1-5 for periods of 1900-2004. .	42
4.3	Time series of global mean temperature change calculated from historical simulations (blue and red), and from single-forcing AA runs (green) and noAA runs (black), using the joint annual mean temperature-precipitation response vector. Different panels correspond to different models and different number of Spheric harmonics, as schematically illustrated in the title. Time series from historical simulations are based on using five Legendre polynomials, whereas time series based on individual forcing runs are based on year-by-year estimation. The response vector is derived from only one ensemble member of the individual forced runs and then used to estimate global mean temperature change time series for each ensemble member from the historical run. . . . .	44

4.4	Time series of global mean temperature change obtained from single-forcing AA runs (green) and noAA runs (red), using the joint annual mean temperature-precipitation response vector derived from the same model (solid) or another model (dashed). Different panels correspond to different models and different number of Spherical harmonics, as schematically illustrated in the title. . . . .	45
4.5	Response vectors that maximize detectability of AA forcing and noAA forcing in two different models (CSIRO and IPSL) based on 20 Laplacian eigenvectors and annual mean fields. The response vector depends on a joint temperature-precipitation relation and therefore is characterized by a pair of patterns, shown in each row. The left column shows temperature fields and the right column shows precipitation fields. The corresponding forcing and model is indicated in the title of each figure. These response patterns were used to estimate the time series shown in fig. 4.4. . . . . .	47
4.6	Maximized signal-to-noise variance ratios (blue line with dots), and the 95% confidence interval under the hypothesis of equal variances (red shading), as a function of the dimensions of fingerprints, based on 20 Laplacian eigenvectors and annual mean fields of each variable, for AA single-forcing runs of CSIRO (top left) and IPSL (bottom left), and noAA single-forcing runs of CSIRO (top right) and IPSL (bottom right). . . . . .	48
4.7	2.5% and 97.5% confidence limits for the coefficients of the five Legendre polynomials (in time) estimated from CSIRO (left column) and IPSL (right column), using 1 Laplacian eigenvector (top row) and 20 Laplacian eigenvectors (bottom row). The dashed line indicates zero. Only the perfect model case, in which all parameters and data come from the same model, is shown.	49

4.8	Potential detectability in a “perfect model” cases, in which the response patterns and noise covariance matrix are from a same model for detecting noAA in CSIRO (a), noAA in IPSL (b), AA in CSIRO (c), and noAA in IPSL (d). Black symbols are for single variables (TAS, PR, PSL). Red symbols are for variable pairs (TAS-PR, PSL-PR, TAS-PSL). Blue symbols are for single variables using variations between seasons and between land and sea. The horizontal axis shows the dimension of the state vector. The precise state vectors are described in the text. The number of Laplacians of each variable used to construct different variable combinations are indicated by marks at top x-axis: black for single variables, red for pairs of variables, blue for variables incorporate information of both seasonal variable and land-sea contrast. Dashed line indicates the threshold for detectability at the 5% significance level. . . . .	51
4.9	Multicollinearity measure (top row), the total-to-noise ratio for noAA (middle row), and the total-to-noise ratio of AA (bottom), for CSIRO (left column) and IPSL (right column), for the variable combinations indicated by colored symbols corresponding to fig. 4.8. . . . .	53
4.10	Same as fig. 4.8, but for “imperfect model” cases, in which response patterns and the noise covariance matrix are from different models, as follows: detecting noAA with response pattern from CSIRO and noise covariance matrix from IPSL (a), detecting noAA with response pattern from IPSL and noise covariance matrix from CSIRO (b), detecting AA with response pattern from CSIRO and noise covariance matrix from IPSL (c), and detecting AA with response pattern from IPSL and noise covariance matrix from CSIRO (d). .	55
4.11	Scatter plot of global average temperature and sea-level pressure over the period 1900-2004. . . . .	56

4.12	Comparison of potential detectability and realized detectability, for noAA forcing (top row) and AA forcing (bottom row), and for CSIRO (left column) and IPSL (right column). potential detectability is computed using <i>TNR</i> ratio estimated from single-forcing run, and realized detectability is computed using <i>TNR</i> ratio estimated from historical run. The colored symbols correspond to the same cases as in fig. 4.8. Grey lines have slope of 1 and cross the origin. Black horizontal lines indicate the threshold of 5% significance level of realized detectability. . . . .	57
4.13	Scatter plot of annual-mean, global-mean temperature and precipitation in simulations from CSIRO (top) and IPSL (bottom), and observations (green symbols). Each panel displays simulations with anthropogenic aerosols only (“AA, blue); with all forcings except anthropogenic aerosols (“noAA, red); with both natural and anthropogenic forcings “his, gold); with no natural and anthropogenic forcing (“ctr, grey). Also shown are observations based on GPCP (green filled circles) and CMAP (green crosses). Temperature in the historical run and observations are centered to have zero mean over the common period 1979-2004, while other simulations are shown relative to the mean of the historical run over this period. Precipitation in the historical run and observations are normalized by their respective means over the period 1979-2004, while other simulations are normalized by the mean of the historical run over the same period. The solid curves show the projection of a low-pass time series of the respective runs. The two lines in the bottom right corner indicate the hydrological sensitivities estimated from the AA and noAA simulations. From fig.3 of DelSole et al. (2015). . . . .	60
5.1	Eigenvectors 2, 3, 4, 5 of the Laplace operator over global land areas. . . . .	62
5.2	Temporal evolution of Legendre polynomials 1-5 for periods of 1901-2013 (left and bottom axes) and 1948-2013 (right and upper axes). . . . .	63

5.3	Potential detectability of different response vectors for noAA in CSIRO (a), noAA in IPSL (b), AA in CSIRO (c), and noAA in IPSL (d). Black symbols are for single variables (TAS, PR, PSL); red symbols are for variable pairs (TAS-PR, PSL-PR, TAS-PSL); blue symbols are for single variables using seasonal variations. The precise state vectors are described in the text. The dimension of each state vector is indicated in the bottom x-axis and the number of Laplacian eigenvectors for each vector is indicated in the top x-axis according to the above color scheme. Dashed line indicates the threshold for detectability at the 5% significance level. . . . .	65
5.4	The signal-to-noise ratio of the noAA (red) and AA (blue) response, and the pattern correlation between the response vectors, in CSIRO (left figures) and IPSL (right figures), as a function of the dimension of Laplacians of the joint annual-mean temperature-precipitation vector. . . . .	67
5.5	Maximized signal-to-noise variance ratios (blue line with dots), and the 95% confidence interval under the hypothesis of equal variances (red shading), as a function of the dimensions of fingerprints, based on 25 Laplacian eigenvectors and annual mean fields of each variable, for AA single-forcing runs of CSIRO (top left) and IPSL (bottom left), and noAA single-forcing runs of CSIRO (top right) and IPSL (bottom right). . . . .	68
5.6	Response vectors that maximize detectability of AA forcing and noAA forcing in two different models (CSIRO and IPSL) based on 25 Laplacian eigenvectors and annual mean fields. The response vector depends on a joint temperature-precipitation relation and therefore is characterized by a pair of patterns, shown in each row. The left column shows temperature fields and the right column shows precipitation fields. The corresponding forcing and model is indicated in the title of each figure. These response patterns were used to estimate the time series shown in fig. 5.10. . . . .	69
5.7	Confidence intervals of estimated coefficients of Legendre polynomials 1 to 5 based on the first (top panels) and the first twenty-five (bottom panels) Laplacian eigenfunctions, separately for noAA (warm colors) and AA (cold colors) of single-forcing runs, historical simulations and CRU observations. Estimates are made based on the joint fields of surface air temperature and precipitation. . . . .	71

5.8	Similar to fig. 5.7 but estimates are made based on only surface air temperature.	74
5.9	Scatter plots of the first (H1) and second (H2) Laplacian time series for surface air temperature over global land, for CSIRO (top) and IPSL (bottom), for historical simulations and observations (left), and AA, noAA and pre-industrial control simulations (right). Different simulations are shown with different colors, as indicated in the legend, and the colored lines show the least squares fit line. The title of each panel shows the correlation of the scatter plots for the respective simulations and observations. The correlation threshold for a 5% significance level is 0.2. . . . .	75
5.10	Response time series derived from historical simulations (blue and red), derived from single-forcing AA runs (green) and noAA runs (black) and from CRU observations (thick red and thick blue) using the temperature response vector (top rows) and the joint temperature-precipitation response vector (bottom rows) associated with the first 25 Laplacian eigenfunctions and Legendre polynomials 1-5. Different columns correspond to different models used for the historical simulation and response vector, as illustrated in the title. The response vector is derived from only one ensemble member of the single-forcing runs and then used to estimate response time series for each ensemble member from the historical run and observations. Time series from historical simulations are based on using five Legendre polynomials, whereas time series based on single-forcing runs are based on year-by-year estimation. . . . .	77
5.11	Similar to fig. 5.7 except that the response vectors of AA and noAA are derived from single-forcing runs from 1948 to 2004, and observations are CRU and GHCN-CAMS data from 1948 to 2013. . . . .	78
5.12	Similar to fig. 5.11 except that the response vectors of AA and noAA are derived from annual mean TAS only. . . . .	79
6.1	Potential detectability of anthropogenic aerosol forcing of different variables, domains, and time averaging using eight-dimensional response vectors from CSIRO. The response vector is defined along the y-axis using the nomenclature in table 6. For instance, ‘tos_NPSST.rsds_ASIALAND.ann’ shows the potential detectability of joint sea surface temperature over North Pacific and surface downward shortwave over Asia land area. Previously studied variables are marked in red, variables to be analyzed in more depth in the future are marked in grey. . . . .	84

6.2 Scatter plots of potential detectability to AA and to noAA, in simulations from CSIRO (top left) and IPSL (top right). And scatter plot of potential detectability to AA in simulations from CSIRO and IPSL (bottom left), and scatter plot of potential detectability to noAA in simulations of CSIRO and IPSL (bottom right). All potential detectability are calculated for variables (combinations) shown in fig. 6.1 and the correspondence is shown by the rank numbers. Line of  $x = y$  is displayed in each panel by the red line. . . 86

## Abstract

### A SYSTEMATIC FRAMEWORK FOR IMPROVING ESTIMATES OF ANTHROPOGENIC AEROSOL COOLING

Xiaoqin Yan, PhD

George Mason University, 2016

Dissertation Director: Dr. Timothy DelSole

One of the most significant uncertainties in climate change projections is the sensitivity of the climate system to increasing greenhouse gas concentrations. Attempts to estimate this sensitivity based on observations over the past century have failed to reduced this uncertainty, primarily because of uncertainties in the cooling effect of aerosols, which have cancelled some of the warming induced by greenhouse gases. This study attempts to improve estimates of aerosol cooling by exploiting new statistical techniques and by identifying variable combinations that are more effective at estimating the response to climate forcings than single variables. The exploration of variable combinations is facilitated using a new measure called potential detectability, which quantifies the extent to which the response to climate forcing can be detected in a model. It is shown that joint temperature-precipitation information over a global domain provides the most accurate estimate of aerosol forced responses in climate models, compared to using temperature, precipitation, or sea level pressure individually or in combination. Unfortunately, observational errors in precipitation are too large to permit estimation of aerosol-induced climate changes. Repeating this estimation using only land-data, where reliable rain gauge data are available, succeeds in estimating aerosol cooling, but is only modestly improved by including precipitation data. We show that the time evolution of the response to aerosol forcing can be estimated from observations. This conclusion goes beyond previous studies by using only the spatial structure of the forced response to detect aerosol forcing.

## Chapter 1: Introduction

There is no serious scientific doubt that the earth is warming and that most of this warming has been caused by human activities, mostly by increasing concentrations of long-lived greenhouse gases (Masson-Delmotte et al. 2013; Myhre et al. 2013; Bindoff and et al. 2013). This warming has been accompanied by diminishing snow and ice (Vaughan et al. 2013), rising sea levels (Church et al. 2013), and changes in extreme weather and climate events (IPCC 2012). Continued emission of greenhouse gases will undoubtedly cause further warming, melting snow and ice, and sea level rise, increasing the likelihood of serious, widespread, and irreversible impacts for people and ecosystems.

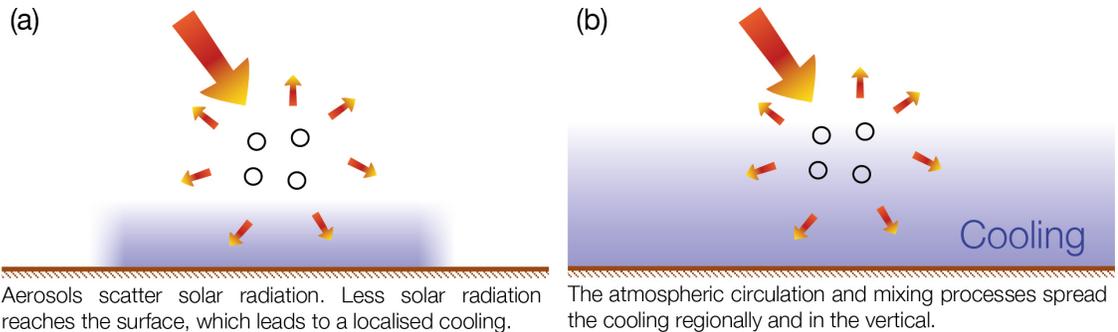
An urgent concern is how climate will change in the future in response to human activity. State of the art climate models predict that global mean temperature will warm by 1.4 to 3.1°C by the end of the 21st century (5 to 95% confidence range, Collins et al. 2013, table 12.2). Unfortunately, this uncertainty range is relatively large: the high end of this prediction, which is in addition to the 0.8°C warming that has already occurred since pre-industrial times, is so far beyond the range of what has been experienced over the past millions of years that abrupt, unpredictable, and potentially irreversible changes cannot be ruled out (e.g., collapse of ecosystems, destabilization of Greenland and Antarctica ice-sheets; Molina et al. 2014; IPCC 2014). The largest source of uncertainty in climate projections is *climate sensitivity* (Knutti et al. 2002; Knutti and Hegerl 2008), which quantifies how a climate system responds to an energy imbalance. One measure of climate sensitivity is the Equilibrium Climate Sensitivity (ECS), defined as the global mean surface warming in response to a doubling of CO<sub>2</sub> after the system has reached a new steady state. The ECS of comprehensive climate models depends on parameters affecting clouds, convection, and other processes in ways that are not well constrained by observations (Flato et al. 2013).

Attempts to constrain ECS using recent climate observations have ruled out low sensitivities ( $ECS < 1^\circ\text{C}$ ), but the upper limit of sensitivity is much less constrained (Forest et al. 2002; Gregory et al. 2002; Knutti and Hegerl 2008). The main reason for this sensitivity is the uncertainty in the radiative forcing due to aerosols. Aerosols are tiny particles suspending in atmosphere, which can be either natural or anthropogenic (human-made). Typical natural aerosols are sea salt, volcanic aerosols and most mineral dust, while typical anthropogenic aerosols (AA) are black carbon, sulphate, nitrate and ammonium that are formed mostly from the incomplete combustion of fossil and biomass based fuels. Aerosols influence the climate mainly through three mechanisms: i) absorption and scattering of incoming solar radiation, called the *direct effect* (left panels of fig. 1.1); ii) local heating due to absorption of solar radiation and the associated change in atmospheric temperature, called the *semi-direct effect* (right panels of fig. 1.1); iii) and changing the microphysical and dynamical behavior of clouds, called the *indirect effect* (fig. 1.2). For the direct effect, atmospheric aerosols absorb or scatter solar radiation, depending on the species of aerosols, and in either case reduce the amount of solar radiation reaching the Earth's surface. For the semi-direct effect, black carbon (BC) can absorb solar radiation and induce warming in the surrounding air, which can prevent cloud formation due to enhanced stability or even evaporate cloud droplets, thereby allowing more solar radiation to reach the surface (Koren et al. 2004). Finally, aerosol particles serve as cloud condensation and ice nuclei and, as such, compete for uptake of water vapor, resulting in smaller cloud droplets and greater overall surface area, reflecting more solar radiation back to space (Twomey 1977; Lohmann 2006). In addition, smaller cloud droplets are less likely to collide to form rain droplets and thereby prolong the life time of the cloud, further reinforcing the cloud albedo effect (Albrecht 1989). Numerous other indirect effects are known to exist but their impact is difficult to quantify and parameterize in global scale climate models.

To quantify the impact of changes in climate properties, scientists use the concept of radiative forcing, which is defined as the net change of the energy balance of the climate system between two particular times, such as pre-industrial to present-day. According to

## Aerosol-radiation interactions

### Scattering aerosols



### Absorbing aerosols

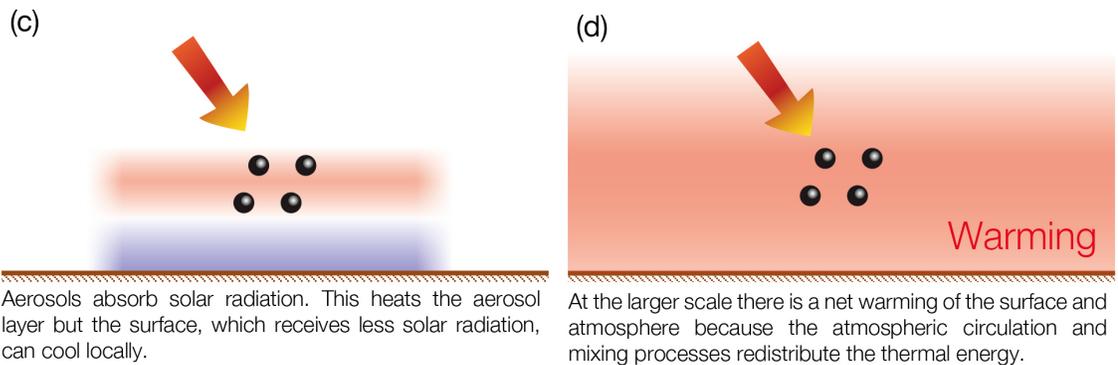
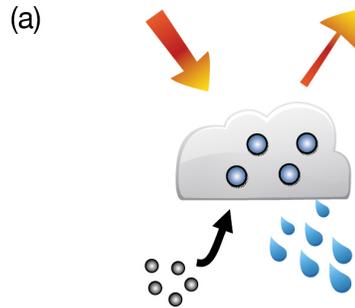


Figure 1.1: Overview of interactions between aerosols and solar radiation and their impact on climate. The left panels show the instantaneous radiative effects of aerosols (direct effects), while the right panels show their overall impact after the climate system has responded to their radiative effects (semi-direct effects). From FAQ 7.2, Figure 1 of IPCC AR5 WG1.

Boucher et al. (2013), the current best estimate of direct effect forcing of aerosols is  $-0.45 \text{ W m}^{-2}$ , with a confidence interval of  $(-0.95, 0.05)$ ; the current best estimate of the semi-direct effect is  $+0.4 (0.05 \text{ to } 0.8) \text{ W m}^{-2}$ ; and, the current best estimates of the total radiative forcing of aerosols is  $-0.9 (-1.9 \text{ to } -0.1) \text{ W m}^{-2}$ . The fact that the net radiative forcing is negative indicates that aerosols tend to *cool* the earth. The large uncertainties associated with these estimates are due primarily to uncertainties in the emission levels, size distribution of aerosols, and certain optical properties of sulfates.

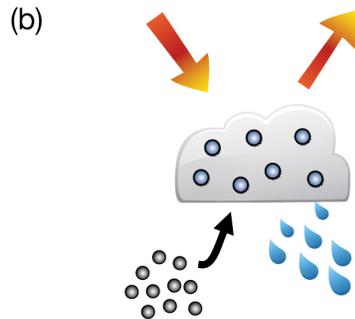
An assessment of the global radiative forcing is summarized in fig. 1.3. Current estimates

## Aerosol-cloud interactions



---

Aerosols serve as cloud condensation nuclei upon which liquid droplets can form.



---

More aerosols result in a larger concentration of smaller droplets, leading to a brighter cloud. However there are many other possible aerosol–cloud–precipitation processes which may amplify or dampen this effect.

Figure 1.2: Overview of aerosol-cloud interactions and their impact on climate (indirect effect). Panels (a) and (b) represent a clean and a polluted low-level cloud, respectively. From FAQ 7.2, Figure 2 of IPCC AR5 WG1.

of global radiative forcing suggest that the forcing due to AA is more uncertain than that due to well-mixed greenhouse gases (GHG), and hence dominates most of the uncertainty in total forcing (Forster et al. 2007, see also fig. 2.20). Also, aerosol cooling partially cancels greenhouse warming. Thus, it is possible that the climate system has a large climate sensitivity, but this sensitivity was masked in the past because of a large cooling from atmospheric aerosols. Also, the large uncertainty in radiative forcing, and the fact that these two forcings are of opposite sign, raises the possibility that climate models might

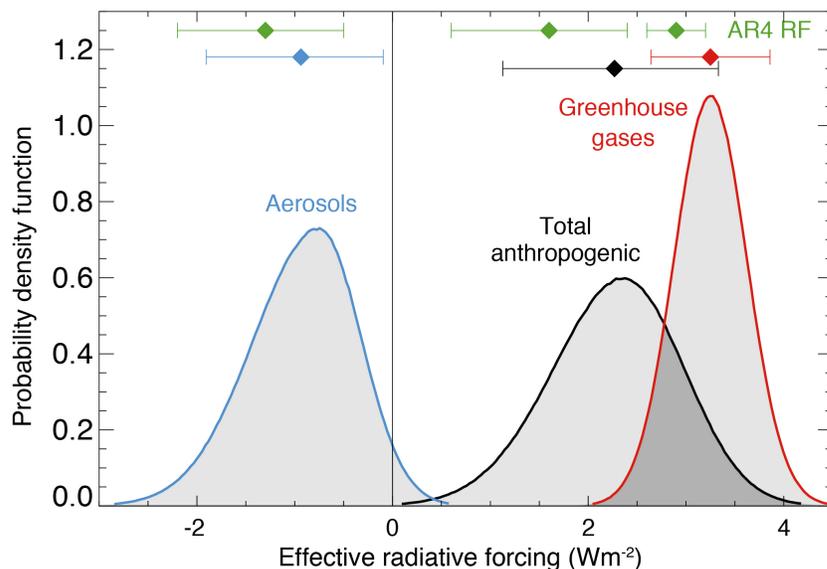


Figure 1.3: Estimated probability distribution functions for the radiative forcing of total anthropogenic (block filled black curve), long-lived greenhouse gases and ozone (block filled red curve), and aerosol direct and cloud albedo (blue dashed). From fig. 8.16 of IPCC AR5 WG1.

respond incorrectly to a given radiative forcing, but still produce realistic simulations of past warming by plausible adjustments in the cooling from aerosols. Therefore it is important to identify the anthropogenic aerosol forced responses in models and observations, in order to validate models and hence to enhance the reliability of future projections.

Aside from its important in estimating climate sensitivity, anthropogenic aerosols also have been implicated as an important source of climate variability (see Boucher et al. 2013, for a comprehensive review). Booth et al. (2012) found that the inclusion of aerosol indirect effects allow model simulation to capture the magnitude, the temporal and spatial structure of SST variability over North Atlantic. Shindell et al. (2012) found that aerosols contribute significantly to precipitation changes over the Sahel and South and East Asia in historical simulations, and inclusion of aerosols in model simulations improves the agreement between simulated and observed precipitation trends. Dong and Zhou (2014) show that GHGs and AAs effectively compete to form the Indian Ocean basin-wide warming pattern

and the positive India Ocean Dipole-like warming pattern along the equator. Model studies of Rotstayn and Lohmann (2002) show that anthropogenic aerosol cooling of the Northern Hemisphere can alter the the low-latitude circulation and rainfall substantially and be the primary cause of the consistent southward shift of ITCZ. Ridley et al. (2015) use precipitation reconstructed from carbon isotope records to show that the anthropogenic aerosol emissions have led to a reduction of rainfall in the northern tropics during the twentieth century. Allen et al. (2014) found from both models and observations that anthropogenic aerosols can modify the Pacific Decadal Oscillation and therefore contribute to the expansion and contraction of the tropical belt during the twentieth century. Multi-model analyses of Chang et al. (2010) show that the anthropogenic sulfate aerosols may have dramatically altered the tropical Atlantic rainfall during the twentieth century. Thus, methods for estimating the response to anthropogenic aerosols would clarify the role of anthropogenic aerosols in climate variability.

In order to estimate the role of anthropogenic aerosols in past climate variability, some assumptions about climate variability need to be made. A basic paradigm in climate science is that climate variability can be partitioned as

$$\begin{array}{rcccl}
 o & = & (f_1 + f_2 + \dots + f_M) & + & n \\
 \text{climate} & & \text{forced} & & \text{internal} \\
 \text{variability} & & \text{variability} & & \text{variability}
 \end{array} \tag{1.1}$$

where  $f_1, f_2, \dots, f_M$  denote forced variability that occurs in response to external forcings (e.g., GHG, anthropogenic aerosols, solar forcing, volcanoes, land-use changes, etc.), and internal variability  $n$  refers to the chaotic variability that arises from internal dynamics of the Earth system in the absence of external forcings. The statistical characteristics of internal variability can be estimated by running climate models with no year-to-year variation in external forcing. These simulations are often called *control runs*. The response to specific forcings can be estimated by running climate models driven by that specific forcing but

starting at slightly different initial conditions. The resulting simulations are called *ensemble members*. For brevity, we call a simulation driven by a specific forcing a *single-forcing run*. Because the climate system is chaotic, different simulations will have different realizations of internal variability. However, if the above paradigm is correct, the different simulations should have the same forced variability, which therefore can be estimated by averaging over the ensemble members while holding time and spatial location fixed. Similarly, climate models can be run with all forcings. These simulations are called *historical* simulations. It should be recognized that estimates of forced and internal variability derived from models may not represent the truth (indeed, they sometimes differ greatly between models).

A first step, then, is to identify the characteristics of the forced responses  $f_m$ . We are especially interested in how the response to anthropogenic aerosols *differ* from other forcings, especially GHG warming, since differences between responses provide the basis for distinguishing them in observations. A major difference in forcing is that greenhouse gases tend to be long-lived and thus well mixed, with relatively weak spatial gradients, whereas anthropogenic aerosols are short lived and thus tend to be more concentrated around their (land) sources. These differences in forcing lead to differences in hemispheric gradients and land-sea gradients in the temperature and precipitation responses. Forcings also differ temporally: forcing from well-mixed greenhouse gases has increased monotonically over the past century while forcing from aerosols has leveled in the last few decades. In addition, both forcings have distinct seasonal cycles. Despite these forcing differences, the associated response patterns tend to be similar to each other. As a representative example, we show in fig. 1.4 estimates of forced variability patterns from Xie et al. (2013), which were derived from a principal component analysis of single-forcing runs averaged across ensemble members and models. The top and bottom rows show the SST and precipitation response, respectively, while the left and right columns show the response to aerosol and GHG forcing, respectively. The pattern correlation amounts to 0.87 for sea surface temperature and 0.67 for precipitation. The similarity of these patterns suggest that separating these responses using data might encounter multicollinearity problems.

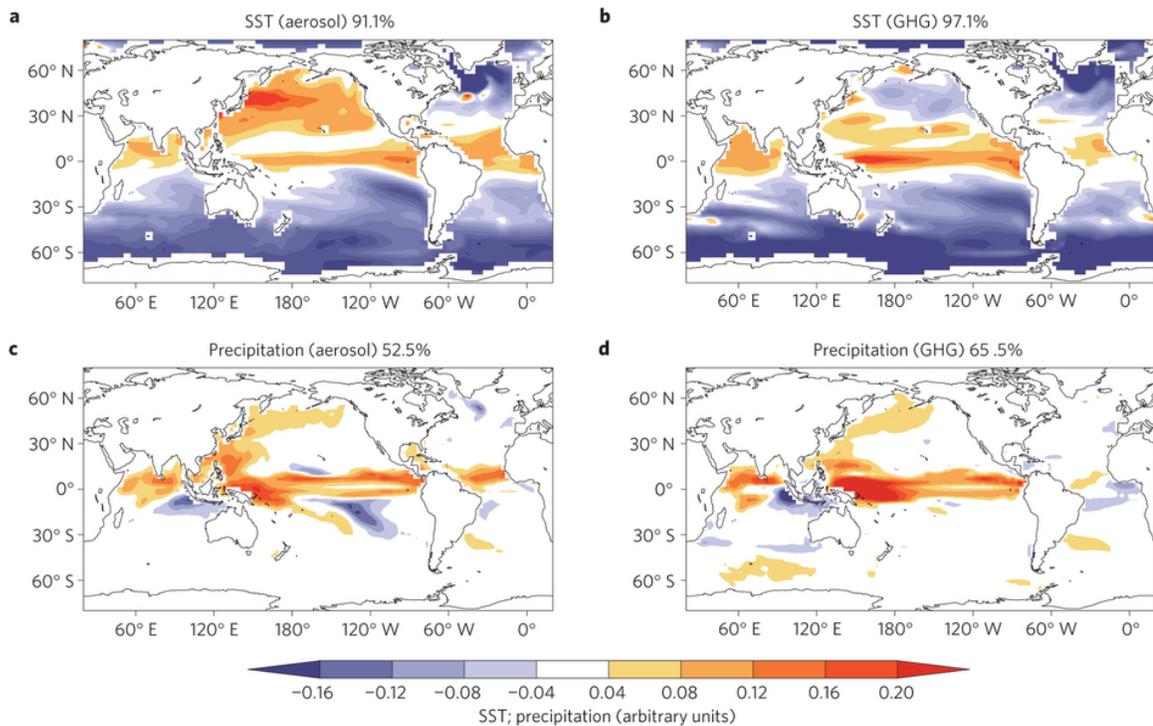


Figure 1.4: Climate response patterns estimated from the leading EOF patterns of multi-model mean CMIP5 simulations for: SST in aerosol runs (a) and GHG runs (b); rainfall in aerosol runs (c) and GHG runs (d). The variances explained by these patterns are indicated in the upper right of each panel. From fig. 2 of Xie et al. (2013).

The most widely used method for estimating forced responses in observations is *optimal fingerprinting* (Allen and Tett 1999; Hegerl et al. 2007; Hegerl and Zwiers 2011), which is essentially generalized least squares (GLS). In this approach, a scaling factor is attached to each forcing in (1.1) to yield the model

$$o = (a_1 f_1 + a_2 f_2 + \dots + a_M f_M) + n. \quad (1.2)$$

Then, the scaling coefficients  $a_1, \dots, a_M$  are estimated using GLS. If the hypothesis

$$H_D : a_m = 0 \quad (1.3)$$

can be rejected, then the response to the  $m$ 'th forcing is said to be *detected*. If the hypothesis

$$H_A : a_m = 1, \tag{1.4}$$

cannot be rejected, then variability is said to have been *attributed* to the  $m$ 'th forced response. If  $H_A$  is rejected and the estimated value of  $a_m$  exceeds one, then the model is said to underestimate the response, otherwise it is said to overestimate the response.

To illustrate, we discuss an analysis by Stott et al. (2006). Stott et al. (2006) assumed that observed surface air temperature could be modeled as (1.2) with three forced response vectors: anthropogenic aerosol forcing, GHG forcing, and natural forcing (e.g., volcanoes and solar insolation). The response vectors and internal variability statistics were estimated from three separate climate models (HadCM3, PCM, GFDL) and used to fit observed surface air temperature over the twentieth century (1900-1999). The resulting scaling coefficients  $a_1, a_2, a_3$  and associated confidence intervals are shown in the top panel of fig. 1.5. Three sets of scaling coefficients are shown because they used three sets of response vectors: response vectors using both global mean and subglobal information (solid); response vectors using global mean alone (dashed); and response vectors using indices of global mean, hemispheric gradient, land-sea contrast and meridional gradient over northern hemisphere (dotted-dashed). Note that confidence intervals for temperature indices tend to be comparable to those using subglobal information, indicating that the information needed to detect forced responses are included mainly in the indices. Also, confidence intervals for response vectors using only global mean information tend to be larger than those for the other response vectors, with the GFDL model showing undetectable aerosol forcing, suggesting that global mean information is not necessarily adequate for detection.

The temperature trend induced by different forcings are calculated by applying the scaling coefficients to the simulated trend in individual models. The resulting estimated trends are shown in the bottom panel of 1.5. For the 100-y period of 1900-1999, the estimated trends show that warming due to GHG forcing (red solid) is detected and attributed consistently for the three models within a range of  $0.7 - 1.3K$ ; cooling due to anthropogenic

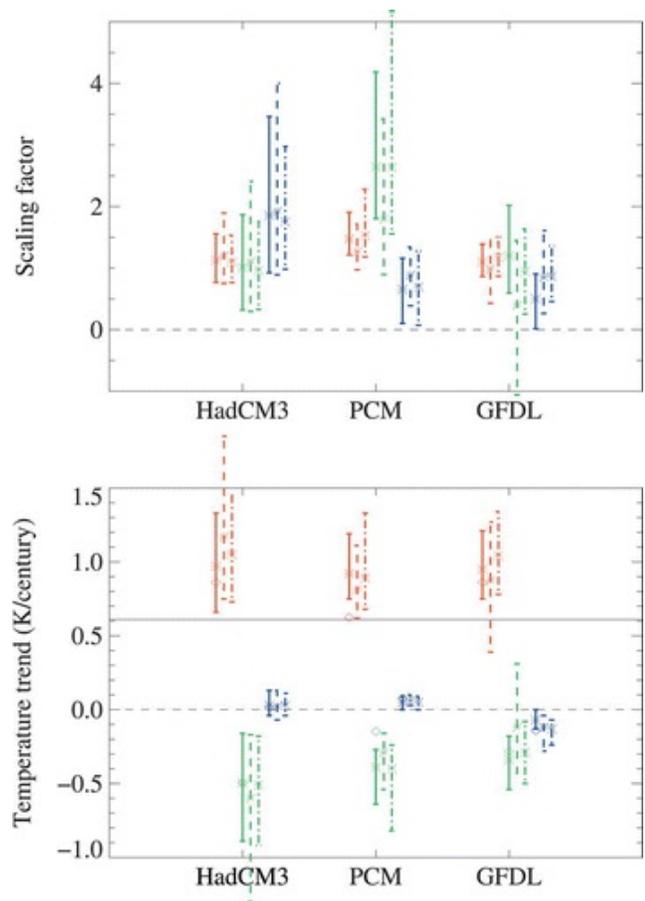


Figure 1.5: Scaling factors (top) and derived trends (bottom) over the twentieth century for the GHG (red), anthropogenic aerosol (green), and natural forcings (blue) contributions for HadCM3, PCM, and GFDL models. Solid lines denote analyses including both global mean and subglobal information, dash lines denote analyses including just global mean data, and dotted-dashed lines denote analyses including indices. The solid black line in the bottom panel shows observed temperature trend over the century. Diamonds show temperature changes from models unscaled by comparison with observations. From fig. 3 of Stott et al. (2006).

aerosol forcing (green solid) is detected, but with larger uncertainty, within a range of  $0.33 - 0.49K$ . Combining results of the three models, estimates of the 5 – 95 percentile range is  $0.4 - 1.4W/m^2$  for net aerosol forcing and is  $2.2 - 4.0K$  for transient climate response (the globally averaged temperature difference between the average temperature over the 20-yr period around the time of  $CO_2$  doubling and the mean temperature of the control

run with CO<sub>2</sub> at preindustrial levels).

Similarly, other studies estimated the scaling coefficients associated with responses to different forcings by applying the above optimal fingerprinting method to data with different length. According to Bindoff and et al. (2013), over the period of 1951-2010, the changes of observed global mean surface temperature is approximately 0.6°C, to which anthropogenic aerosols contributed -0.6°C to 0.1°C (5 to 95% uncertainty). Note that even the sign of aerosol forced responses is not well determined yet, as can be seen more clearly from fig. 10.4 of Bindoff and et al. (2013).

The question arises as to whether the above results can be improved upon. Despite the word “optimal,” additional improvement is possible because the word “optimal” refers only to the fact that the scaling factors derived from fingerprinting have minimal variance *for the given response vectors and noise covariances*. Nothing in this definition prohibits *other vectors* from estimating aerosol cooling and greenhouse warming with less uncertainty. There are several ways in which alternative response vectors can improve estimates of aerosol cooling and greenhouse gas warming:

- a vector could have a stronger response to a specific forcing.
- a vector could have less internal variability.
- a vector could be less collinear with the vectors associated with other forcings.

The goal of our research is to improve estimates of the response to anthropogenic aerosols by characterizing the forced responses in different (and hopefully better) ways.

One approach to improving estimates of forced responses is to identify response vectors with less noise. Most optimal fingerprinting studies estimate the forced response using the ensemble mean of single-forcing simulations. To improve upon this approach, we select the response vectors that maximize the ratio of variances of total variability over internal variability. The resulting vector maximally discriminates between forced and internal variability. This vector can be obtained using the method of Jia and DelSole (2012) (discussed shortly). Although these vectors maximize detectability of a single forcing, they may not be

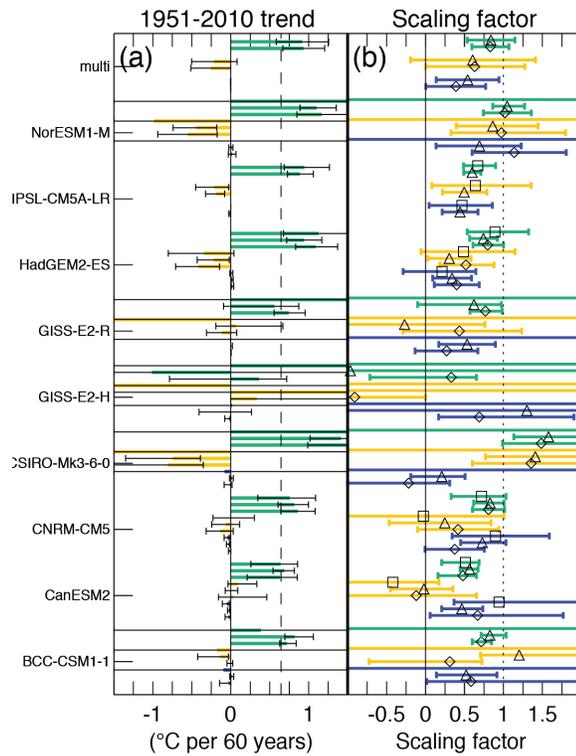


Figure 1.6: (a) Estimated contributions of greenhouse gas (GHG, green), anthropogenic aerosols (yellow) and natural (blue) forcing components to observed global mean surface temperature changes over the 1951-2010 period. (b) Corresponding scaling factors by which simulated responses to GHG (green), other anthropogenic (yellow) and natural forcings (blue) must be multiplied to obtain the best fit to Hadley Centre/Climatic Research Unit gridded surface temperature data set 4 (HadCRUT4; Morice et al., 2012) observations based on multiple regressions using response patterns from nine climate models individually and multi-model averages (multi). Results are shown based on an analysis over the 1901-2010 period (squares, Ribes et al. (2013)), an analysis over the 1861-2010 period (triangles, Gillett et al. (2013)) and an analysis over the 1951-2010 period (diamonds, Jones et al. (2013)). Coloured bars show best estimates of the attributable trends (a) and 5 to 95% confidence ranges of scaling factors (c). Vertical dashed lines show the best estimate HadCRUT4 observed trend over the period concerned. From fig. 10.4 of Bindoff and et al. (2013).

the best variables to *distinguish* different forcings. For instance, two forcings may produce highly detectable, but virtually identical responses in a given set of variables, in which case those variables could not be used to distinguish the forced responses. Our goal is to find the best variables for *separating* forced responses from each other and from internal variability.

Some hints about the features of the response that are important for estimating anthropogenic aerosols can be inferred from the study of Stott et al. (2006). In particular, they examined the sensitivity of their estimates to including or excluding features like the differential warming rates between the hemispheres, between land and ocean, and between mid- and low-latitudes. They found that removing one or more of these features from the response vector increased the final uncertainty. These features will be explored in this thesis.

Another approach to improving estimates of anthropogenic aerosol cooling is to enhance the signal by using new variables or by combining variables. Indeed, previous studies have found that combining different variables can improve inferences, including using joint changes of temperature at the surface and aloft (Jones et al. 2003), joint changes of temperature, diurnal temperature range and precipitation (Schnur and Hasselmann 2005), joint changes of temperature and salinity (Stott and Jones 2009; Pierce et al. 2012), joint changes of maximum amount and location of zonal mean precipitation (Marvel and Bonfils 2013), and joint changes on river flows, winter air temperature and snow pack over the Western United States (Barnett et al. 2008). While these studies illustrate the success that can be achieved using multiple variables, several questions arise. For instance: is there an objective way to demonstrate that combining variables does indeed improve the final estimate? How can the relative merits of different variables be compared?

The purpose of our research is to quantify the effectiveness of different features, different variables, and different methods for detecting the response to aerosol forcing, and thereby improve estimates of anthropogenic aerosol cooling. One could attempt to do this by brute force: simply apply various guesses for response vectors to observations and see what happens. Doing this for different variables, domains, combinations, and time averages would produce an overwhelming amount of data that would be difficult to interpret. We have developed a much more systematic procedure that also is rigorous and insightful. A key quantity in our methodology is *potential detectability*, which measures the detectability of a forced response *in a model*. For a single forcing, potential detectability equals the total-to-noise ratio of the forced response, defined as the ratio of the total variance (forced plus

unforced) to the variance of internal variability. For two forcings, potential detectability equals the product of the total-to-noise ratio of the given forcing times a measure of multicollinearity between the two forcings (as we will show). Both quantities can be estimated from single-forcing simulations. Moreover, potential detectability is directly related to the statistic for testing the detection hypothesis (1.3) and can be compared across vectors and models. Thus, this approach not only allows us to compare the detectability of different response vectors, but can be used to diagnose why certain vectors are not as good as others (e.g., the vectors have weak responses or are close to collinear).

In addition, most optimal fingerprinting studies distinguish different forced responses by exploiting differences in forced response patterns that contain information of both spatial and temporal evolution. Therefore, the relative importance of spatial versus temporal structure for separating different forced responses is unclear. The extent to which different forced responses can be separated based on time series of global mean quantities has been investigated extensively (Stern and Kaufmann 2000; Lean and Rind 2009; Zhou and Tung 2012). In this study, we not only investigate the extent to which forced responses can be separated based on forced response vectors including information of both spatial and temporal structure, but also explore the extent to which forced responses can be separated based on solely spatial structures. Thus, we explore an approach to optimal fingerprinting that differs from previous approaches in that we do not use the full space-time evolution of the climate change signal as a predictor. Instead, we use only the spatial structure of a variable, and its co-variability with other variables, as input to fingerprinting. The output is an estimate of the time histories of the forced responses. This approach avoids making the (strong) assumption that the climate model correctly simulates the spatial-temporal evolution of the forced response. The consistency of optimal fingerprinting methods with different assumptions will also be investigated.

In chapter 2, a brief review of standard optimal fingerprinting is given. Then, we propose that instead of using the ensemble mean to estimate the forced response, we use vectors

that maximize the total-to-noise ratio of the forced response. Another modification of optimal fingerprinting is proposed whereby only the spatial structure of the forced response is included in optimal fingerprinting, leaving the time history to be inferred from data. The proposed methodologies are illustrated using climate simulations and observations of surface air temperature. In chapter 3, potential detectability of forced response vectors are derived for the different fingerprinting methods. In chapter 4, potential detectability is evaluated in climate simulations over a global domain in a perfect model setting. This analysis reveals that the *joint* temperature and precipitation response vector has the most potential detectability of any vector on a global domain. However, evidence is presented that global precipitation observations are not accurate enough to use such vectors to estimate anthropogenic forced responses. In chapter 5, the potential detectability of land-only variables, observations of which are considered to be more reliable, is evaluated. These results are then used to guide analysis of observations. Potential detectability of other variables and domains are summarized in chapter 6. Finally, chapter 7 gives a summary of conclusions.

## Chapter 2: Improving Detection and Attribution by Filtering Noise and Relaxing Assumptions

### 2.1 Traditional Fingerprinting

The goal of our research is to improve estimates of forced responses by exploiting advances in optimal fingerprinting. Accordingly, we first review optimal fingerprinting. The material presented here is standard (Allen and Tett 1999; Hegerl et al. 2007; Hegerl and Zwiers 2011). A basic paradigm in climate science is to partition climate variability into two parts: *forced* variability, which occurs in response to external forcings, such as volcanic eruptions, solar variations, and human-induced changes in atmospheric composition, and *internal* variability, which occurs in the absence of external forcings, and arises from internal dynamics of the atmosphere-ocean-land-cryosphere system. Optimal fingerprinting assumes that observed variability  $\mathbf{o}$  can be modeled as a linear sum of responses to individual forcings  $\mathbf{f}_1, \dots, \mathbf{f}_M$  plus random internal variability  $\mathbf{n}$ ,

$$\mathbf{o} = \sum_{i=1}^M \mathbf{f}_i a_i + \mathbf{n}, \quad (2.1)$$

where  $a_i$  are unknown scaling factors to account for amplitude errors. The internal variability  $\mathbf{n}$  is assumed to be a random vector from a multivariate distribution with covariance matrix  $\tilde{\Sigma}_N$ . To estimate the scaling factors, the response to individual forcings  $\mathbf{f}_1, \dots, \mathbf{f}_M$  and the statistics of the noise need to be pre-specified. The noise covariance matrix is estimated from a long control simulation that has no year-to-year variation in climate forcing. The forced response vector  $\mathbf{f}_m$  is estimated from climate simulations driven by the  $m$ th forcing. By convention, the response vector is normalized such that the model amplitude is  $a_m = 1$ .

The scaling factors are then estimated using a generalized version of regression that accounts for covariance structure in the noise. The solution is most conveniently expressed by writing model (2.1) equivalently as

$$\mathbf{o} = \mathbf{F}\mathbf{a} + \mathbf{n} \quad (2.2)$$

where,

$$\mathbf{a} = (a_1 \ a_2 \ \dots \ a_m)^T \quad \text{and} \quad \mathbf{F} = (\mathbf{f}_1 \ \mathbf{f}_2 \ \dots \ \mathbf{f}_m)^T. \quad (2.3)$$

Then the generalized least squares estimate of the amplitudes (Allen and Tett 1999) is

$$\hat{\mathbf{a}} = (\mathbf{F}^T \tilde{\Sigma}_N^{-1} \mathbf{F})^{-1} \mathbf{F}^T \tilde{\Sigma}_N^{-1} \mathbf{o} \quad (2.4)$$

where the caret  $\hat{\ }^$  denotes a sample estimate. The estimate is unbiased, in the sense that  $E[\hat{\mathbf{a}}] = \mathbf{a}$ , and has covariance matrix

$$\text{cov}[\hat{\mathbf{a}}] = (\mathbf{F}^T \tilde{\Sigma}_N^{-1} \mathbf{F})^{-1}. \quad (2.5)$$

If the noise is normally distributed, then  $\hat{\mathbf{a}}$  also is normally distributed, and a 95% confidence interval for the  $m$ 'th regression coefficient is

$$\left( \hat{a}_m - 1.96 \left( \left( \mathbf{F}^T \tilde{\Sigma}_\epsilon^{-1} \mathbf{F} \right)^{-1} \right)_{mm}^{1/2}, \quad \hat{a}_m + 1.96 \left( \left( \mathbf{F}^T \tilde{\Sigma}_\epsilon^{-1} \mathbf{F} \right)^{-1} \right)_{mm}^{1/2} \right). \quad (2.6)$$

The hypothesis  $a_m = 0$  is rejected when the above confidence interval excludes zero. This step is called *detection*. If the hypothesis  $a_m = 1$  cannot be rejected, then we say the observed change can be *attributed* to the  $m$ 'th forcing. As a result, this procedure is called *detection and attribution analysis*. Again, the above results are standard.

## 2.2 Fingerprinting Based on Spatial Response Patterns

This thesis explores a novel modification of optimal fingerprinting: instead of using the full spatio-temporal evolution of the model forced response, only the spatial structure of the forced response is used, leaving the time evolution of the response to be inferred from data. The resulting detection and attribution analysis is argued to be more convincing than the traditional approach because less information from models is assumed as input to the fingerprinting, and more information about the forced responses (e.g., the time evolution) is derived from observations and then compared to models.

Let the observed change at the  $n$ 'th time step and  $s$ 'th spatial location be denoted  $o_{n,s}$ . Thus  $\mathbf{O}$  is an  $N \times S$  matrix for a total of  $N$  time steps and  $S$  spatial locations. As before, observed changes  $\mathbf{O}$  are modeled as a linear sum of simulated changes from forcings  $\mathbf{f}_1, \dots, \mathbf{f}_M$  plus random internal variability  $\eta$ ,

$$\mathbf{O} = \sum_{m=1}^M \mathbf{f}_m a_m + \eta, \quad (2.7)$$

where  $a_1, \dots, a_M$  are scaling factors to account for amplitude errors. A different symbol is used for internal variability compared to (2.2) because we make different assumptions about it. In particular, we assume internal variability  $\eta$  is white noise such that each time step is drawn independently from a Gaussian distribution with zero mean and  $S$ -dimensional covariance matrix  $\Sigma_\eta$ . One might question whether the (temporal) white noise assumption for internal variability is appropriate. It is found in this study that the white noise assumption gives reasonable estimates of uncertainty, in the sense that the results derived from separate ensemble members frequently lie (at the expected rate) within the confidence interval derived from any one of the members. This consistency suggests that the above model (including the white noise assumption) is reasonable.

The forced response is represented in a form that is separable in space and time, namely

$$\mathbf{f}_m = \mathbf{r}_m \mathbf{p}_m^T, \quad (2.8)$$

where  $\mathbf{r}_m$  is an  $N$ -dimensional vector specifying the response time series,  $\mathbf{p}_m$  is an  $S$ -dimensional vector specifying the spatial pattern, and superscript  $T$  denotes the matrix transpose. The pattern  $\mathbf{p}_m$  is derived from single-forcing simulations in a manner that will be discussed in the next section. The goal of the modified fingerprinting method is to estimate the time series  $\mathbf{r}_m$  (which subsumes the scaling factor  $a_m$ ). This approach requires information of only the spatial structure of the forced response and does not impose any constraints on the forced time variability. Unfortunately, the uncertainties become unacceptably large when time series for two or more forcings are estimated for each year separately (e.g., confidence intervals for the scaling factors include 0, implying that the forced response can not be detected). Consequently, this study assumes that the response evolves slowly in time, which allows the estimation method to pool data across years to reduce uncertainty. Specifically, the temporal variation is represented by a  $J$ 'th order polynomial in time:

$$r_m(t) = \sum_{j=1}^J c_{m,j} L_j(t), \quad (2.9)$$

where  $L_j(t)$  is the  $j$ 'th Legendre polynomial in time  $t$ . This formulation requires estimating  $J$  coefficients per forcing. An illustration of Legendre polynomials 1-5 over a time period from 1950 to 2004 are shown in Figure 2.1. The polynomial representation is most questionable during volcanic eruptions, but since the dominant response to volcanic eruptions spans only a few years (Iles et al. 2013), smoothing over periods of major eruptions leads to only minor errors on multidecadal time scales. It should be noted that though the shape of each Legendre polynomial is the same for varying time periods, the sign might change since there is no convention for the signs. Therefore we will display the Legendre polynomials each time when they are used.

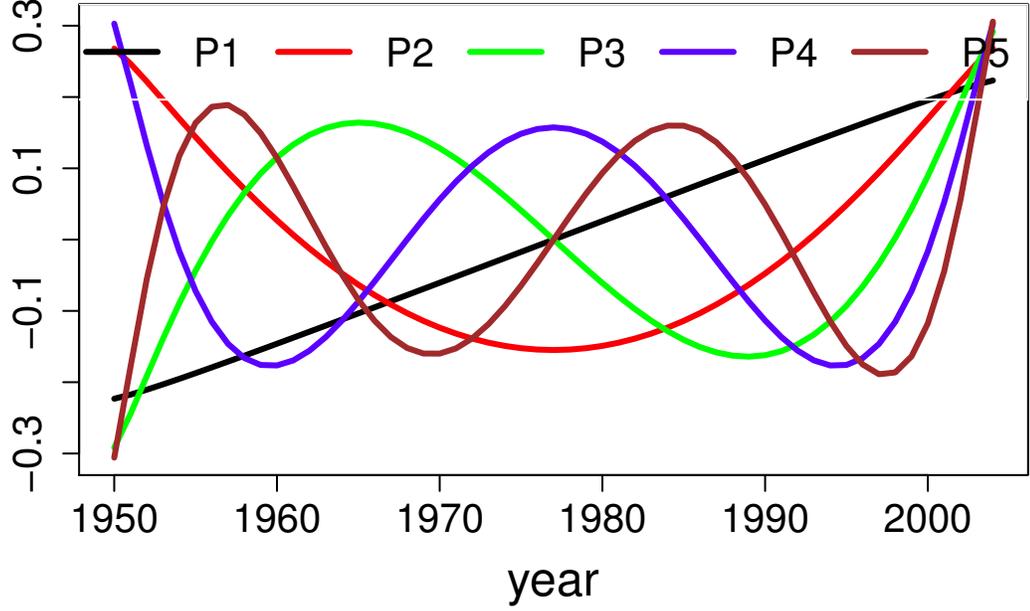


Figure 2.1: Temporal evolution of Legendre polynomials 1-5 for periods of 1950-2004.

Let  $\mathbf{L}$  be the  $N \times J$  matrix whose columns are the leading  $J$  Legendre polynomials. Then, the regression model (2.7) becomes

$$\mathbf{O} = \mathbf{LCP} + \eta, \quad (2.10)$$

where  $\mathbf{C}$  is an  $J \times M$  matrix of coefficients. The above regression model assumes that the spatial structure of the forced responses  $\mathbf{P}$  are known but leaves the associated coefficients of the Legendre polynomials  $\mathbf{C}$  (which gives the time series  $\mathbf{r}_m$ ) to be inferred from data. This re-formulated model also can be expressed in the form of optimal fingerprinting, hence solutions can be derived from standard fingerprinting equations. A detailed derivation of these solutions is given in Appendix A. The result is that the coefficient estimate is

$$\hat{\mathbf{C}} = (\mathbf{P}^T \Sigma_\eta^{-1} \mathbf{P})^{-1} \mathbf{P}^T \Sigma_\eta^{-1} \mathbf{O} \mathbf{L}, \quad (2.11)$$

and the corresponding 95% confidence interval for the  $j$ 'th Legendre polynomial of the  $m$ 'th forced response is

$$\left( \hat{c}_{m,j} - 1.96\sqrt{\left((\mathbf{P}^T\Sigma_\eta^{-1}\mathbf{P})^{-1}\right)_{mm}}, \hat{c}_{m,j} + 1.96\sqrt{\left((\mathbf{P}^T\Sigma_\eta^{-1}\mathbf{P})^{-1}\right)_{mm}} \right). \quad (2.12)$$

## 2.3 Reducing Noise Using Discriminant Analysis

As stated earlier, the goal of our research is to improve estimates of forced responses. One way to improve estimates from optimal fingerprinting is to filter out more noise from the estimated forced response vector. In the standard approach, the forced response vector  $\mathbf{f}_m$  is estimated by running a climate model with the specified forcing and then computing an average over an ensemble of such simulations. For finite ensemble size, some residual internal variability still remains and therefore contaminates the estimated forced response vector. A standard approach is to account for this uncertainty using the Errors in Variables technique (Allen and Stott 2003).

In this thesis, we propose an alternative approach to reducing sampling error. Specifically, the forced response will be estimated using an optimization procedure that maximizes the signal-to-noise (S/N) ratio, i.e., maximizes the ratio of the variance of the forced response to the variance of internal variability. By optimizing the S/N ratio, the procedure removes noise from the forced response vector. These response vectors can be obtained from the method of Jia and DelSole (2012), which decomposes a forced response into a sum of components ordered by S/N ratio, where each component is a fixed pattern multiplied by an associated time series (similar to the way principal component analysis decomposes data by variance). In all cases, only one component has statistically significant S/N ratio, implying that other forced patterns, if they exist, are not detectable after the most detectable pattern has been removed. Jia and DelSole (2012) show that the pattern  $\mathbf{p}_m$  that maximizes the S/N ratio satisfies the eigenvalue problem

$$\Sigma_m \Sigma_\eta^{-1} \mathbf{p}_m = \lambda \mathbf{p}_m, \quad (2.13)$$

where  $\Sigma_m$  is the covariance matrix for the single-forcing simulation for the  $m$ 'th forcing, and anomalies are measured relative to the control simulation. The above technique can estimate the forced response pattern  $\mathbf{p}_m$  from a single realization of the control and forced runs. In particular, this method does not require averaging over an ensemble. The eigenvalue  $\lambda$  equals the total-to-noise ratio, or equivalently,  $S/N + 1$ . Accordingly, we call the eigenvector associated with the largest eigenvalue the *most detectable component*.

## 2.4 Data

To illustrate the above method, model simulations from phase 5 of the Coupled Model Intercomparison Project (Taylor et al. 2012) are used. Estimates of internal variability are derived from 500 years of pre-industrial control runs, which have no year-to-year variation in external forcing. The response to particular forcings are estimated from simulations that are forced over the period 1850-2004 only by anthropogenic aerosols, called AA simulations, and by all forcings *except* anthropogenic aerosols, called noAA simulations. Only two CMIP5 models, CSIRO-QCCCE.CSIRO-Mk3-6-0 (CSIRO) and IPSL.IPSL-CM5A-LR (IPSL), have both noAA and AA simulations, thus only these models are considered. Both CSIRO and IPSL models use the same aerosol emission inventory (Lamarque et al. 2010) over the historical period from 1850 to 2000 in decadal increments at a horizontal resolution of  $0.5^\circ$  in latitude and longitude. CSIRO scaled all BC emissions by 1.25 and all OA emissions by 1.5 (Rotstayn et al. 2012). Both models include aerosol indirect effects and represent a wide range of aerosol sensitivity (Gettelman et al. 2015).

We also use simulations of the recent past (1950-2004) driven by anthropogenic and natural forcing, called *historical* simulations. These simulations will be used as a substitute for observations to test the new methodologies. Analyzing historical simulations can be interpreted as fingerprinting in a “perfect model world” in the sense that observational errors vanish (since data is generated by a climate model) and fingerprinting has nearly perfect knowledge of the forced response pattern and the statistics of internal variability (since these were estimated from dedicated simulations from the same model). Applying optimal

fingerprinting to data from historical simulations, as if the simulations were observations, allows us to test whether optimal fingerprinting is extracting the forced response accurately.

Observational temperature data is from version 4 of the Hadley Centre/Climatic Research Unit data set (HadCRUT4)<sup>1</sup>. Annual means were computed when at least 10 months per calendar year were available. A grid point is included only when annual means are available for every year in the period 1950-2004. The resulting map of available grid points covers 75% of the global surface area and was used to analyze both observations and model data.

The Legendre polynomials used in this section is displayed in fig. 2.1.

## 2.5 Regularizing Optimal Fingerprinting

In the above data set, the spatial dimension far exceeds the temporal dimension. As a result, the covariance matrices  $\Sigma_m$  and  $\Sigma_\eta$  are singular (i.e., they have no inverse). Therefore, in practice, we need to *regularize* the problem by constraining it in some way. The most common approach to regularizing the estimation problem is to project the data onto the leading empirical orthogonal functions (EOFs) of internal variability, i.e., to represent the data in a reduced EOF space. Unfortunately, EOFs are data dependent and lead to difficulties in interpretation and statistical overfitting (Lawley 1956; Allen and Tett 1999; Jones et al. 2013; DelSole and Tippett 2015). Moreover, EOFs of internal variability may not project strongly on the forced responses. Therefore this study uses a different basis set to avoid these problems, namely, the eigenvectors of the Laplacian operator. These eigenvectors depend only on the domain and hence are data independent, and the leading eigenvectors often have simple physical interpretation. The eigenfunctions can be ordered by decreasing spatial scale. For a global domain, Laplacian eigenfunctions are equivalent to spherical harmonics. For restricted domains, such as ocean-only or land-only, the Laplacian eigenfunctions are computed by the method discussed in DelSole and Tippett (2015), which is based on a Green’s function method. The Laplacian eigenvectors for the HadCRUT4

---

<sup>1</sup>downloaded from <http://www.cru.uea.ac.uk/cru/data/temperature/>

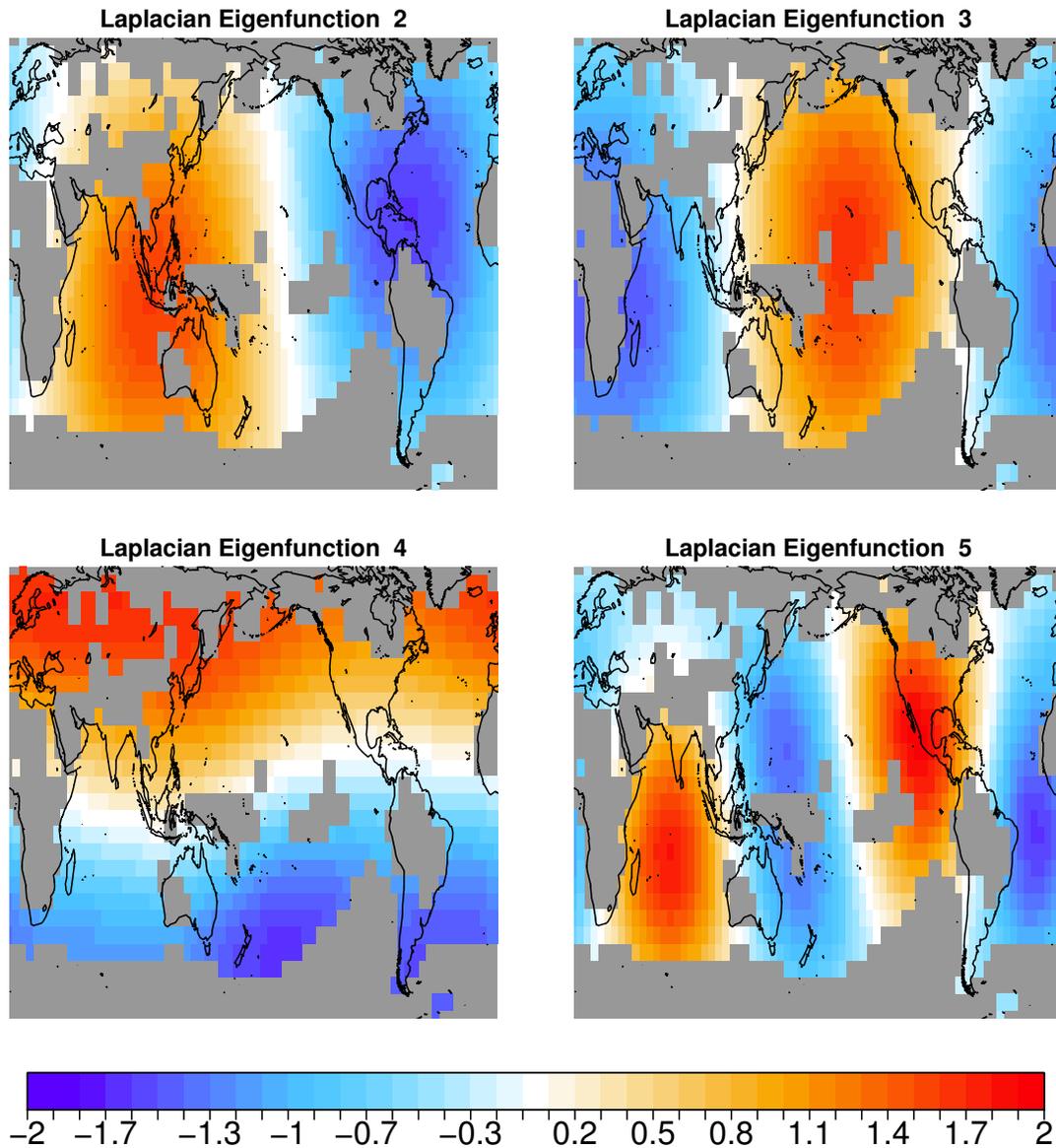


Figure 2.2: Eigenvectors 2-5 of the Laplace operator over the HadCRUT4 observational grid.

observational grid discussed in the previous section are shown in fig. 2.2. For these restricted domains, the first Laplacian eigenfunction still corresponds to a spatially uniform pattern over the domain (and hence is not shown in fig. 2.2), and subsequent eigenvectors correspond to dipoles, tripoles, quadrapoles, etc. of decreasing length scale.

## 2.6 Results

In this section, we illustrate the methodologies discussed in the previous sections using annual mean 2m air temperature (TAS). A key question is how many Laplacian eigenvectors should be chosen to perform fingerprinting— if the dimension is too small then the forced response might escape detection (because it cannot be resolved by the basis vectors), but if the dimension is too large then overfitting becomes problematic. We explored different dimensions to find a choice such that the forced response estimated from historical simulations was consistent with the forced response derived from the single-forcing simulations (i.e., the “truth”). This exploratory analysis suggested that twenty Laplacian eigenvectors was a good choice to keep the consistency, so this choice will be used in the following.

In order to compare different methodologies, each of which involves different response vectors, we estimate individual forced responses using fingerprinting, but then compute the global mean temperature change due to the linear trend in the forced response. This allows the results to be compared using a common variable. This computation is fairly trivial in our framework, since all fingerprints contain the first Laplacian eigenvector, which gives the area mean temperature over the domain, and the trend is either easily computed from the time series multiplied by the scaling factor, or an immediate result of the fingerprinting.

The change in near-global mean temperature over 1950-2004 predicted by climate models and estimated from observations by different methods is shown in fig. 2.3. The thick error bars with crosses shows estimates from HadCRUT4, and error bars with dots show estimates derived from historical simulations. For comparison, thin error bars show estimates from individual ensemble members of the single-forcing runs, which can be viewed as estimates of the “truth” in a perfect model world. In general, estimates from historical simulations (in which AA and noAA are estimated simultaneously) tend to under-estimate the responses in the perfect model world. This bias is expected since errors in the forced response vectors are neglected in our fingerprinting method but tend to bias the estimated forced responses toward zero (Allen and Stott 2003).

We consider the results one method at a time. For traditional fingerprinting, responses

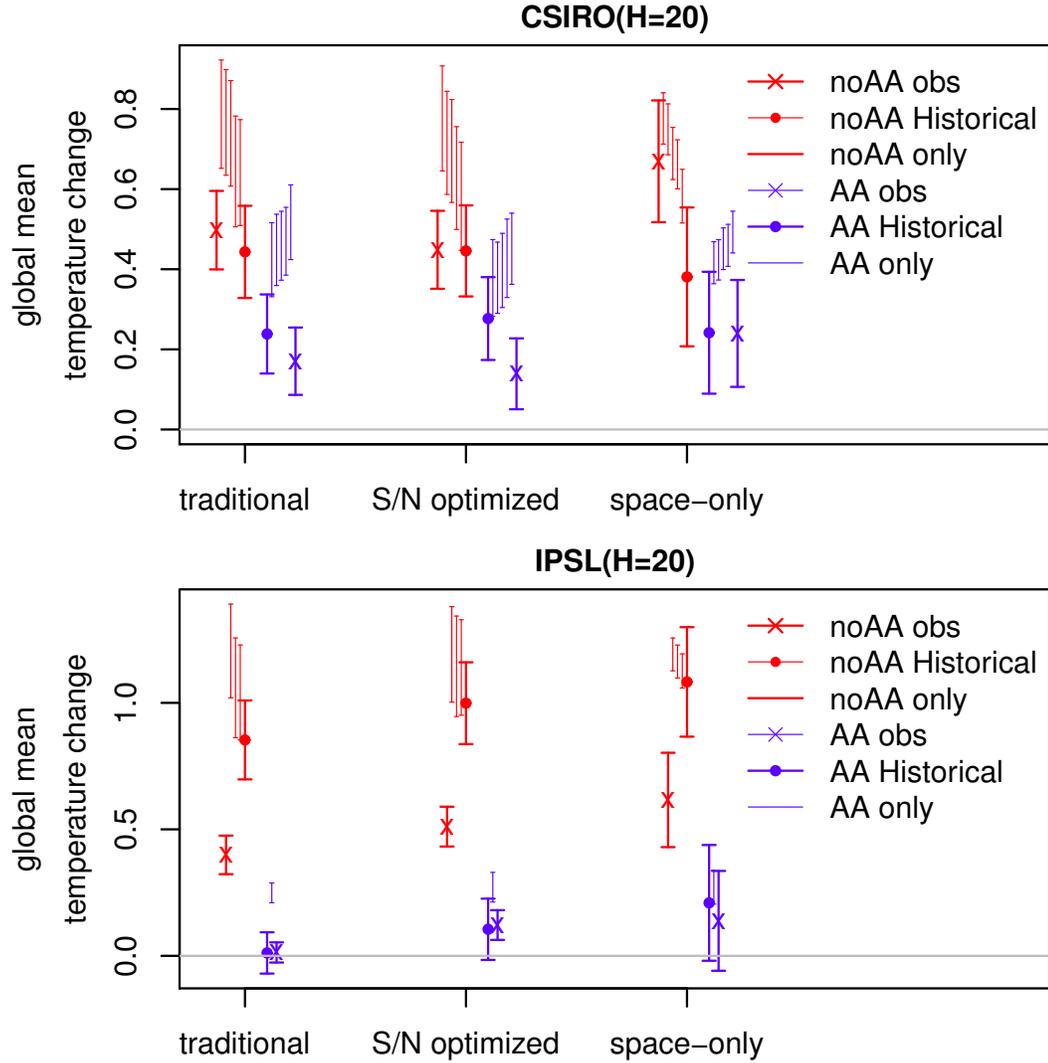


Figure 2.3: 95% confidence intervals of global-mean temperature changes associated with the linear trend due to AA (blue) and noAA (red) forcing in CSIRO (upper) an IPSL (lower), as estimated from HadCRUT4 (thick error bars with crosses), historical simulations (thick error bars with dots) and from single-forcing simulations (thin error bars correspond to different ensemble members), for three different calculation methods: “traditional” means that the forced response vectors are derived from the ensemble means of single forcing simulations; “S/N optimized” means the forced response vectors are derived from the first discriminants of the forced response; and “space-only” means the forced response vectors include only the forced response patterns of single forcing simulations. Signs of changes due to AA have being flipped to be positive.

to AA and noAA forcing are detected using CSIRO fingerprints, and the observed estimates are consistent with model predictions (i.e., the error bars with the dots and crosses overlap in fig. 2.3). In contrast, the response to AA cannot be detected using IPSL fingerprints (i.e., the blue error bars include zero), and the response to noAA is detectable but significantly weaker than the model predicted response. These results imply that IPSL over-estimates the warming due to noAA forcing and the cooling due to AA forcing. Moreover, the structure of the warming due to AA and noAA forcing is not consistent with the observed changes (since the error bars between historical and observations do not overlap).

To investigate whether the results from traditional fingerprinting can be improved, we use forced response vectors that maximize detectability in single-forcing runs. These vectors are determined by solving (2.13). The most detectable components derived from CSIRO for AA and noAA forcing are shown in fig. 2.4, and the most detectable components derived from IPSL are shown in fig.2.5 in the same format. The top rows of fig. 2.4 and 2.5 show the maximized ratios as a function of fingerprints dimension. After an exploration of all truncations from 1 to 25, it is found that the results illustrated here are not sensitive to truncation choices. The top panel shows that only the leading component has statistically significant variance ratio, which implies that there is only one distinct pattern that is detectable in models for this variable, for each forcing. This result confirms that each forced response can be expressed in the separable form (2.8). The middle panels show the projection of the component onto the control run (black) and onto the twentieth century runs (blue and red lines). The ensemble member used in the maximization is shown in red, while an independent ensemble member is shown in blue. If the component were pure internal variability, then there would be no correlation between the time series. The strong agreement between the red and blue curves shows that the method accurately infers the time series of the forced response from a single realization. The temperature patterns associated with the most detectable component are shown in the bottom panels. The patterns for AA and noAA forcing differ slightly in their relative amplitudes of the hemispheric gradient and east-west gradient.

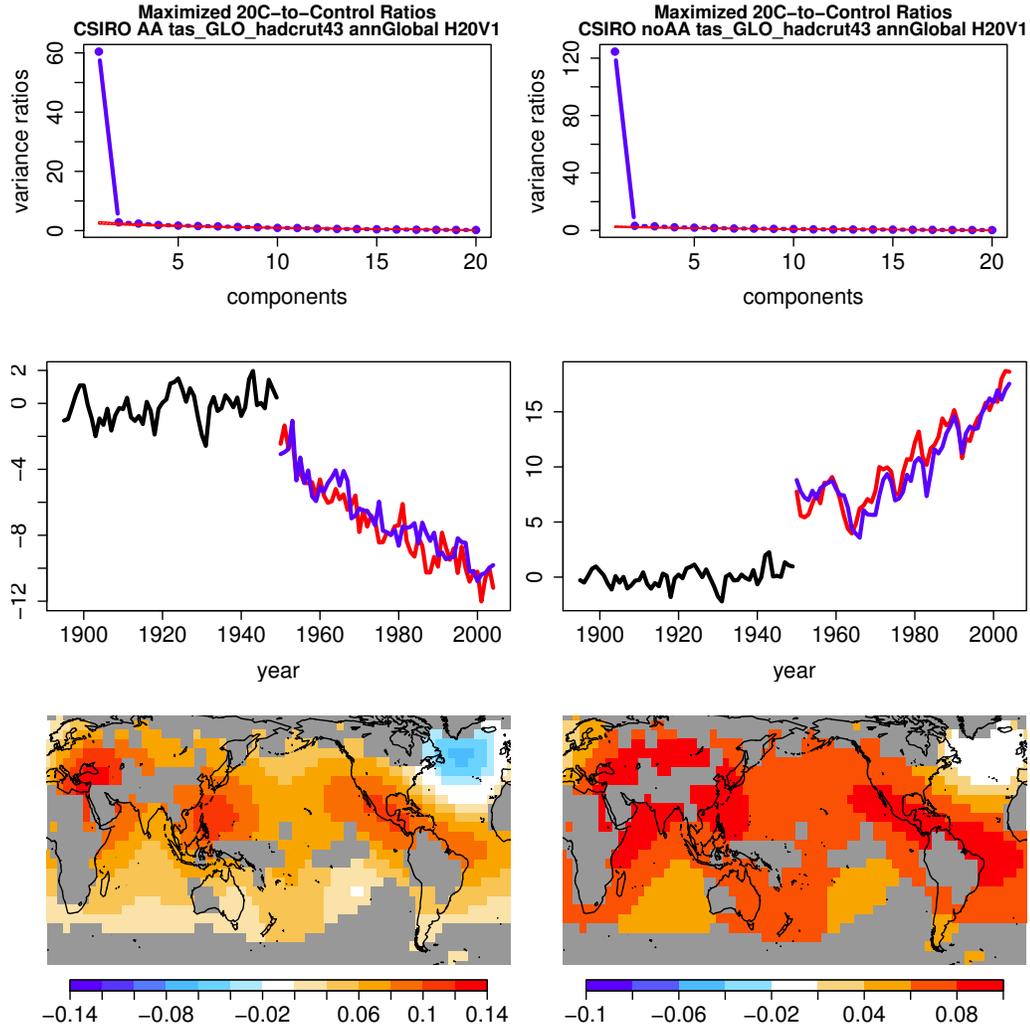


Figure 2.4: Results of maximizing detectability of anthropogenic aerosols (left) and all forcings but anthropogenic aerosols (right) in the CSIRO model for annual mean temperature. The top row shows maximized total-to-noise variance ratios (blue line with dots), and the 95% confidence interval under the hypothesis of equal variances (red shading), as a function of the dimensions of laplacian eigenfunctions. The second row shows time series of the most detectable component in a 55-year segment of the control runs (black) and in the twentieth century (blue and red lines), when the component is represented by 20 laplacian eigenfunctions. The blue curve shows the ensemble member used in the maximization while the red curve shows the component in an independent ensemble member. The third row shows the temperature patterns, respectively, associated with the most detectable anomaly using 20 laplacian eigenfunctions. The temperature pattern has units Kelvin per unit time series.

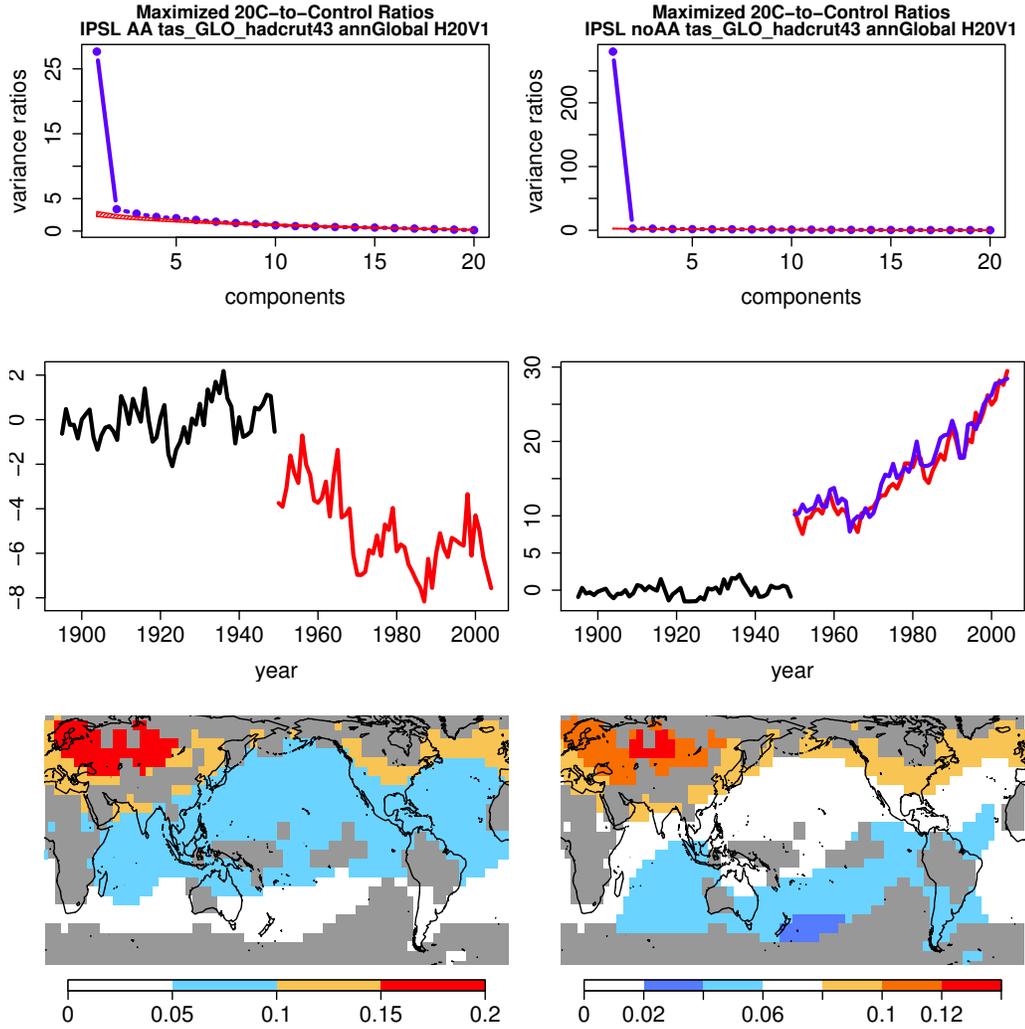


Figure 2.5: Similar to fig. 2.4, but for IPSL

The result of estimating global mean temperature changes using the most detectable components are shown in the second column of fig. 2.3. A comparison between the first and second columns of fig. 2.3 shows that the estimates are generally consistent between the two methods. In particular, the uncertainties are comparable between the two methods. Thus, although we expected to reduce the uncertainties by maximizing detectability, no such improvement was found. Nevertheless, the result represents an advance because fingerprinting based on optimized patterns is much more transparent, in the sense that the

spatial pattern and time series characterizing the forced response can be visualized in a single figure (in contrast, the forced response in traditional fingerprinting is characterized by a full spatio-temporal vector that is not generally separable in space and time).

Next, we consider the result of estimating global mean temperature changes using only the spatial structures of forced responses to AA and noAA forcings, shown in the third column of fig. 2.3. As expected, the uncertainty of the detected changes increases compared to the other two methods, since less information is encoded in the fingerprints. Nevertheless, the results are consistent with the previous two methods. This, too, represents an advance, since it demonstrates that AA and noAA forcing can be detected *and the associated time evolution can be inferred from observations*.

## 2.7 Summary

This chapter proposed novel modifications of optimal fingerprinting for improving estimates of forced responses. First, we proposed filtering out noise from the response vectors by maximizing detectability with respect to single-forcing simulations. Second, we proposed that instead of using the full spatio-temporal evolution of the model forced response, only the spatial structure of the forced response is used in the response vector, leaving the time evolution of the response to be inferred from data. The resulting detection and attribution analysis is argued to be more convincing than the traditional approach because less information from models is assumed as input to the fingerprinting, and more information about the forced responses (e.g., the time evolution) is extracted from observations and compared to models. Because the time evolution is represented by an orthogonal set of temporal polynomials, detection and attribution of time signals can be performed for each polynomial separately, allowing consistency of trends and multi-decadal variability to be assessed separately. Spatial structures were represented using eigenvectors of the Laplace operator over a global domain, which provides an objective basis set for filtering small scale spatial structure that is independent of data.

The results presented in this chapter, and also in later chapters, suggest that optimizing

the forced response patterns does not significantly reduce the uncertainty in estimated responses. Despite this, the result represents an advance, because it demonstrates that much of the signal in the forced response is “rank-1,” which means that the spatial structure and temporal evolution of the forced response can be illustrated in a single figure, which physically clarifies the forcing that is being detected. In addition, we estimated forced responses based only on spatial structure. This approach produced estimates with more uncertainty, as expected since more information was extracted from observations, but the results were nevertheless consistent with those of traditional fingerprinting. We argue that this latter approach gives a compelling demonstration that models correctly simulate the spatial structure of the response to AA and noAA, so much so that the multi-decadal time evolution of these patterns can be derived from observations.

## Chapter 3: Improving Detection and Attribution by Augmenting Additional Variables

The purpose of this thesis is to investigate ways to improve estimates of the response to anthropogenic aerosols. The last chapter explored the use of discriminant analysis to filter out noise from the response vectors. In this chapter, we explore attempts to enhance the signal by using new variables or by combining variables. In principle, we could re-do the calculations presented in the previous chapter for different variables, combinations of variables, and basis vectors, and then compare the results, but this approach would be too overwhelming to comprehend. As part of our thesis research, we developed a framework for comparing the effectiveness of different variables for detecting a specific forced response. A central quantity in this framework is *potential detectability*, which measures the extent to which a forced response can be detected *in a perfect model scenario*. Although observations are not used to evaluate potential detectability, significant potential detectability is a necessary condition for detection in observations— if a forced response cannot be detected in a perfect model scenario, then it cannot provide a legitimate basis for detection and attribution in observations. Thus, it provides a basis for discarding variables that are less effective than temperature and for identifying new variables that could be more effective than temperature. Another attractive property is that it can be evaluated from single-forcing runs. Consequently, potential detectability allows the detectability of a forced response in a model to be compared across a variety of variables so that the best choice of variables for detecting a given forced response can be identified prior to analyzing observations.

The purpose of this chapter is to derive potential detectability. As a result, this chapter is rather technical. However, we can summarize the essence of the measure as follows. The degree to which a forced response can be detected is quantified by the total-to-noise ratio (TNR), which is defined as the ratio of variance of the total variability to the variance of

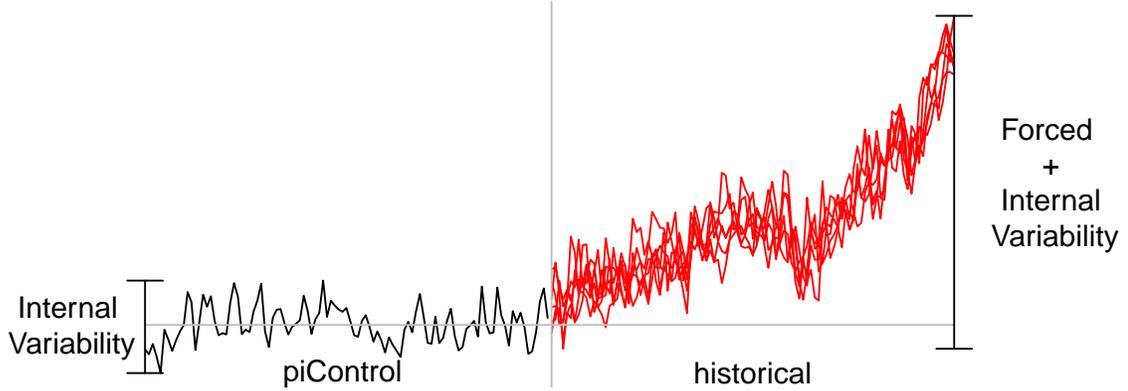


Figure 3.1: Schematic illustration of total, forced and internal variability. Internal variability is simulated by a control simulation (black curve). Total variability is simulated by a historical simulation that contains both internal and forced variability (red curves). Multiple red curves indicate independent ensemble members of the historical simulation.

internal variability. Looking at the schematic in fig. 3.1, it is obvious that the stronger the total-to-noise ratio, the easier forced variability can be detected in a background of internal variability. For a single forcing, potential detectability is identical to the total-to-noise ratio. When the forcing is partitioned into two response vectors (e.g., AA and noAA), potential detectability is proportional to the product of the total-to-noise ratio of the forced response, times a measure of multicollinearity between the two response vectors. That is,

$$\text{potential detectability} = (\text{total-to-noise-ratio}) (1 - \rho^2), \quad (3.1)$$

where  $\rho$  is the cosine of the angle between the two response vectors (in a certain space). To our knowledge, this equation has not appeared in the literature, despite its simplicity and fundamental importance. This result substantially clarifies the relation between detectability, signal-to-noise ratio, and multicollinearity. In particular, a forcing may produce a strong response, and thus have large TNR ratio, but if this response is very similar to the other response being separated, then  $\rho \approx 1$  and the detectability may be weak. Conversely, a response may be relatively weak in a certain set of variables, hence small TNR ratio, but

that response could be very different than the response in a different set of variables, in which case  $\rho \approx 0$  and the detectability could be large.

### 3.1 Potential Detectability for Traditional Fingerprinting

The traditional fingerprinting model is (2.2). The  $m$ 'th forced response is said to be detected if one can reject the null hypothesis  $H_m$  that the coefficient for the  $m$ 'th forced response vanishes:

$$H_m : a_m = 0. \quad (3.2)$$

A standard result is that, under this null hypothesis, the statistic

$$\phi_m = \frac{\hat{a}}{\sqrt{((\mathbf{F}^T \tilde{\Sigma}_N^{-1} \mathbf{F})^{-1})_{mm}}}, \quad (3.3)$$

has a standardized normal distribution. This quantity has the form of a “total-to-noise ratio.” Large absolute values of  $\phi_m$  favors rejection of the null hypothesis. Typically, the forced response  $\mathbf{f}_m$  is normalized so that in a perfect model  $a_m = 1$ . Therefore, the square of the expected value of the detection statistic is

$$\phi_m^2 = ((\mathbf{F}^T \tilde{\Sigma}_N^{-1} \mathbf{F})^{-1})_{mm}^{-1}. \quad (3.4)$$

If there is only a single forced response, say  $\mathbf{f}_m$ , then the detection statistic is

$$\phi_m^2 = \mathbf{f}_m^T \tilde{\Sigma}_N^{-1} \mathbf{f}_m. \quad (3.5)$$

Thus, the TNR ratio for a single forcing is

$$\text{TNR}_m = \mathbf{f}_m^T \tilde{\Sigma}_N^{-1} \mathbf{f}_m. \quad (3.6)$$

If there are two forced responses,  $\mathbf{f}_1$  and  $\mathbf{f}_2$ , then this detection statistic takes the form

$$\mathbf{F}^T \tilde{\Sigma}_N^{-1} \mathbf{F} = \begin{pmatrix} (\mathbf{f}_1^T \tilde{\Sigma}_N^{-1} \mathbf{f}_1) & (\mathbf{f}_1^T \tilde{\Sigma}_N^{-1} \mathbf{f}_2) \\ (\mathbf{f}_2^T \tilde{\Sigma}_N^{-1} \mathbf{f}_1) & (\mathbf{f}_2^T \tilde{\Sigma}_N^{-1} \mathbf{f}_2) \end{pmatrix}. \quad (3.7)$$

Since this is a  $2 \times 2$  matrix, the inverse can be computed analytically as

$$\left( \mathbf{F}^T \tilde{\Sigma}_N^{-1} \mathbf{F} \right)^{-1} = \begin{pmatrix} (\mathbf{f}_2^T \tilde{\Sigma}_N^{-1} \mathbf{f}_2) & -(\mathbf{f}_1^T \tilde{\Sigma}_N^{-1} \mathbf{f}_2) \\ -(\mathbf{f}_2^T \tilde{\Sigma}_N^{-1} \mathbf{f}_1) & (\mathbf{f}_1^T \tilde{\Sigma}_N^{-1} \mathbf{f}_1) \end{pmatrix} \frac{1}{(\mathbf{f}_1^T \tilde{\Sigma}_N^{-1} \mathbf{f}_1)(\mathbf{f}_2^T \tilde{\Sigma}_N^{-1} \mathbf{f}_2)(1 - \rho^2)}, \quad (3.8)$$

where  $\rho^2$  is

$$\rho^2 = \frac{(\mathbf{f}_1^T \tilde{\Sigma}_N^{-1} \mathbf{f}_2)^2}{(\mathbf{f}_1^T \tilde{\Sigma}_N^{-1} \mathbf{f}_1)(\mathbf{f}_2^T \tilde{\Sigma}_N^{-1} \mathbf{f}_2)}. \quad (3.9)$$

The parameter  $\rho$  can be interpreted as the cosine of the angle between  $\mathbf{f}_1$  and  $\mathbf{f}_2$  in “whitened space” (i.e., a space in which the covariance matrix of the noise is the identity matrix) (DelSole and Tippett 2007). Using the above result to evaluate (3.4) gives

$$\phi_m^2 = (\mathbf{f}_m^T \tilde{\Sigma}_N^{-1} \mathbf{f}_m)(1 - \rho^2). \quad (3.10)$$

This is the equation described in (3.1). Note that we set  $a_m = 1$  to derive this equation, so the above  $\phi_m$  measures detectability in a perfect model world.

## 3.2 Potential Detectability for Responses that Are Separable in Space and Time

We showed in the previous chapter that the forced response can be assumed to be separable in space and time with no practical loss of generality. Accordingly, we assume the response vector is of the form (2.8). To use standard optimal fingerprinting equations, the response

vector needs to be “stacked” into a single vector. This can be done as

$$\mathbf{f}_m = \mathbf{r}_m \otimes \mathbf{p}_m, \quad (3.11)$$

where  $\otimes$  indicates the Kronecker product. Substituting this expression into (3.4) and simplifying (and skipping tedious algebraic details) yields

$$TNR_m = \phi_m^2 = (\mathbf{p}_m^T \tilde{\Sigma}_N^{-1} \mathbf{p}_m)(\mathbf{r}_m^T \mathbf{r}_m). \quad (3.12)$$

In the case of two response vectors, we obtain

$$\mathbf{F}^T \tilde{\Sigma}_N^{-1} \mathbf{F} = \begin{pmatrix} (\mathbf{p}_1^T \tilde{\Sigma}_N^{-1} \mathbf{p}_1)(\mathbf{r}_1^T \mathbf{r}_1) & (\mathbf{p}_1^T \tilde{\Sigma}_N^{-1} \mathbf{p}_2)(\mathbf{r}_1^T \mathbf{r}_2) \\ (\mathbf{p}_2^T \tilde{\Sigma}_N^{-1} \mathbf{p}_1)(\mathbf{r}_2^T \mathbf{r}_1) & (\mathbf{p}_2^T \tilde{\Sigma}_N^{-1} \mathbf{p}_2)(\mathbf{r}_2^T \mathbf{r}_2) \end{pmatrix} \quad (3.13)$$

Computing the inverse of this matrix analytically and following similar steps as before gives,

$$\phi_m^2 = TNR_m(1 - \rho_{time}^2 \rho_{space}^2) \quad (3.14)$$

where,

$$\rho_{time} = \frac{\mathbf{r}_1^T \mathbf{r}_2}{\sqrt{\mathbf{r}_1^T \mathbf{r}_1} \sqrt{\mathbf{r}_2^T \mathbf{r}_2}} \quad (3.15)$$

and

$$\rho_{space} = \frac{\mathbf{p}_1^T \tilde{\Sigma}_N^{-1} \mathbf{p}_2}{\sqrt{\mathbf{p}_1^T \tilde{\Sigma}_N^{-1} \mathbf{p}_1} \sqrt{\mathbf{p}_2^T \tilde{\Sigma}_N^{-1} \mathbf{p}_2}}. \quad (3.16)$$

The parameter  $\rho_{time}$  measures the cosine of the angle between the response time series. The parameter  $\rho_{space}$  measures the cosine of the angle between the response vectors in “whitened space” (DelSole and Tippett 2007), and can be interpreted as a generalized spatial correlation. The squared temporal (pattern) correlation is always between 0 and 1, where one implies that the response time series are perfectly collinear while zero implies

that the response time series are orthogonal. Expression (3.14) shows that detectability decreases as either the response time series or spatial patterns become more collinear.

The total-to-noise ratio  $TNR_m$ , temporal correlation  $\rho_{time}$  and pattern correlation  $\rho_{space}$  can be computed immediately after the forced response has been obtained from the single-forcing runs and separated in space and time. Accordingly, we define *potential detectability* to be equation (3.14) after substituting  $TNR_m$  from single-forcing runs. It follows that potential detectability can be computed directly from the single-forcing runs and does not require simulations with multiple forcings. We also define the *realized detectability* to be equation (3.14) after substituting  $TNR_m$  derived from coefficient estimates  $\hat{a}_m$  from historical simulations

### 3.3 Potential Detectability Using only Spatial Information

For fingerprinting based on spatial patterns only, the regression model is (2.10). The  $m$ 'th forced response is said to be detected if one can reject the null hypothesis  $H_m$  that all polynomial coefficients for the  $m$ 'th forced response vanishes:

$$H_m : c_{m,1} = c_{m,2} = \dots = c_{m,J} = 0. \quad (3.17)$$

Under this null hypothesis, the statistic

$$\phi_m^2 = \frac{\sum_{j=1}^J \hat{c}_{m,j}^2}{\left( (\mathbf{P}^T \Sigma_\eta^{-1} \mathbf{P})^{-1} \right)_{mm}}, \quad (3.18)$$

has a chi-squared distribution with  $J$  degrees of freedom. The quantity  $\phi_m^2$  measures the *detectability* of the  $m$ 'th forced response and must exceed a critical value for a forced response to be detected.

If only one forced response vector (i.e.,  $M = 1$ ) is considered, the covariance matrix for the Legendre coefficients reduces to the scalar  $\text{cov}[\hat{\mathbf{c}}_j] = (\mathbf{p}_m^T \Sigma_\eta^{-1} \mathbf{p}_m)^{-1}$ . Then the

detectability (3.18) of one forced response vector becomes

$$\phi_m^2[M = 1] = \left( \sum_{j=1}^J \hat{c}_{m,j}^2 \right) (\mathbf{p}_m^T \Sigma_\eta^{-1} \mathbf{p}_m). \quad (3.19)$$

Note that this expression is of the form of a total-to-noise ratio:

$$\phi_m^2[M = 1] = \frac{\sum_{j=1}^J \hat{c}_{m,j}^2}{\text{cov}[\hat{\mathbf{c}}_j]} = TNR_m, \quad (3.20)$$

where  $TNR_m$  denotes for the total-to-noise ratio of the forced response relative to the internal noise

Now we consider the case of two forced response patterns  $\mathbf{p}_1$  and  $\mathbf{p}_2$ . In this case,

$$\mathbf{P}^T \Sigma_\eta^{-1} \mathbf{P} = \begin{pmatrix} \mathbf{p}_1^T \Sigma_\eta^{-1} \mathbf{p}_1 & \mathbf{p}_1^T \Sigma_\eta^{-1} \mathbf{p}_2 \\ \mathbf{p}_2^T \Sigma_\eta^{-1} \mathbf{p}_1 & \mathbf{p}_2^T \Sigma_\eta^{-1} \mathbf{p}_2 \end{pmatrix}. \quad (3.21)$$

The inverse of this matrix can be computed analytically as

$$(\mathbf{P}^T \Sigma_\eta^{-1} \mathbf{P})^{-1} = \frac{1}{\det[\mathbf{P}^T \Sigma_\eta^{-1} \mathbf{P}]} \begin{pmatrix} (\mathbf{p}_2^T \Sigma_\eta^{-1} \mathbf{p}_2) & -(\mathbf{p}_1^T \Sigma_\eta^{-1} \mathbf{p}_2) \\ -(\mathbf{p}_2^T \Sigma_\eta^{-1} \mathbf{p}_1) & (\mathbf{p}_1^T \Sigma_\eta^{-1} \mathbf{p}_1) \end{pmatrix}. \quad (3.22)$$

The determinant is

$$\det[\mathbf{P}^T \Sigma_\eta^{-1} \mathbf{P}] = (\mathbf{p}_1^T \Sigma_\eta^{-1} \mathbf{p}_1) (\mathbf{p}_2^T \Sigma_\eta^{-1} \mathbf{p}_2) (1 - \rho^2), \quad (3.23)$$

where

$$\rho = \frac{\mathbf{p}_1^T \Sigma_\eta^{-1} \mathbf{p}_2}{\sqrt{\mathbf{p}_1^T \Sigma_\eta^{-1} \mathbf{p}_1} \sqrt{\mathbf{p}_2^T \Sigma_\eta^{-1} \mathbf{p}_2}}. \quad (3.24)$$

Consolidating the above results into the detection statistic (3.18) gives

$$\phi_m^2[M = 2] = \left( \sum_{j=1}^J \hat{c}_{m,j}^2 \right) (\mathbf{p}_m^T \Sigma_\eta^{-1} \mathbf{p}_m) (1 - \rho^2). \quad (3.25)$$

A comparison of (3.19) and (3.25) shows that the detectability of the response when two forced responses are present is equal to the detectability of the response associated with one response vector scaled by a collinearity term  $(1 - \rho^2)$ , i.e.,

$$\phi_m^2 = TNR_m (1 - \rho^2). \quad (3.26)$$

The parameter  $\rho_{space}$  measures the cosine of the angle between the response vectors in “whitened space” (DelSole and Tippett 2007), and can be interpreted as a generalized spatial correlation. The squared pattern correlation  $\rho_{space}^2$  is always between 0 and 1, where one implies that the response vectors are perfectly collinear while zero implies that the response vectors are orthogonal. Expression (3.26) shows that detectability decreases as the response vectors become more collinear.

The total-to-noise ratio  $TNR_m$  and pattern correlation  $\rho_{space}$  can be computed immediately after the response vectors have been derived from the single-forcing runs. Accordingly, we define *potential detectability* to be equation (3.26) after substituting  $TNR_m$  from single-forcing runs. It follows that potential detectability can be computed directly from the single-forcing runs and does not require simulations with multiple forcings. Also, we define *realized detectability* to be equation (3.26) after substituting  $TNR_m$  derived from coefficient estimates  $\hat{c}_{m,j}$  from historical simulations

## **Chapter 4: What Surface Observations are Important for Separating the Influences of Anthropogenic Aerosols From Other Forcings?**

In this chapter, we apply the concept of potential detectability, as derived in the previous chapter, to identify variables or combinations thereof that may be effective for detecting anthropogenic aerosol cooling. This investigation will be performed in a perfect model world, using no observations, to illustrate the methodology and to identify the range of possibilities for improving estimates of aerosol cooling.

We use the same model data as discussed in sec. 2.4. Moreover, we analyze annual means of three commonly used variables for climate detection: 2m air temperature (TAS), precipitation (PR) and sea level pressure (PSL). However, no observations are used. Since there are no restrictions due to missing data, we consider global fields. The Laplacian eigenvectors used to represent global data are displayed in fig. 4.1. The first eigenfunction, not shown, is merely a constant and corresponds to the global mean. The second eigenfunction measures the gradient between Northern and Southern Hemispheres, the next two measure the east-west gradients, and the fifth measures the pole-to-equator gradient. The Legendre polynomials shown in fig. 4.2 are used to represent time variability of the forced response.

### **4.1 Detection and Attribution Based on Temperature and Precipitation**

Since detection and attribution with multiple variables has not been discussed yet, we first illustrate our procedure by analyzing a single case in detail. Accordingly, we perform a detection analysis of AA and noAA forced responses using joint temperature and precipitation data. The detection analysis is performed on a single ensemble from a historical simulation,

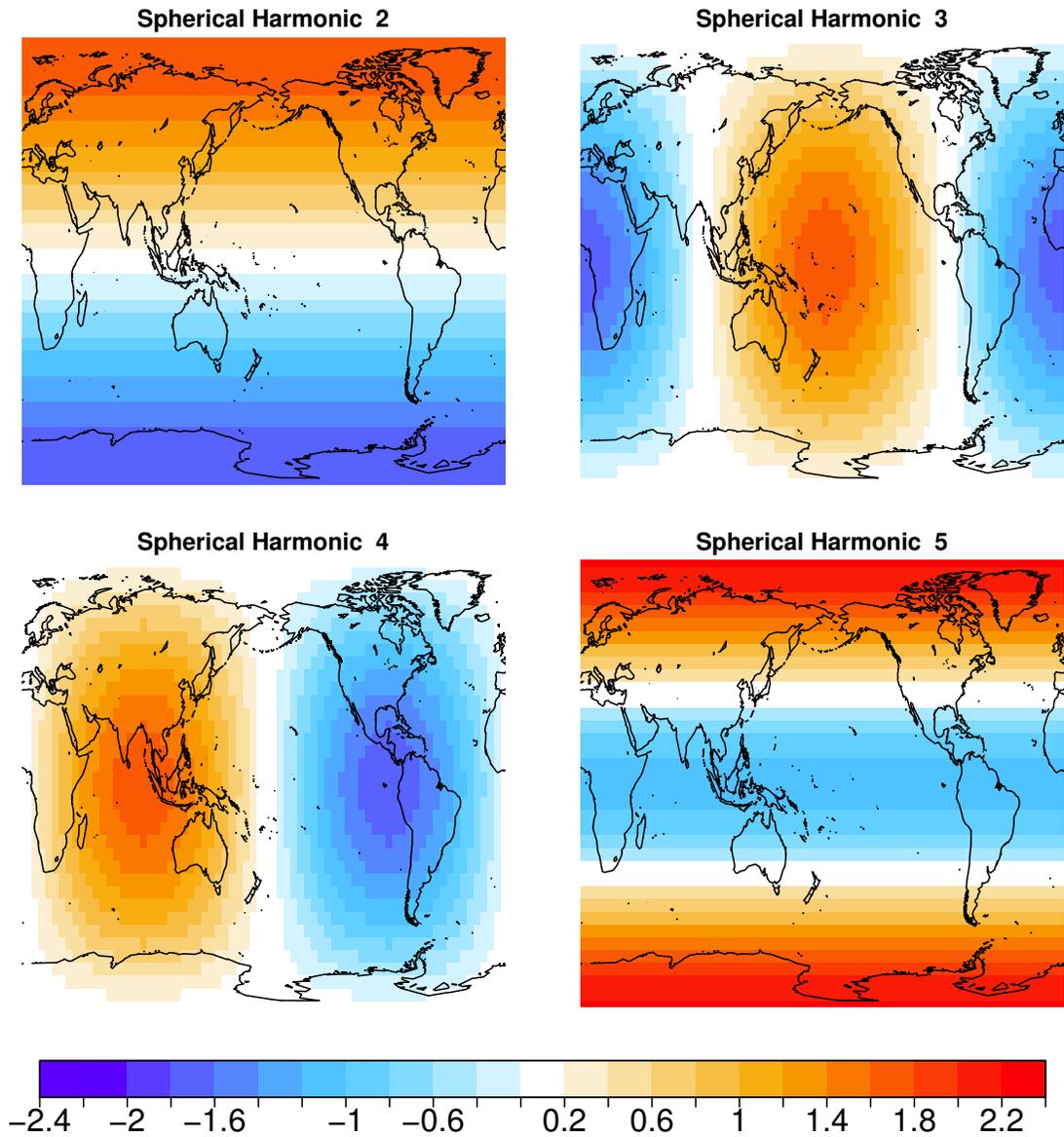


Figure 4.1: Spatial patterns of Spherical Harmonics 2, 3, 4, 5

which serves as a surrogate for observations, thereby avoiding complications due to model error and observational error. There are ten such ensemble members for CSIRO and six such ensemble members for IPSL. We first consider a state vector based on the first Laplacian eigenvector (i.e., the spatially uniform pattern) for temperature and precipitation. At one

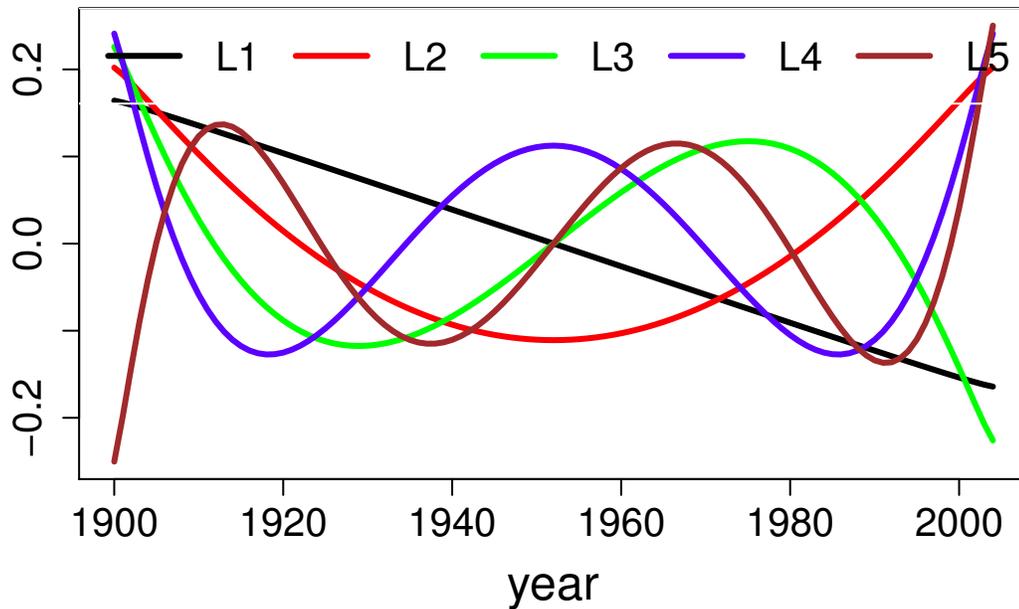


Figure 4.2: Temporal evolution of Legendre polynomials 1-5 for periods of 1900-2004.

time step, this vector contains only two elements: global average temperature and precipitation. The response vectors  $\mathbf{p}_{AA}$  and  $\mathbf{p}_{noAA}$  that maximize detectability with respect to internal noise are determined. Because the 2-dimensional vectors are unique up to a scaling factor, only the ratio of the two elements is unique, which has units of  $\%/K$  and is called *hydrological sensitivity*. Thus, for one Laplacian eigenvector each, fingerprinting based on temperature and precipitation is tantamount to estimating response time series based on differences in hydrological sensitivity between different forcings. For comparison, we also consider response vectors using 20 spherical harmonics for temperature and 20 spherical harmonics for precipitation.

The estimated time series of global mean temperature change using five Legendre polynomials are shown in fig. 4.3. Different curves of the same color show results for different ensemble members of the historical run (10 for CSIRO; 6 for IPSL). The figure also shows

global mean temperature change estimated from the AA and noAA simulations separately without temporal smoothing (A.25). The top and bottom rows show results derived based on 1 and 20 spherical harmonics, respectively, while the left and right columns show results derived from CSIRO and IPSL, respectively. As is evident from the figure, the global mean temperature changes for AA and noAA are estimated very well in this “perfect model” case. In fact, the time series of global mean temperature change estimated from historical runs look essentially like smoothed versions of those obtained from the AA and noAA runs, even though the two time series were estimated independently. Moreover, global mean temperature change time series for different ensemble members are relatively close to each other, indicating that the estimates are not sensitive to sampling errors.

The impact of adding spatial gradient information in the forced response patterns is shown in the bottom row. Based on twenty harmonics for temperature and twenty harmonics for precipitation, the time series estimated from the CSIRO historical run (lower left panel) are generally closer together (hence less sensitive to internal variability) and closer to the single-forcing runs (hence more accurate). In contrast, time series for IPSL have changed relatively little, suggesting that detectability is not strongly improved by adding spatial gradient information. Thus, the importance of spatial gradient information for separating forced responses is model dependent.

So far in this study, optimal fingerprinting has been applied in the context of a “perfect model world,” in the sense that the same model generates the data and generates the forced response used in fingerprinting. In practice, the forced response vectors used in fingerprinting differ from the “true” responses that influence observations. To assess model error, one often examines the sensitivity of estimated responses to response vectors derived from different models. Accordingly, we show in fig. 4.4 the change in global mean temperature in single-forcing simulations attributable to AA and noAA forced response vectors derived from different models. The top and bottom rows show results based on 1 and 20 spherical harmonics, respectively, while the left and right columns show results derived from CSIRO and IPSL single-forcing runs, respectively. The figure shows that the change in global mean

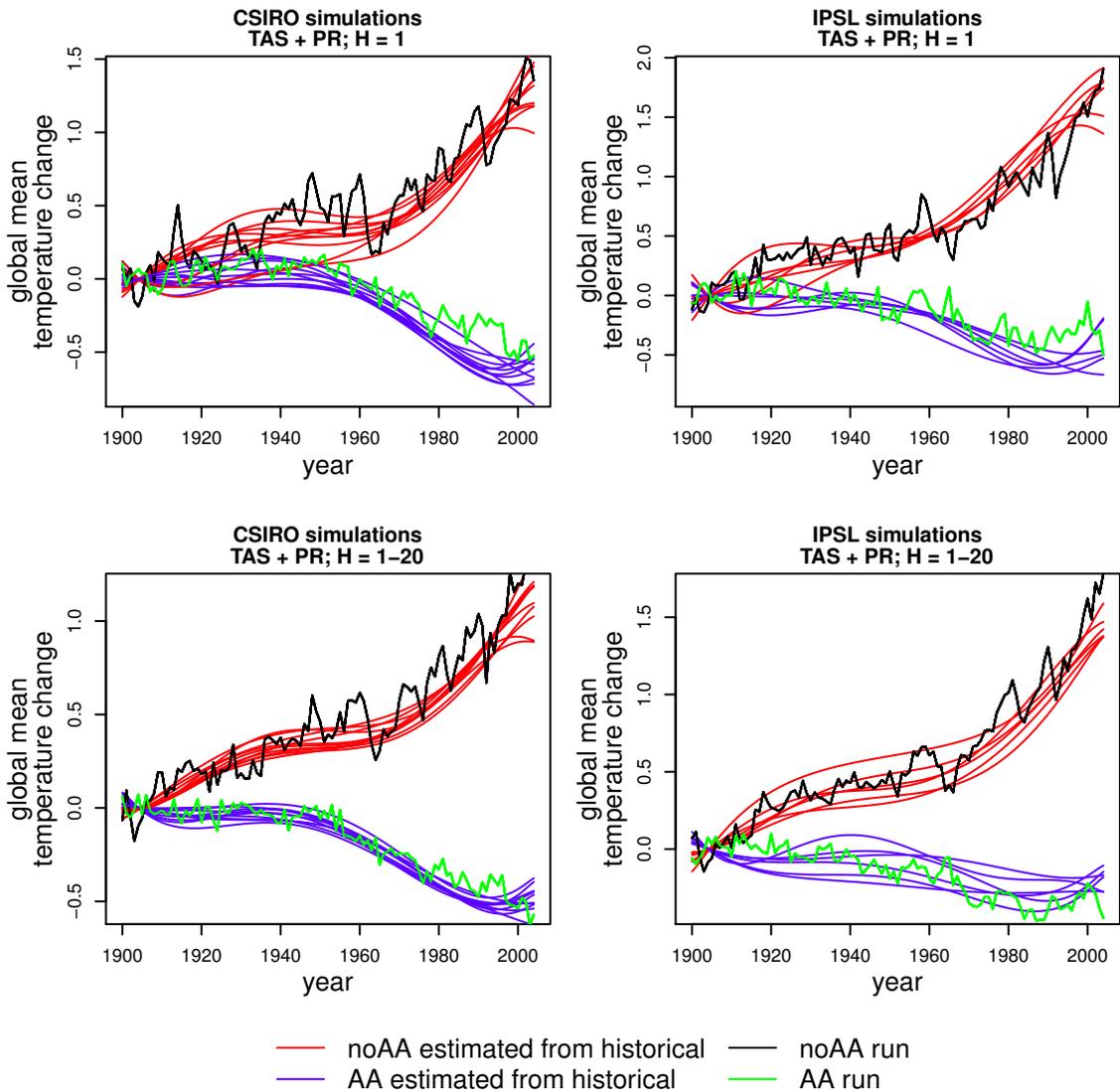


Figure 4.3: Time series of global mean temperature change calculated from historical simulations (blue and red), and from single-forcing AA runs (green) and noAA runs (black), using the joint annual mean temperature-precipitation response vector. Different panels correspond to different models and different number of Spheric harmonics, as schematically illustrated in the title. Time series from historical simulations are based on using five Legendre polynomials, whereas time series based on individual forcing runs are based on year-by-year estimation. The response vector is derived from only one ensemble member of the individual forced runs and then used to estimate global mean temperature change time series for each ensemble member from the historical run.

temperature is relatively insensitive to the model response vector. Thus, while response vectors differ between models, these differences do not lead to significantly different conclusions

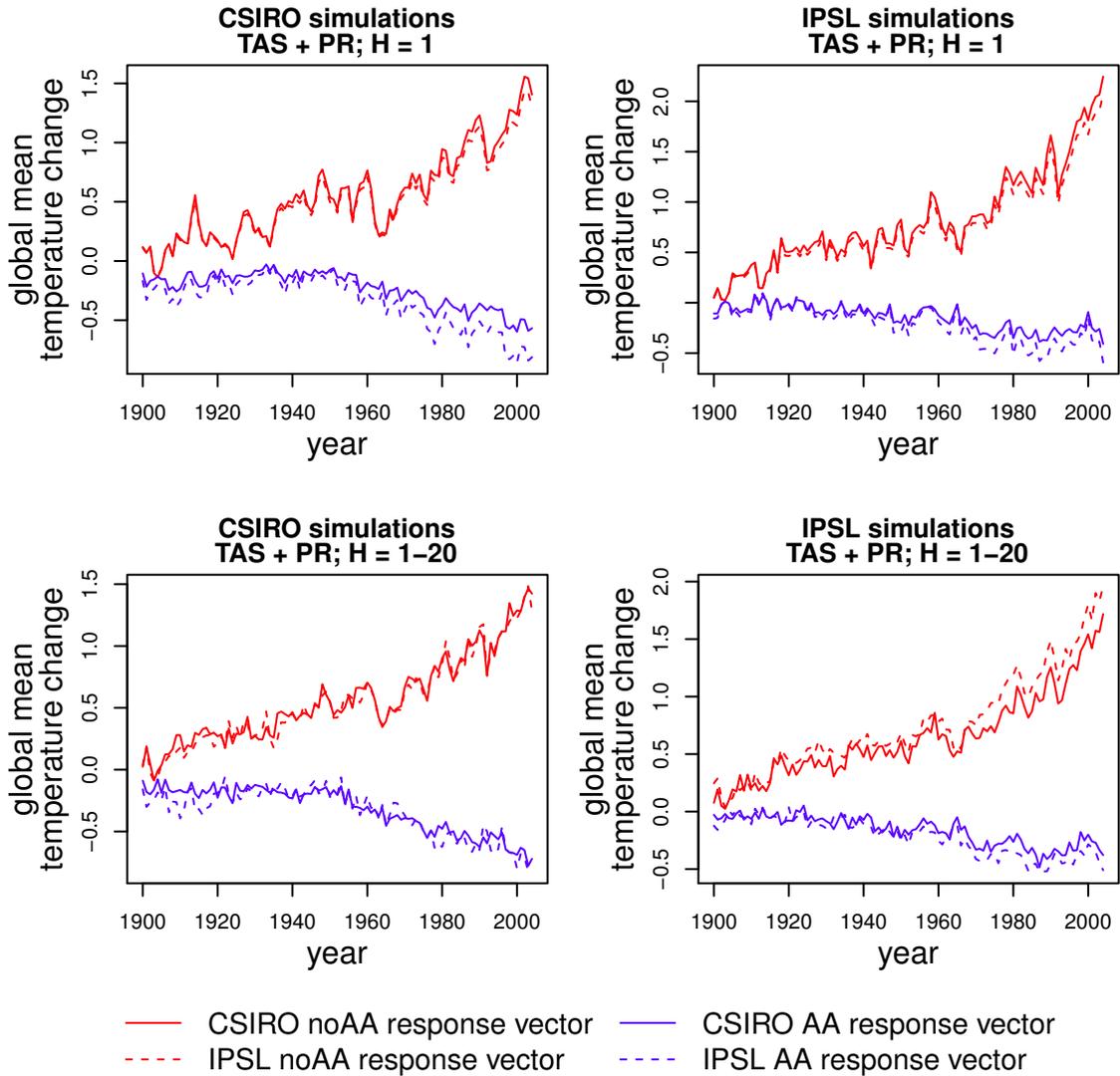


Figure 4.4: Time series of global mean temperature change obtained from single-forcing AA runs (green) and noAA runs (red), using the joint annual mean temperature-precipitation response vector derived from the same model (solid) or another model (dashed). Different panels correspond to different models and different number of Spherical harmonics, as schematically illustrated in the title.

about climate changes attributable to AA and noAA.

The response patterns  $\mathbf{p}_{AA}$  and  $\mathbf{p}_{noAA}$  used to derive time series in bottom panels of fig. 4.3 and 4.4 are shown in fig. 4.5. Recall that for each model's single-forcing run, this pattern is the only component with statistically significant signal-to-noise ratio (fig.

4.6). Because the state vector depends on the joint temperature-precipitation fields, the response pattern is characterized by a *pair* fields, shown along each row. Multiplying these patterns by their respective time series gives the spatio-temporal response to AA and to noAA forcing in each model. The models agree that noAA forcing tends to warm the globe and AA forcing tends to cool the globe. However, CSIRO suggests that the noAA-warming is approximately equally distributed between Northern and Southern Hemispheres, whereas IPSL suggests that noAA-warming is dominated in the Northern Hemisphere. The models also agree that noAA will tend to enhance precipitation and AA will tend to decrease precipitation over the western Pacific, near the maritime continent, and over the Atlantic and Pacific centered near central America. However, the models disagree in the direction of precipitation changes over the Indian ocean. The fact that the AA and noAA response patterns are similar implies that the patterns have a large squared pattern correlation  $\rho^2$ . Despite this large pattern correlation, the total-to-noise ratio is sufficiently strong to overcome multicollinearity problems to produce stable time series estimates.

Confidence intervals for the polynomial coefficients (in the perfect model case) are shown in fig. 4.7. A forced response is said to be detected when the interval derived from the historical run does not include zero, and is said to be attributable to AA or noAA forcing when the interval from the historical run overlaps with the appropriate interval from the single-forcing run. Almost all forced responses are attributable to the appropriate forcing, as expected because the data comes from a perfect model experiment. For both CSIRO and IPSL, the linear trend and the quadratic growth are detected (i.e., coefficients associated with the first two Legendre polynomials differ significantly from zero). Based on all variable combinations that we have analyzed, adding more Laplacian eigenvectors tends to improve the ability to detect and attribute changes in higher order polynomials in time. However, as the order of the polynomial increases, so too does the number of fitted parameters and the likelihood of overfitting. We have found that 20 Laplacian eigenvectors produces stable results in a perfect model scenario. Since multidecadal variability is associated with polynomials of degree 3-6, these results suggest that detecting and attributing multidecadal

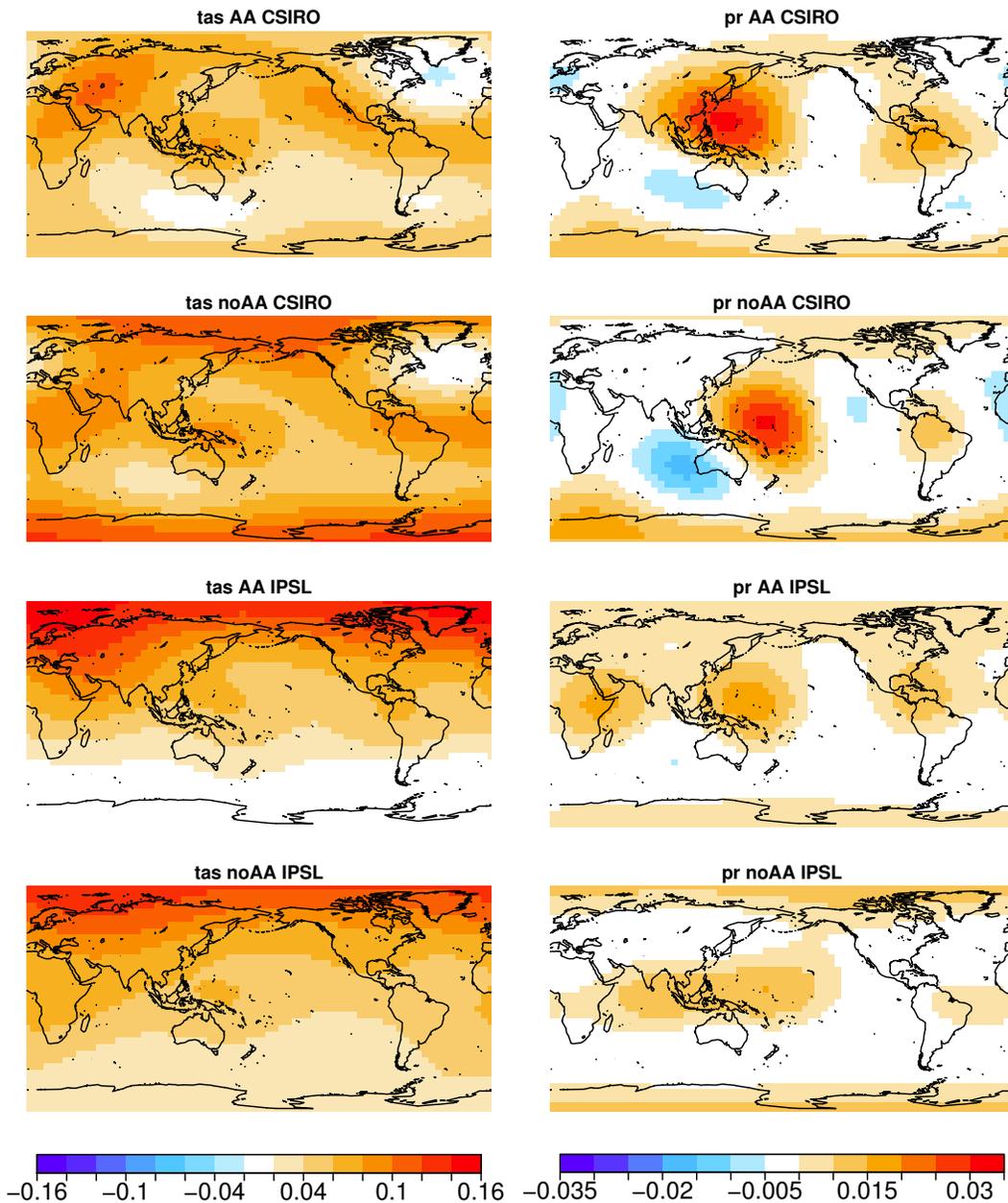


Figure 4.5: Response vectors that maximize detectability of AA forcing and noAA forcing in two different models (CSIRO and IPSL) based on 20 Laplacian eigenvectors and annual mean fields. The response vector depends on a joint temperature-precipitation relation and therefore is characterized by a pair of patterns, shown in each row. The left column shows temperature fields and the right column shows precipitation fields. The corresponding forcing and model is indicated in the title of each figure. These response patterns were used to estimate the time series shown in fig. 4.4.

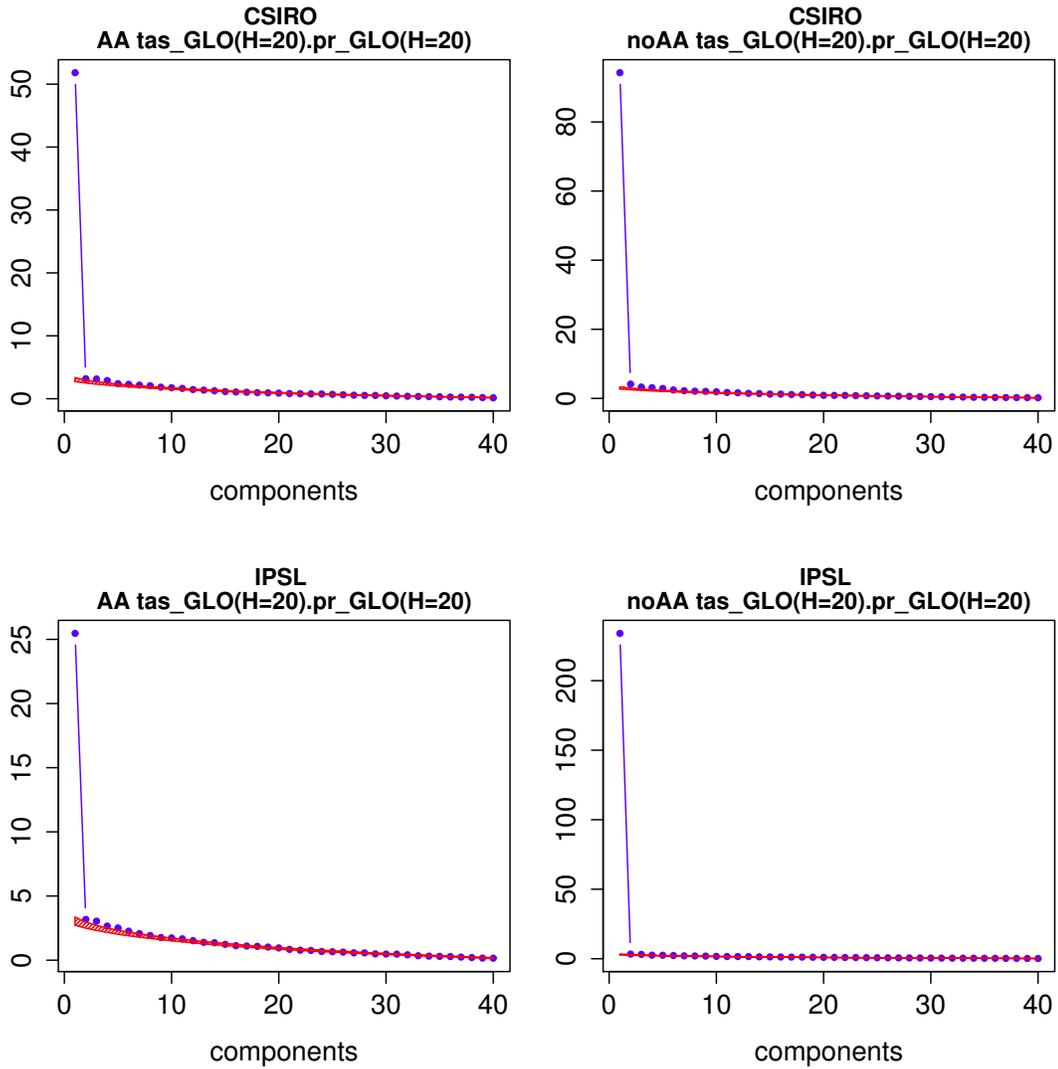


Figure 4.6: Maximized signal-to-noise variance ratios (blue line with dots), and the 95% confidence interval under the hypothesis of equal variances (red shading), as a function of the dimensions of fingerprints, based on 20 Laplacian eigenvectors and annual mean fields of each variable, for AA single-forcing runs of CSIRO (top left) and IPSL (bottom left), and noAA single-forcing runs of CSIRO (top right) and IPSL (bottom right).

variability to forcings in a perfect model requires spatial gradient information.

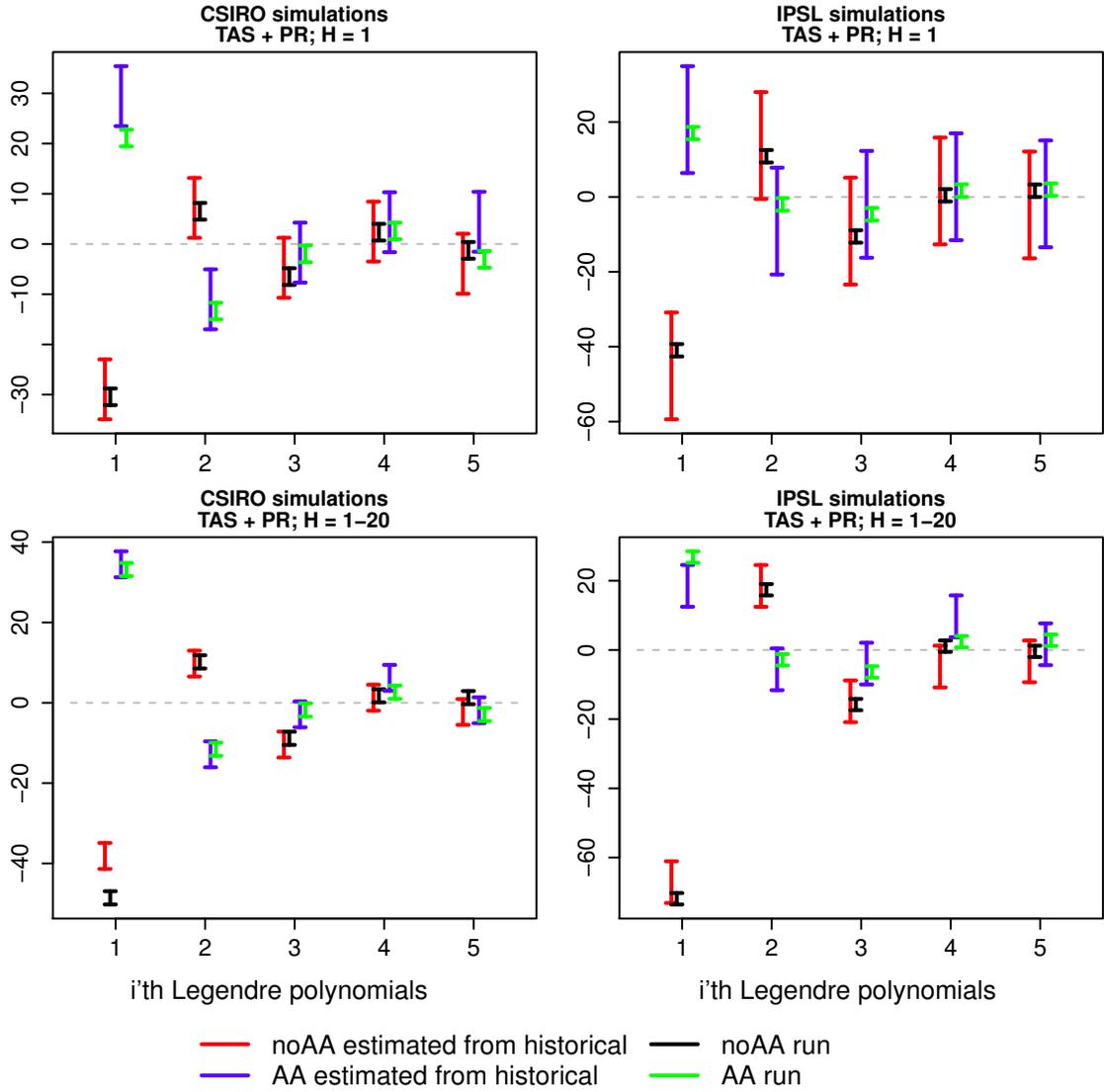


Figure 4.7: 2.5% and 97.5% confidence limits for the coefficients of the five Legendre polynomials (in time) estimated from CSIRO (left column) and IPSL (right column), using 1 Laplacian eigenvector (top row) and 20 Laplacian eigenvectors (bottom row). The dashed line indicates zero. Only the perfect model case, in which all parameters and data come from the same model, is shown.

## 4.2 Detectability Under Different Combinations of Variables, Spatial Structures, and Seasons

We now explore detectability based on various combinations of variables. To summarize the results, we show the potential detectability measure  $\phi_m^2$  in (3.25) using the total-to-noise

ratios derived from the single-forcing runs. To precisely state the state vector, let  $Y_{v,l}^{\Delta T, \text{domain}}$  denote the  $v$ 'th physical variable– surface air temperature (TAS), precipitation (PR) or sea level pressure (PSL)– for the  $l$ 'th Laplacian eigenvector for time averaging period  $\Delta T$  and geographic “domain.” For a variable  $v$  on a global domain averaged over a year (“ann”), the state vector is of the form

$$\mathbf{p} = \left( Y_{v,1}^{\text{ann, globe}} \quad Y_{v,2}^{\text{ann, globe}} \quad Y_{v,3}^{\text{ann, globe}} \quad \dots \right)^T. \quad (4.1)$$

In this case, as the dimension of the vector increases, more spatial gradient information is included through higher order spherical harmonics. The vector is calculated by maximizing total-to-noise ratio in the first ensemble member of the appropriate forced run from each model. The potential detectability  $\phi_m^2$  as a function of vector dimension using 3 Legendre polynomials are shown as black symbols in fig. 4.8 (the other symbols will be discussed shortly).

Potential detectability slightly above the significance threshold implies that the signal is marginally above the noise level for detection. While such values are desirable because they meet the criterion for detection, in practice, we are interested in response vectors whose potential detectability is several times higher than the threshold, because these vectors can detect a signal that is several times larger than the noise level. To be clear, we would not select the vector that maximizes detectability *in the training sample*, because of overfitting.

The figure shows that, in general, TAS and PR lead to stronger detection than PSL. PR gives stronger detection of AA than any other single variable. PSL cannot detect AA forced responses if too few spherical harmonics ( $< 8$ ) are included in the state vector. In CSIRO, the potential detectability jumps after including the second and fifth spherical harmonic, suggesting that the hemispheric gradient and pole-to-equator gradient provides critical information for separating AA and noAA forced responses. Such jumps are not evident in IPSL, indicating that the importance of spatial gradient information to separating AA and noAA forced responses is model dependent. Note that AA or noAA forcing individually

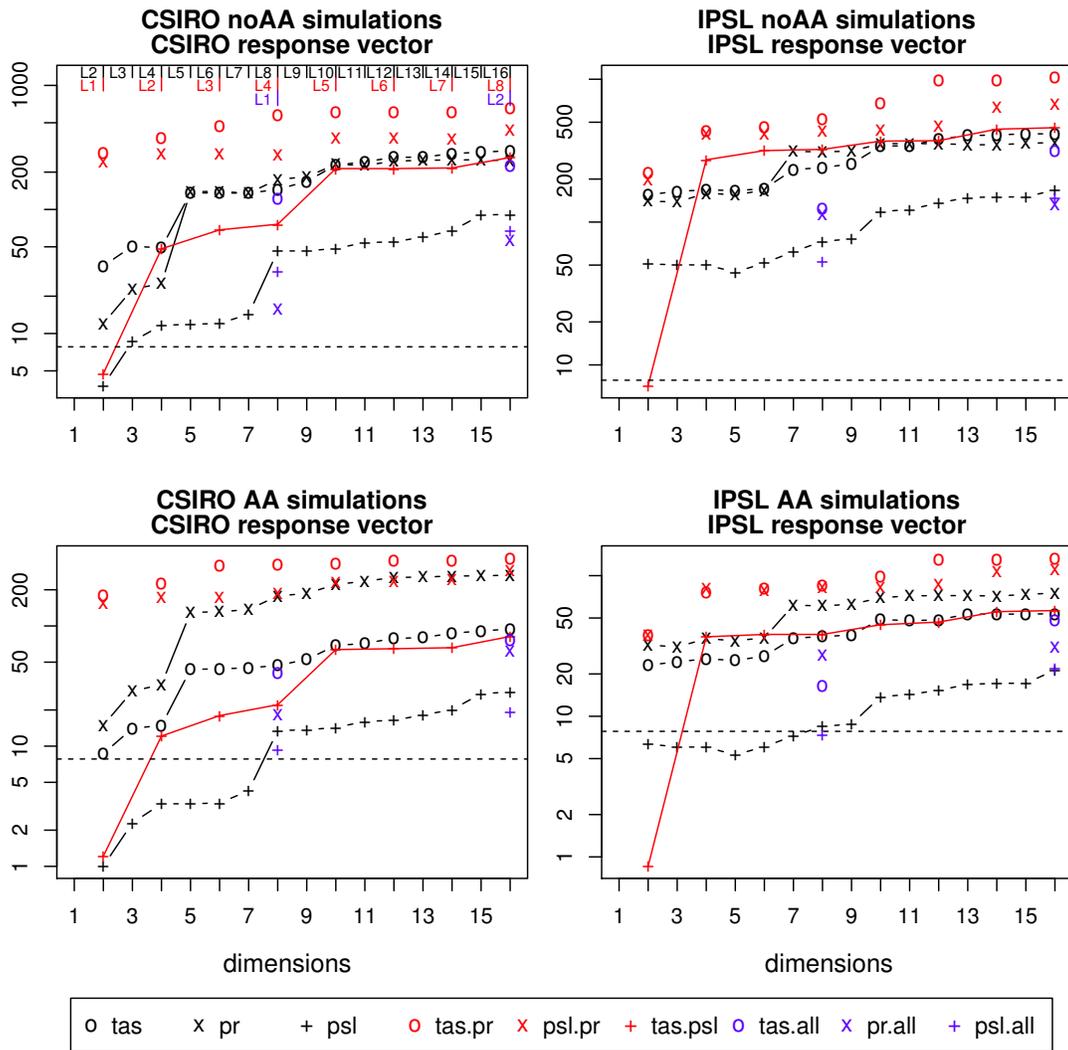


Figure 4.8: Potential detectability in a “perfect model” cases, in which the response patterns and noise covariance matrix are from a same model for detecting noAA in CSIRO (a), noAA in IPSL (b), AA in CSIRO (c), and noAA in IPSL (d). Black symbols are for single variables (TAS, PR, PSL). Red symbols are for variable pairs (TAS-PR, PSL-PR, TAS-PSL). Blue symbols are for single variables using variations between seasons and between land and sea. The horizontal axis shows the dimension of the state vector. The precise state vectors are described in the text. The number of Laplacians of each variable used to construct different variable combinations are indicated by marks at top x-axis: black for single variables, red for pairs of variables, blue for variables incorporate information of both seasonal variable and land-sea contrast. Dashed line indicates the threshold for detectability at the 5% significance level.

produce spatial gradients. The above result implies that the degree to which such spatial gradients *discriminate* between forcings is model dependent.

Recall that the potential detectability  $\phi^2$  is proportional to the product of the total-to-noise ratio and multicollinearity measure  $(1 - \rho^2)$ . These individual terms are shown in fig. 4.9. For the single-variable case (i.e., black symbols), the figure shows that PSL has relatively low total-to-noise ratio and high degree of collinearity compared to TAS or PR. This explains why PSL is less effective than other variables at detecting AA. PR has the least collinearity between AA and noAA responses than any other single variable.

We next consider potential detectability for pairs of variables. Specifically, for a pair of physical variables  $i$  and  $j$  on a global domain averaged over a year, the state vector is of the form

$$\mathbf{p} = \left( Y_{i,1}^{\text{ann, globe}} \quad Y_{j,1}^{\text{ann, globe}} \quad Y_{i,2}^{\text{ann, globe}} \quad Y_{j,2}^{\text{ann, globe}} \quad Y_{i,3}^{\text{ann, globe}} \quad Y_{j,3}^{\text{ann, globe}} \quad \dots \right)^T. \quad (4.2)$$

For instance, if  $i = \text{TAS}$  and  $j = \text{PR}$ , then the first two elements give the projection of these variables on the first spherical harmonic, the next two elements give the projection on the second spherical harmonic, and so on. The potential detectability for this case is shown as red symbols in fig. 4.8. Note that in this case the vector dimension does not equal the number of spherical harmonics, since the vector is formed by concatenating *pairs* of variables. Accordingly the dimension of the state vector is shown in the lower x-axis while the number of spherical harmonics is shown on the upper x-axis of fig. 4.8. The figure reveals that state vectors based on TAS-PR variable pairs give the strongest potential detectability relative to other vectors of the same dimension. Interestingly, potential detectability based on TAS-PR does not increase strongly beyond the second spherical harmonic. The strong potential detectability using TAS-PR is because the associated response vectors are less collinear than other vectors (see fig. 4.9). For example, other variables in CSIRO can have larger total-to-noise ratios but lower potential detectability due to high collinearity.

As mentioned in the introduction, many studies suggest that seasonal variations and

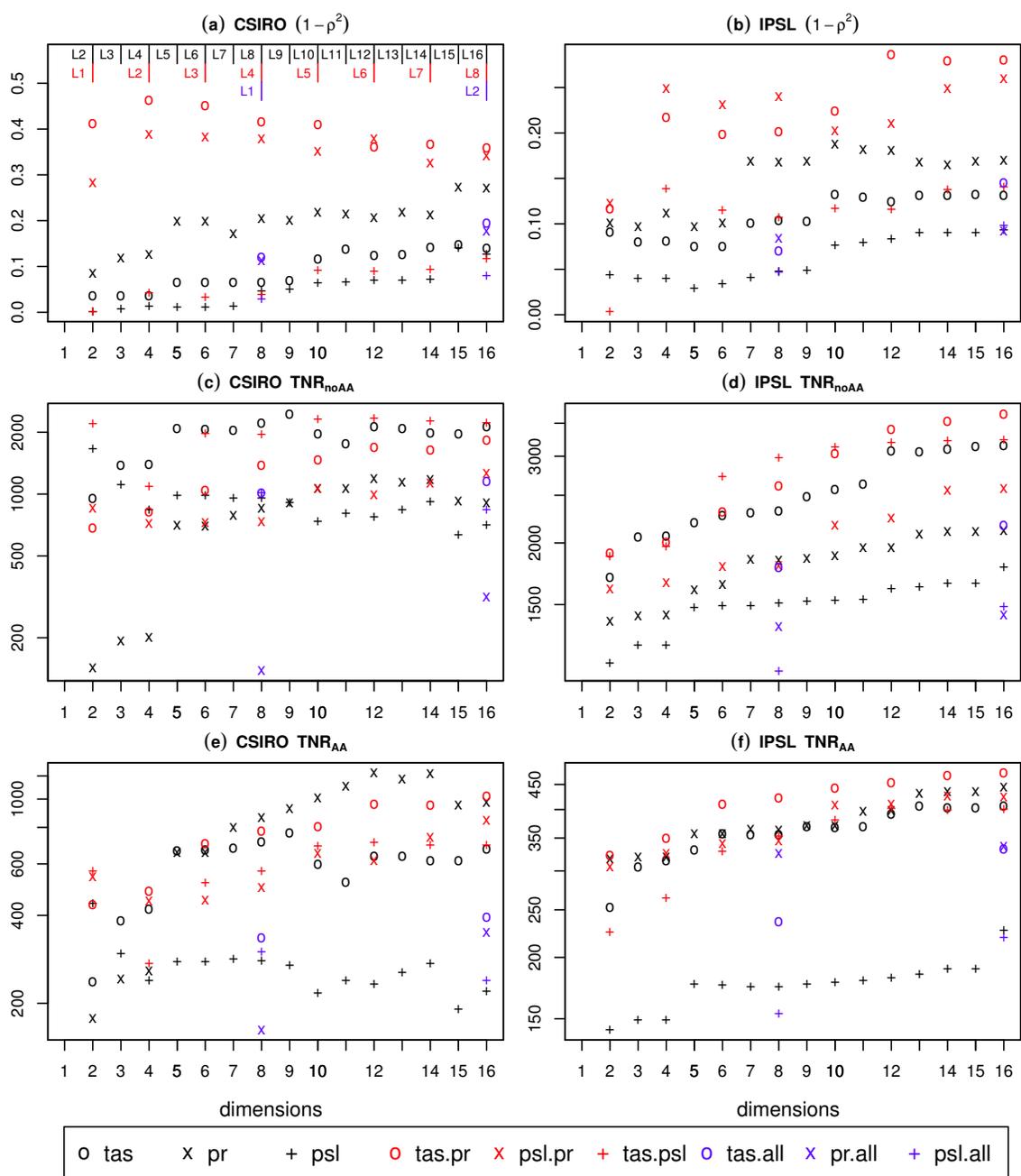


Figure 4.9: Multicollinearity measure (top row), the total-to-noise ratio for noAA (middle row), and the total-to-noise ratio of AA (bottom), for CSIRO (left column) and IPSL (right column), for the variable combinations indicated by colored symbols corresponding to fig. 4.8.

land-sea gradients are important for detection. We have examined state vectors based on this type of information, but potential detectability often is worse than just using spatial

gradients. As an example, we select state vectors that include both seasonal variations and land-sea gradients of the form

$$\mathbf{p} = \left( Y_{v,1}^{\text{djf, land}} \quad Y_{v,1}^{\text{mam, land}} \quad Y_{v,1}^{\text{jja, land}} \quad Y_{v,1}^{\text{son, land}} \quad Y_{v,1}^{\text{djf, ocean}} \quad Y_{v,1}^{\text{mam, ocean}} \quad \dots \right)^T, \quad (4.3)$$

where the superscripts indicate the obvious 3-month mean and the domain over which Laplacian eigenvectors are computed. For this vector, the first eight elements correspond to the first Laplacian eigenvector (i.e., the spatial average over the domain), and the second eight elements correspond to the second Laplacian eigenvector in the respective domain. Potential detectability based on these vectors is shown as blue symbols in fig. 4.8. For temperature and precipitation, this vector leads to less potential detectability than a vector of the same dimension based on spherical harmonics (especially for IPSL). The total-to-noise ratios for AA forcing are especially low for vectors based on (4.3) (see fig. 4.9). These results suggest that seasonal variations or land-sea contrast information does not significantly enhance potential detectability relative to global, annual-mean indices.

The above conclusions pertain to a “perfect model” analysis. The question arises as to whether our conclusions hold in an imperfect model setting. Figure 4.10 shows results similar to those in fig. 4.8, except for the “imperfect model” case in which response vectors and noise covariances are estimated from a model that differs from the model that generated the historical run. As in the perfect model case, precipitation still gives the strongest potential detectability of AA, and combining precipitation with either temperature or sea level pressure gives the strongest potential detectability of all. In contrast, jumps in potential detectability occur at the second and fifth Laplacian in IPSL rather than CSIRO, implying that the jumps are due to resolving gradient information in the response vector and the noise covariance matrix rather than the historical runs used. The use of seasonal variations or land-sea contrast information is more mixed in the imperfect model case: such information can slightly enhance potential detectability relative to vectors of the same dimension using annual mean information and spherical harmonics, but these increases are

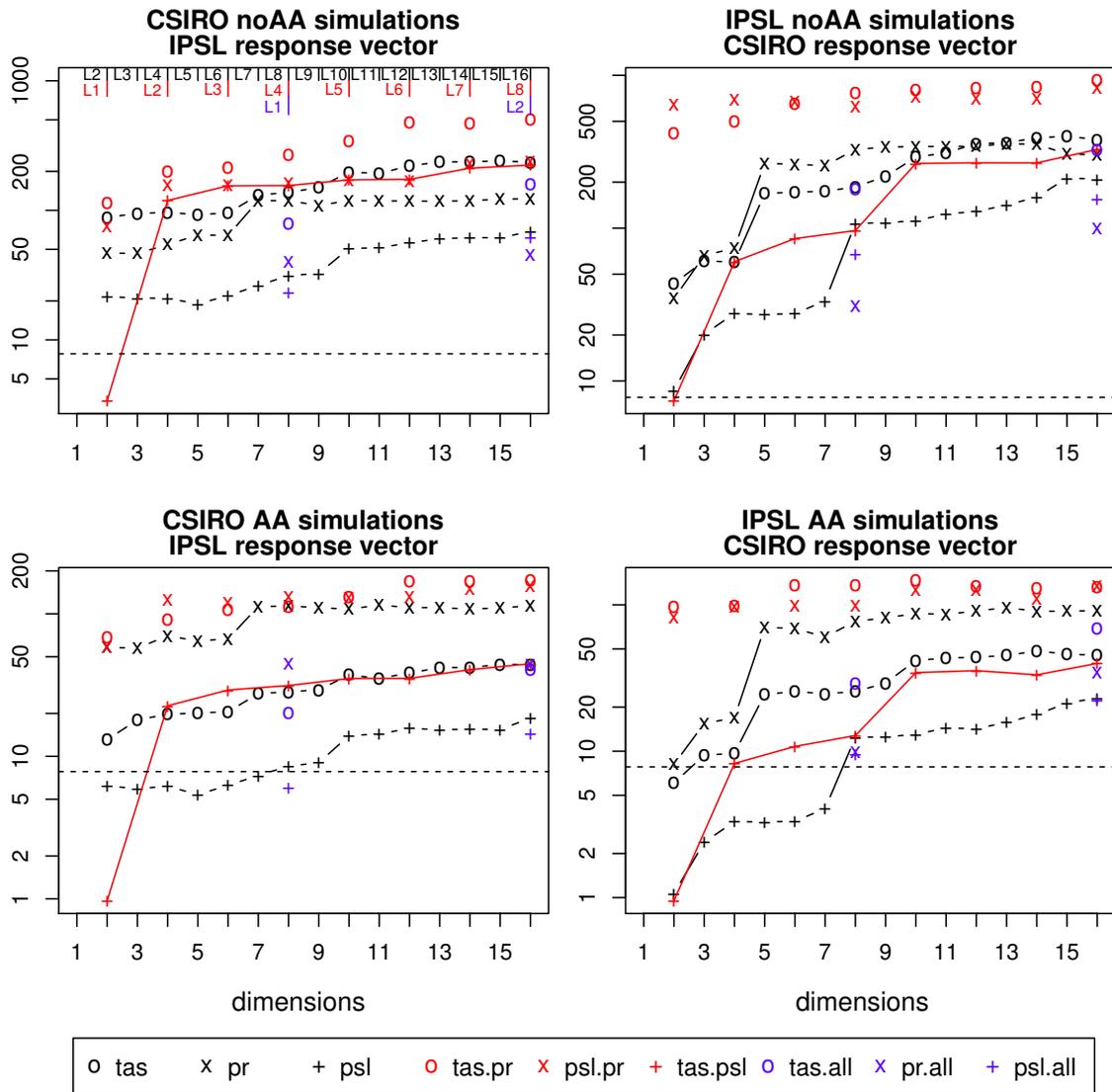


Figure 4.10: Same as fig. 4.8, but for “imperfect model” cases, in which response patterns and the noise covariance matrix are from different models, as follows: detecting noAA with response pattern from CSIRO and noise covariance matrix from IPSL (a), detecting noAA with response pattern from IPSL and noise covariance matrix from CSIRO (b), detecting AA with response pattern from CSIRO and noise covariance matrix from IPSL (c), and detecting AA with response pattern from IPSL and noise covariance matrix from CSIRO (d).

generally small, whereas the losses of potential detectability due to seasonal variations or land-sea contrast information can be substantial (especially for precipitation).

Note that TAS-PSL produces weaker potential detectability than other variable pairs.

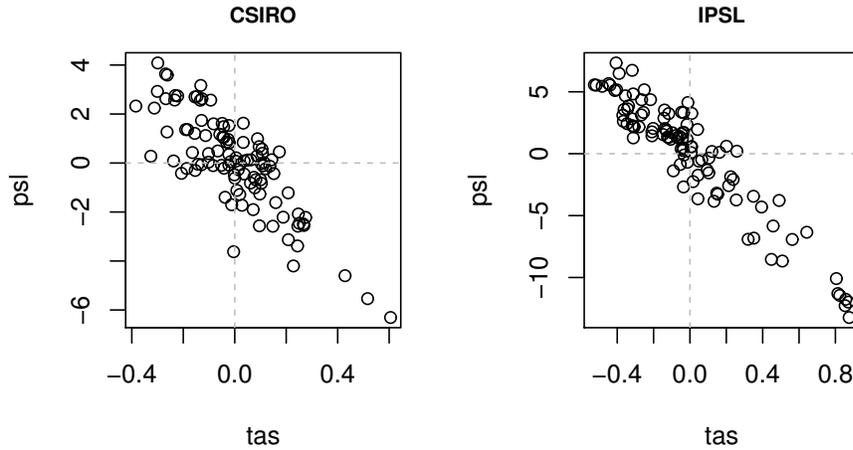


Figure 4.11: Scatter plot of global average temperature and sea-level pressure over the period 1900-2004.

PSL is atmospheric pressure extrapolated to sea level and is computed partly *from* temperature data. The question arises as to whether PSL enhances potential detectability beyond TAS. In some cases, TAS and PSL are highly correlated, as shown in fig. 4.11 for the global mean, which explains why the AA response is not detectable using only one Laplacian eigenvector of TAS and PSL. In all cases shown in fig. 4.8, TAS always has stronger potential detectability than PSL when based on the same number of Laplacian eigenvectors. In rare cases in fig. 4.10, PSL has only marginally larger potential detectability than TAS. In general, then, PSL does not enhance potential detectability significantly beyond that already available from the temperature data used to compute PSL.

The potential detectabilities shown above are computed from single-forcing runs without actually analyzing data in which both forcings are present. Realized detectability also can be computed from (3.25) using data from historical simulations, in which both forcings are present. A comparison between the potential detectability and realized detectability for the different vectors is shown in fig. 4.12. In general, potential detectability tends to overestimate the realized detectability. Nevertheless, the two measures tend to agree more strongly as the realized detectability increases beyond the significance threshold.

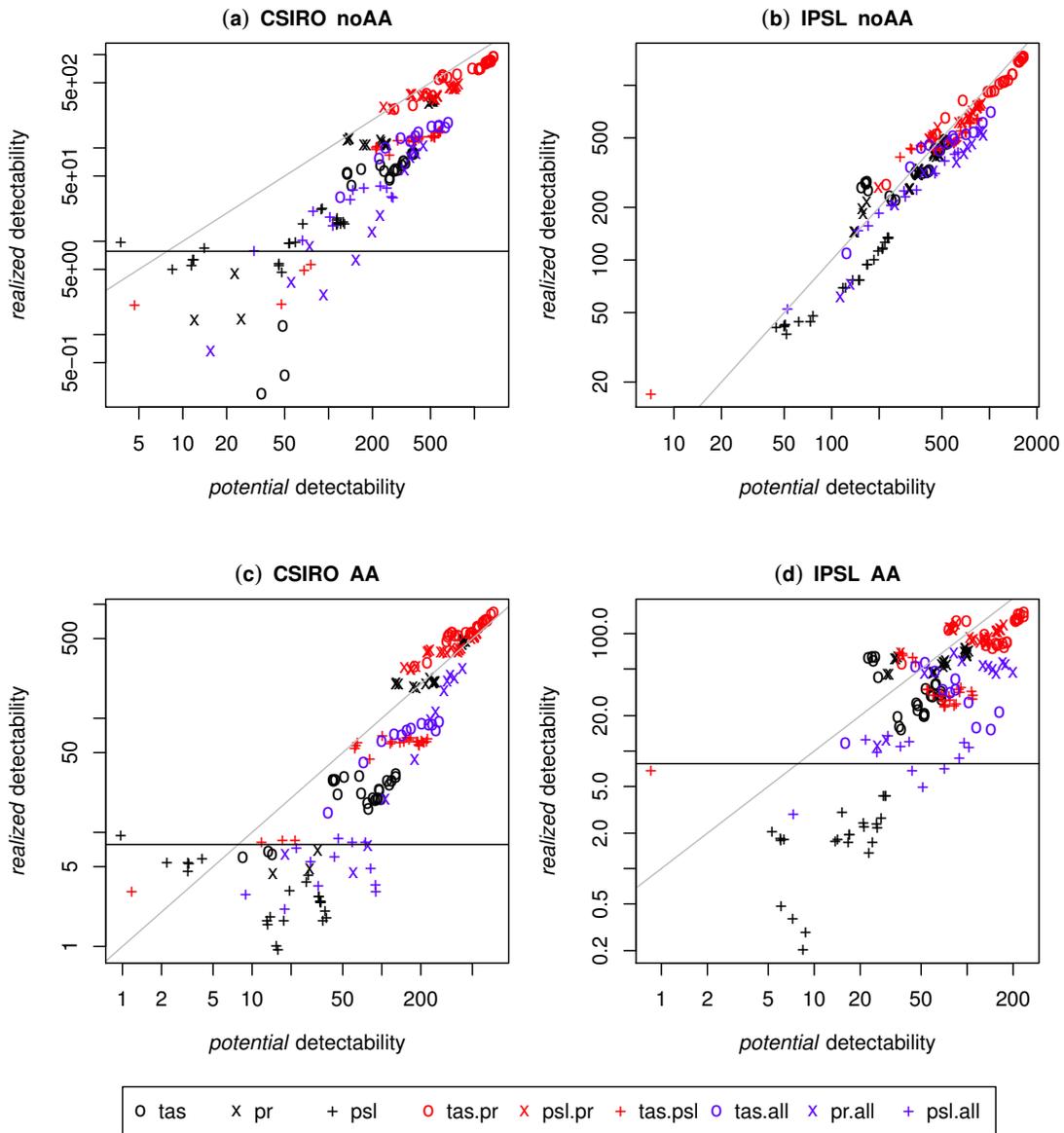


Figure 4.12: Comparison of potential detectability and realized detectability, for noAA forcing (top row) and AA forcing (bottom row), and for CSIRO (left column) and IPSL (right column). potential detectability is computed using  $TNR$  ratio estimated from single-forcing run, and realized detectability is computed using  $TNR$  ratio estimated from historical run. The colored symbols correspond to the same cases as in fig. 4.8. Grey lines have slope of 1 and cross the origin. Black horizontal lines indicate the threshold of 5% significance level of realized detectability.

### 4.3 Summary and Discussion

This study showed that joint temperature-precipitation information over a global domain provides more accurate estimates of aerosol forced responses in climate models, compared to using temperature, precipitation, or sea level pressure individually or in combination. To clarify the importance of spatial versus temporal information, we use only spatial structure or co-variability between variables as input to fingerprinting and infer the associated time series from fingerprinting (which subsumes the traditional scaling factor). The response vector was chosen to maximize signal-to-noise ratio using the method of Jia and DelSole (2012). Typically, only one vector is detectable for any given forcing and model. To do this, the data were represented in terms of a few Laplacian eigenvectors instead of EOFs. Laplacian eigenvectors enhance physical interpretation (since they often are of the form of monopoles, dipoles, tripoles, etc) and are data independent and hence can be compared across models. Also, the time evolution is assumed to be low-frequency, in the sense that it is representable by a low order polynomial in time. Since no temporal information is included in the response vector, the results clarify how detectability depends on spatial structure and co-variability with other variables.

To illustrate the above methodology, we used it to estimate the change in global average temperature in historical simulations due to different forcings using joint temperature-precipitation response vectors. The inferred time series closely matched the time series computed from single-forcing runs, even if only the first spherical harmonic was used, and even if the response vector of one model is used to infer the change in a different model. In some models, the accuracy of the estimates improves as more spherical harmonics are included in the response vector. A third order polynomial in time was found to be sufficient for capturing most of the response over a century.

The above methodology was applied to identify the most effective variables in a model for estimating the cooling attributable to anthropogenic aerosols. We explored three different variables— surface temperature, precipitation, sea-level pressure— and various combinations of spatio-temporal information. We found that joint temperature-precipitation response

vectors yielded the strongest detection of aerosol cooling, even in the absence of spatial and temporal gradient information. The best single variable for detecting anthropogenic aerosols is annual mean precipitation. We call attention to the fact that sea-level pressure is computed partly from temperature data and hence is not independent of temperature. We find no evidence that sea-level pressure provides any significant gain in detectability beyond that already available in the surface temperature data from which sea-level pressure is derived. We also used state vectors that included seasonal variations and that separated land and ocean domains, but these vectors generally produced less detectability than vectors of the same dimension based on annual mean variables of Laplacian eigenvectors. These conclusions hold even in “imperfect model” setting, in which the response pattern and noise covariance matrix are from different models. In one model, jumps in detectability occurred after the second and fifth spherical harmonics were included, which measure, respectively, the North-South hemispheric gradient and equator-to-pole gradient. Further comparison between “perfect” and “imperfect” model analyses indicate that these jumps are due to resolving gradient information in the response vectors and the noise covariance matrix.

As emphasized throughout this study, our results are derived from model experiments in which errors in the forced response and observations can be neglected. There is no guarantee that response vectors that work well in models also will work well in observations. In fact, DelSole et al. (2015) applied optimal fingerprinting to observations using global-mean joint temperature and precipitation information, but observed, global, long-term precipitation was inconsistent with models and even between data sets, resulting to inconsistent hydrological sensitivity between models and observations (fig. 4.13). Thus, detection analysis based on observed global-mean joint temperature and precipitation turns out to be difficult, despite the fact that the global-mean joint temperature and precipitation is an efficient detection variable combination in a model setting. Nevertheless, the methodology illustrated here is expected to provide a useful starting point for deciding which variables to use for detection and attribution analysis of observations.

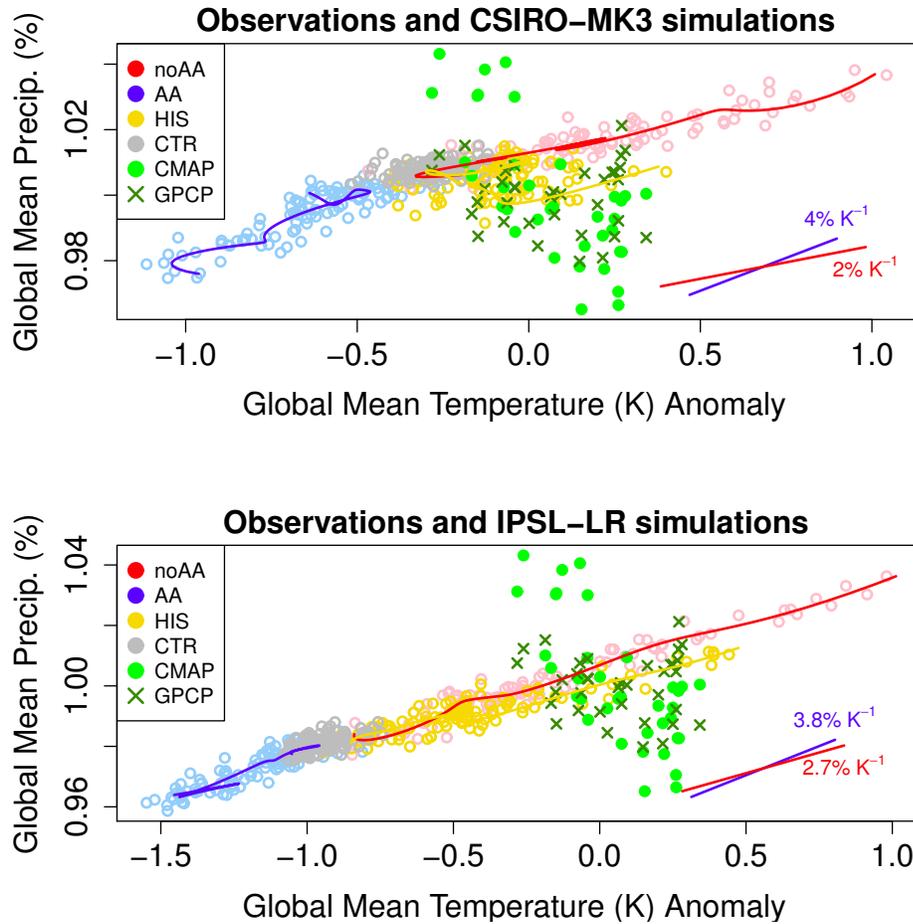


Figure 4.13: Scatter plot of annual-mean, global-mean temperature and precipitation in simulations from CSIRO (top) and IPSL (bottom), and observations (green symbols). Each panel displays simulations with anthropogenic aerosols only (“AA, blue); with all forcings except anthropogenic aerosols (“noAA, red); with both natural and anthropogenic forcings “his, gold); with no natural and anthropogenic forcing (“ctr, grey). Also shown are observations based on GPCP (green filled circles) and CMAP (green crosses). Temperature in the historical run and observations are centered to have zero mean over the common period 1979-2004, while other simulations are shown relative to the mean of the historical run over this period. Precipitation in the historical run and observations are normalized by their respective means over the period 1979-2004, while other simulations are normalized by the mean of the historical run over the same period. The solid curves show the projection of a low-pass time series of the respective runs. The two lines in the bottom right corner indicate the hydrological sensitivities estimated from the AA and noAA simulations. From fig.3 of DelSole et al. (2015).

## Chapter 5: Inferring Anthropogenic Aerosol Cooling From Land Observations

The last chapter showed that the *joint* variability of surface air temperature and precipitation is sufficiently distinct from that of other forcings that it can be used to estimate aerosol cooling in a “perfect model world.” Unfortunately, global average precipitation is inconsistent between models and observations, and even between different observational data sets (Adler et al. 2011; Wentz et al. 2007; Liepert and Previdi 2009; DelSole et al. 2015). One possible source of this inconsistency is observational errors, since estimates of global mean precipitation differ considerably between data sets, and precipitation estimates over the ocean are derived largely from satellite measurements, which lack ground truth for calibration (Adler et al. 2011; Wentz et al. 2007; Liepert and Previdi 2009; DelSole et al. 2015). Over land, however, long records of reliable gauge observations are available in addition to satellite measurements. Therefore, we consider variables only over land, where rain gauge data are available for constraining satellite estimates of precipitation. More precisely, this chapter investigates whether combining precipitation and temperature data *over land* can improve estimates of anthropogenic aerosol cooling.

### 5.1 Data

We use the same model data as described in sec. 2.4. But by considering the length of observation data, the periods are of 1901-2004 and 1948-2004 separately. Note that model data ends in 2004 for historical simulations. Two sets of observation datasets are used in this study: the CRU TS 3.22 monthly gridded fields (Harris et al. 2014) and the GHCN-CAMS monthly gridded fields (Fan and van den Dool 2008). Each dataset contains observational fields of both surface air temperature (TAS) and precipitation (PR). The CRU data spans

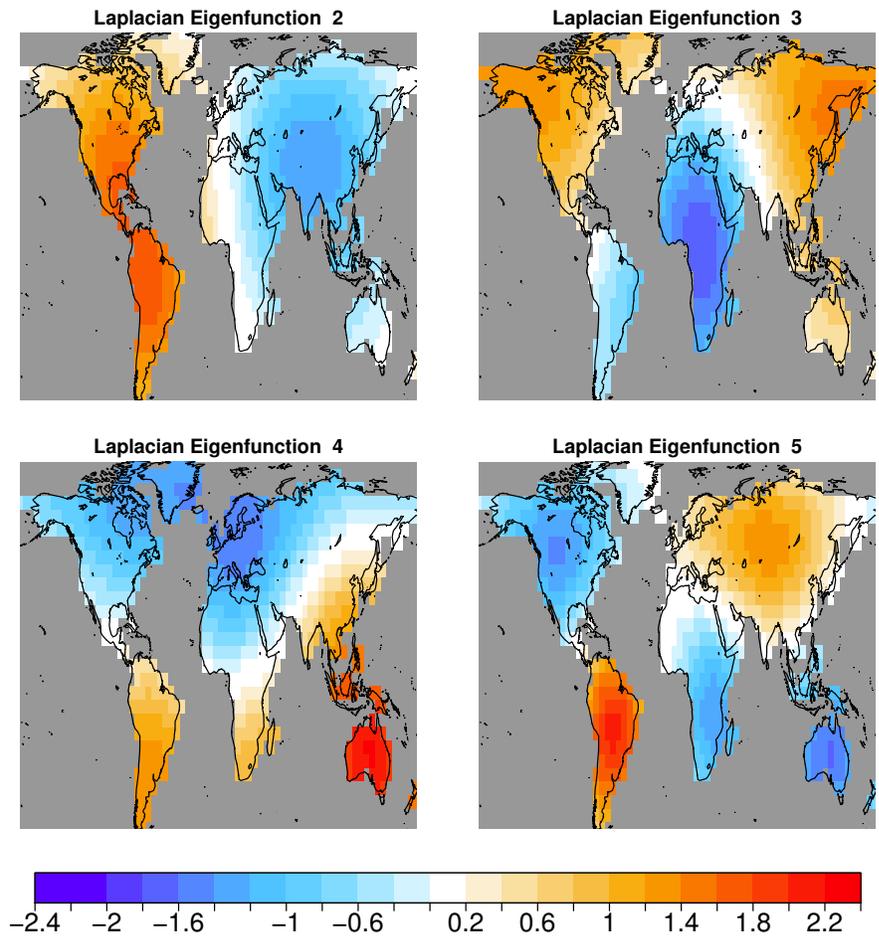


Figure 5.1: Eigenvectors 2, 3, 4, 5 of the Laplace operator over global land areas.

from 1901 to 2013 and the GHCN-CAMS data spans from 1948 to 2013. In analyses of the longer period (1901-2013), only CRU data is used. In analyses of the shorter period (1948-2013), both datasets are used. Observations of sea level pressure (PSL) are not analyzed in this study because this variable has less potential detectability than TAS and PR, as will be shown in the next section. All model simulations and observations are interpolated onto a common  $5^\circ \times 5^\circ$  grid. A common mask is constructed based on the two observational datasets and applied to all model simulations and observations.

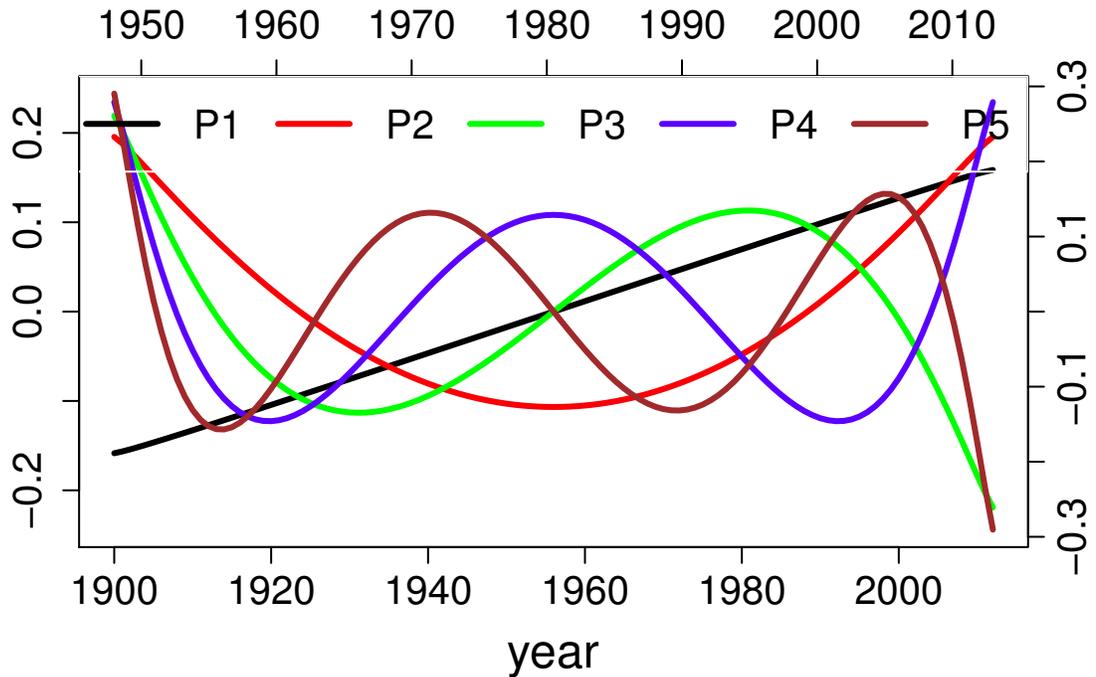


Figure 5.2: Temporal evolution of Legendre polynomials 1-5 for periods of 1901-2013 (left and bottom axes) and 1948-2013 (right and upper axes).

## 5.2 Potential Detectability Over Land

Even if certain variables were found to improve detection of anthropogenic forcing on a global domain, there is no guarantee that the same variables, restricted to land, would improve detection. In this section, we investigate whether temperature and precipitation over land areas allows detection of anthropogenic aerosols in a perfect model world. To answer this question, we compare the potential detectability of different variables, variable-combinations, and spatial filters. We also included PSL in the mix. These results will be used to rank variable combinations by their ability to detect a specific forcing, and demonstrate that certain fingerprints should be discarded because they would fail to detect a forcing even in a perfect model.

We are interested in whether forced responses can be detected using only spatial or co-variability information over land (i.e., no temporal information). For this purpose, we

use Laplacian eigenvectors over land, displayed in fig. 5.1. Also, the Legendre polynomials used in this section for observations is displayed in fig. 5.2. We also consider the time period 1948-2013, for which the Legendre polynomials have the same structure as in fig. 5.2 but with an x-axis ranging between 1948 and 2013 (see top axis in fig. 5.2). The Legendre polynomials used in this section for model data (not shown) have the same structure as those in fig. 5.2 but with x-axis ranging between 1901 and 2004, and 1948 and 2004 separately corresponding to the longer and shorter observations.

Let  $Y_{v,l}^{\Delta T}$  denote the  $v$ 'th physical variable– TAS, PR or PSL– for the  $l$ 'th Laplacian eigenfunction for time averaging period  $\Delta T$ . For an annual average ( $\Delta T = \text{“ann”}$ ) single variable  $v$  on a global land domain, the state vector is of the form

$$\mathbf{p} = \left( Y_{v,1}^{\text{ann}} \quad Y_{v,2}^{\text{ann}} \quad Y_{v,3}^{\text{ann}} \quad \dots \right)^T, \quad (5.1)$$

where vector dimension  $d$  equals the number of Laplacian eigenvectors. For a pair of physical variables  $i$  and  $j$  on a global land domain averaged over a year, the state vector is of the form

$$\mathbf{p} = \left( Y_{i,1}^{\text{ann}} \quad Y_{j,1}^{\text{ann}} \quad Y_{i,2}^{\text{ann}} \quad Y_{j,2}^{\text{ann}} \quad Y_{i,3}^{\text{ann}} \quad Y_{j,3}^{\text{ann}} \quad \dots \right)^T, \quad (5.2)$$

where vector dimension  $d$ , for even dimension, corresponds to  $d/2$  Laplacian eigenvectors for variable  $i$  and  $d/2$  Laplacian eigenvectors for variable  $j$ . For a variable based on seasonal variations, the state vector is of the form

$$\mathbf{p} = \left( Y_{v,1}^{\text{djf}} \quad Y_{v,1}^{\text{mam}} \quad Y_{v,1}^{\text{jja}} \quad Y_{v,1}^{\text{son}} \quad Y_{v,2}^{\text{djf}} \quad Y_{v,2}^{\text{mam}} \quad \dots \right)^T, \quad (5.3)$$

where vector dimension  $d$  corresponds to  $d/4$  Laplacian eigenvectors for variable  $v$ , assuming  $d$  is a multiple of 4, and  $\Delta T$  corresponds to the obvious 3-month average.

The potential detectability of different variables and variable-combinations are shown in fig. 5.2. Different colors indicated different categories of variables: black for single

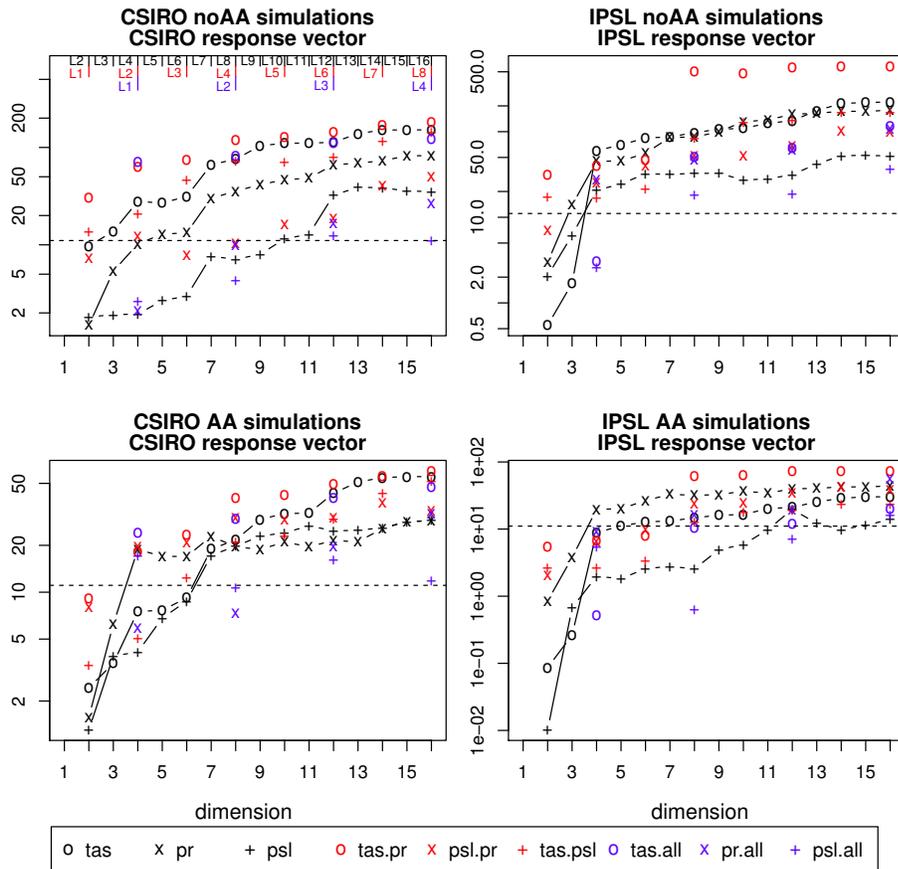


Figure 5.3: Potential detectability of different response vectors for noAA in CSIRO (a), noAA in IPSL (b), AA in CSIRO (c), and noAA in IPSL (d). Black symbols are for single variables (TAS, PR, PSL); red symbols are for variable pairs (TAS-PR, PSL-PR, TAS-PSL); blue symbols are for single variables using seasonal variations. The precise state vectors are described in the text. The dimension of each state vector is indicated in the bottom x-axis and the number of Laplacian eigenvectors for each vector is indicated in the top x-axis according to the above color scheme. Dashed line indicates the threshold for detectability at the 5% significance level.

annual-mean variables, red for pairs of annual-mean variables, and blue for single seasonal-mean variables. Vector dimension is indicated on the bottom x-axis and the number of Laplacian eigenvectors is indicated on the upper x-axis according to the color scheme. The figure shows that, for sufficiently large dimension, the joint temperature-precipitation vector based on annual averages gives the largest potential detectability in both models and for both forcings (e.g., the red circles lie above all other symbols at large dimension).

However, in the case of CSIRO, the land-only vectors for TAS are comparable to those for the joint TAS-PR vector; that is, combining temperature and precipitation does not enhance detectability very much compared to the temperature-only vector. This result is in contrast to the global-average vectors analyzed in the last chapter, and in contrast to IPSL. Inclusion of seasonal variations of TAS can improve potential detectability in CSIRO, but not dramatically, and not in IPSL. PSL always shows the lowest potential detectability in both models. Note, however, that while the joint TAS-PR vector and the single-variable TAS vector have comparable potential detectability for equal dimension, the former uses fewer Laplacian eigenvectors and therefore filters out more small scale variability.

We also have computed potential detectability in an “imperfect model world” in which response patterns and noise covariance matrix derived from one model are used to detect forced responses in another model’s single-forcing run (not shown). The above conclusions hold in this case too. We conclude that in a “perfect model world” the joint temperature-precipitation data over land can detect AA and noAA forced responses. However, temperature data also can detect AA and noAA forced responses given enough Laplacian eigenvectors (i.e., given enough small scale structure).

### 5.3 Signal-to-Noise Ratio and Multicollinearity

As indicated in (3.25), potential detectability depends on two factors: the strength of the forced response relative the noise (i.e.,  $TNR_m$ ) and the degree of collinearity between the response vectors (i.e.,  $\rho$ ). The value of these quantities for the joint temperature-precipitation vector based on annual means is shown in fig. 5.4. Recall that the vector is constructed according to (4.2), so if the dimension is even then the number of Laplacian eigenvectors equals half the vector dimension.

The signal-to-noise ratio increases with the number Laplacian eigenvectors, as expected: the vectors are chosen to maximize the signal-to-noise ratio at each specified dimension. The pattern correlations decrease with the number of Laplacian eigenvectors, but in quite different ways in the two models: a sharp decrease occurs in IPSL from the 3rd to 4th Laplacian

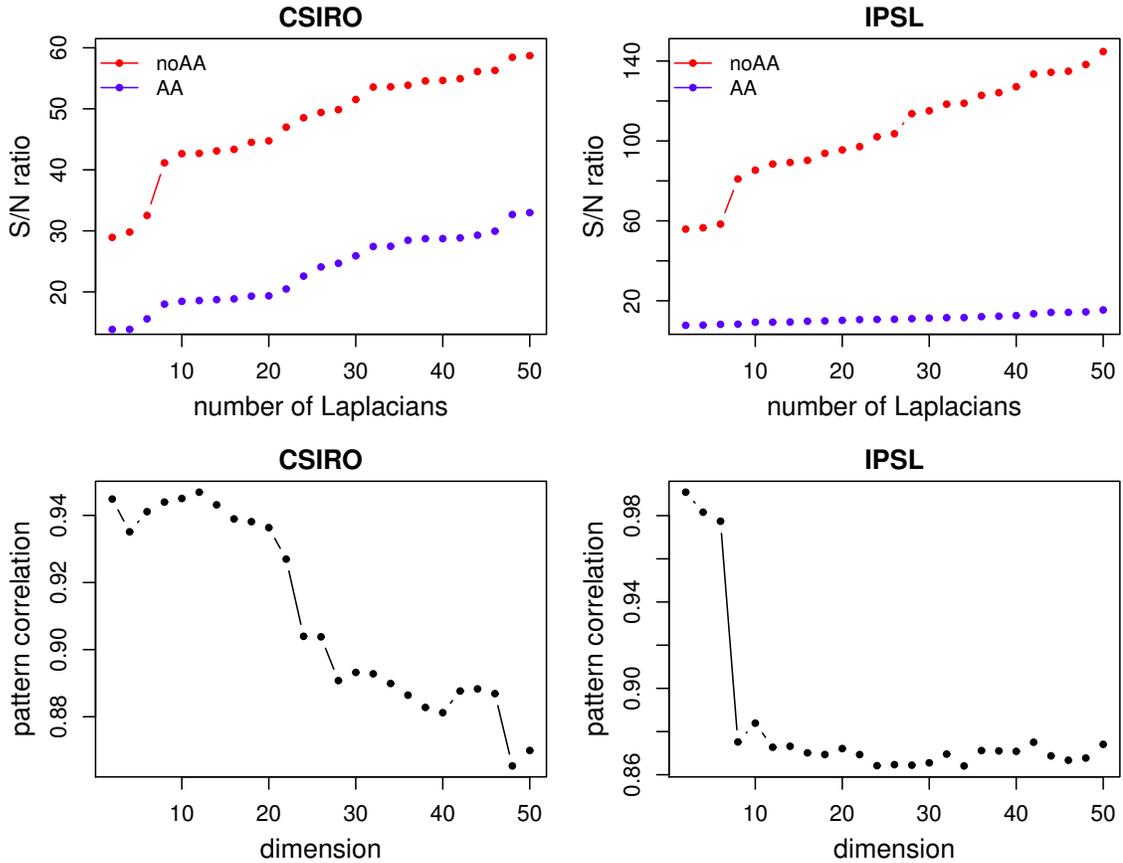


Figure 5.4: The signal-to-noise ratio of the noAA (red) and AA (blue) response, and the pattern correlation between the response vectors, in CSIRO (left figures) and IPSL (right figures), as a function of the dimension of Laplacians of the joint annual-mean temperature-precipitation vector.

eigenvector, while the decrease in CSIRO is more gradual. The abrupt decrease implies that the 4th Laplacian eigenvector contains structure (i.e., the hemispheric gradient, as can be seen from fig. 5.1) that helps to distinguish between AA and noAA responses in IPSL. A variety of factors could be responsible for this difference in model response patterns. For instance, aerosol-cloud interactions (Ma et al. 2015) and precipitation simulations (Wehner et al. 2014; Li et al. 2015) are known to be sensitive to a model’s horizontal resolution and emissions. CSIRO has stronger black carbon and organic aerosol emissions (Rotstayn et al. 2012) and higher longitudinal resolution (Dufresne et al. 2013) than IPSL. In any case, this abrupt change in pattern correlation will be seen to affect the detection and attribution

results shown later.

## 5.4 Spatial Patterns

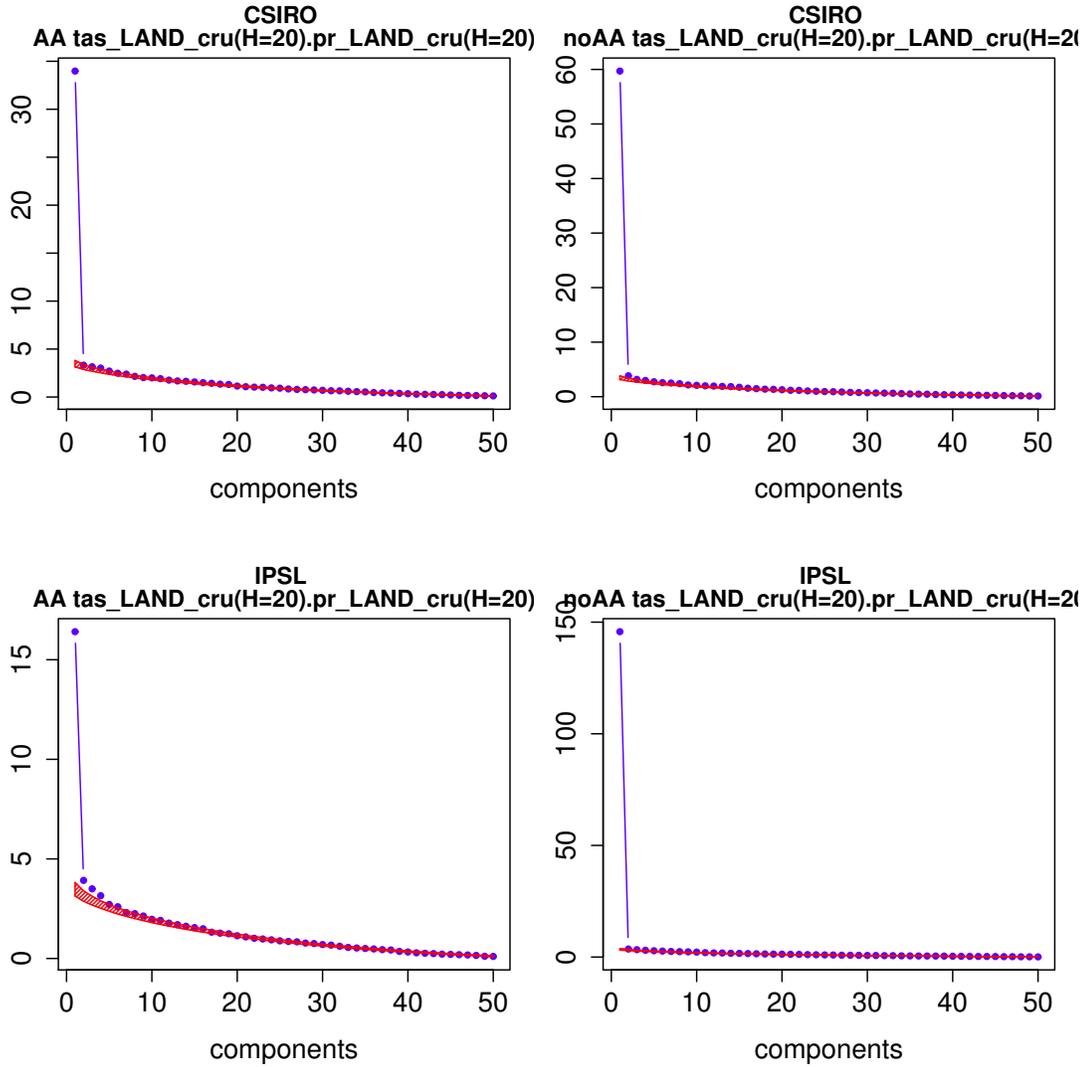


Figure 5.5: Maximized signal-to-noise variance ratios (blue line with dots), and the 95% confidence interval under the hypothesis of equal variances (red shading), as a function of the dimensions of fingerprints, based on 25 Laplacian eigenvectors and annual mean fields of each variable, for AA single-forcing runs of CSIRO (top left) and IPSL (bottom left), and noAA single-forcing runs of CSIRO (top right) and IPSL (bottom right).

We find that twenty-five Laplacian eigenfunctions produce stable results for observations.

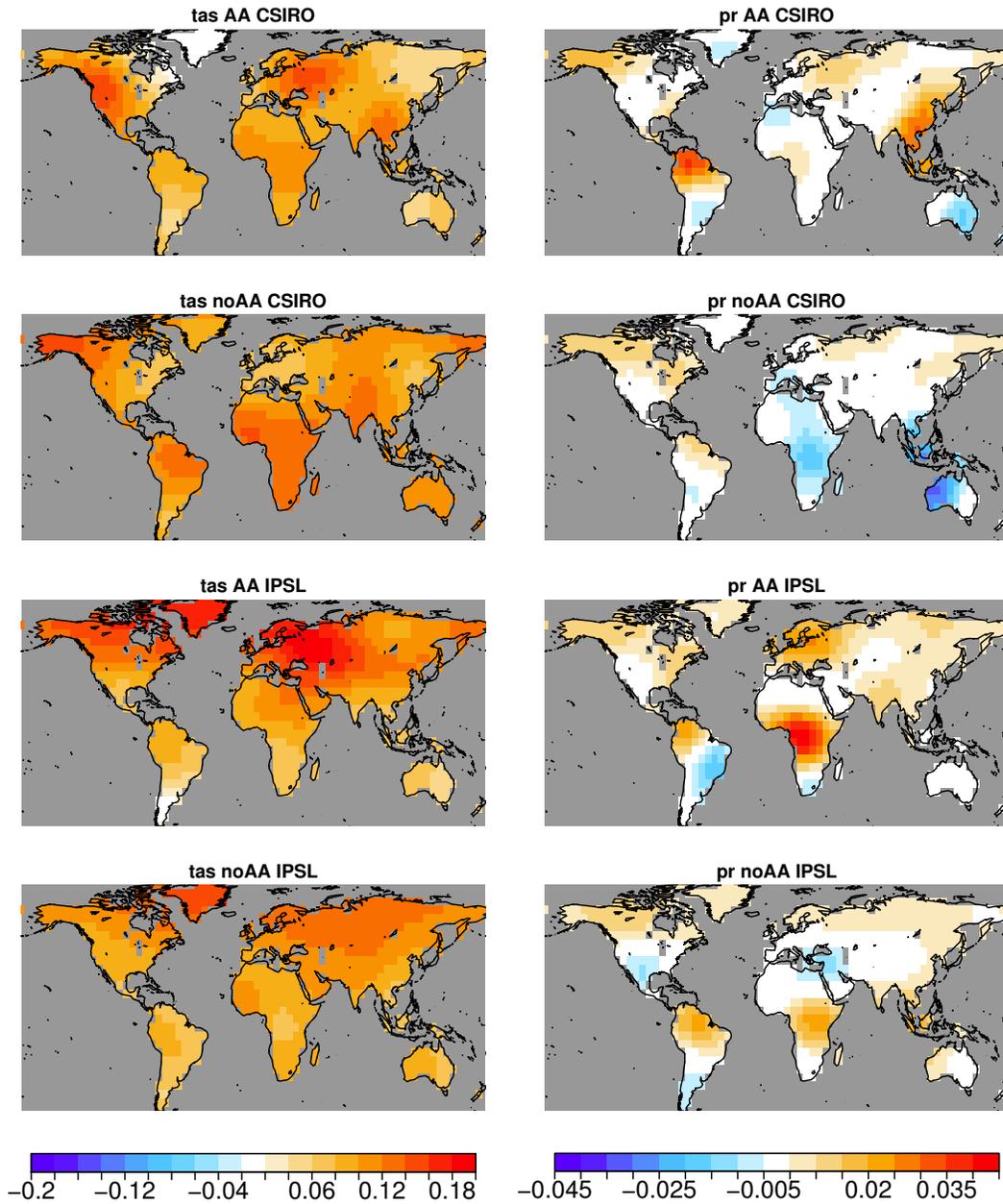


Figure 5.6: Response vectors that maximize detectability of AA forcing and noAA forcing in two different models (CSIRO and IPSL) based on 25 Laplacian eigenvectors and annual mean fields. The response vector depends on a joint temperature-precipitation relation and therefore is characterized by a pair of patterns, shown in each row. The left column shows temperature fields and the right column shows precipitation fields. The corresponding forcing and model is indicated in the title of each figure. These response patterns were used to estimate the time series shown in fig. 5.10.

Given this many Laplacian eigenfunctions, the most detectable annual-mean temperature patterns for AA and noAA responses are quite similar to the temperature patterns in the most detectable joint temperature-precipitation pattern. Note that for each single-forcing simulations, only the most detectable component are statistically significant fig. 5.5. Therefore we show just the most detectable joint temperature-precipitation patterns in fig. 5.6. Because these are joint patterns, each vector contains both a temperature and precipitation pattern, which are shown along a row. When the response patterns of surface air temperature and their respective time series (thin black and green curves in fig. 5.10) are multiplied, both models indicate cooling due to AA and warming due to noAA over most global land areas, as expected. For precipitation, models show diversities in regions of equatorial Africa, the maritime continent and Australia. In most global land regions, the precipitation patterns are still fairly consistent across models: decreasing precipitation due to AA and increasing precipitation due to noAA. Note that the direction of response depends on the product of response pattern and time series, and that the time series for AA is mostly negative while that of noAA is mostly positive.

## 5.5 Confidence Intervals for the Polynomial Coefficients

### 5.5.1 Joint Temperature-Precipitation Fingerprints

Confidence intervals for the coefficients of Legendre polynomials 1-5 are shown in fig. 5.7 for the joint annual-mean temperature-precipitation vector using 1 and 25 Laplacian eigenfunctions. For each coefficient, warm colors show confidence intervals for noAA and cold colors show those of AA. Each confidence interval is shown separately for CRU observations and for the first ensemble member of historical runs and single-forcing runs. Confidence intervals that exclude zero indicate *detection* of a given forced response. Confidence intervals excluding zero and overlapping with those of individual forced responses indicate *attribution* of the change to the forced response.

Not surprisingly, the response for both AA and noAA forcing projects most strongly on

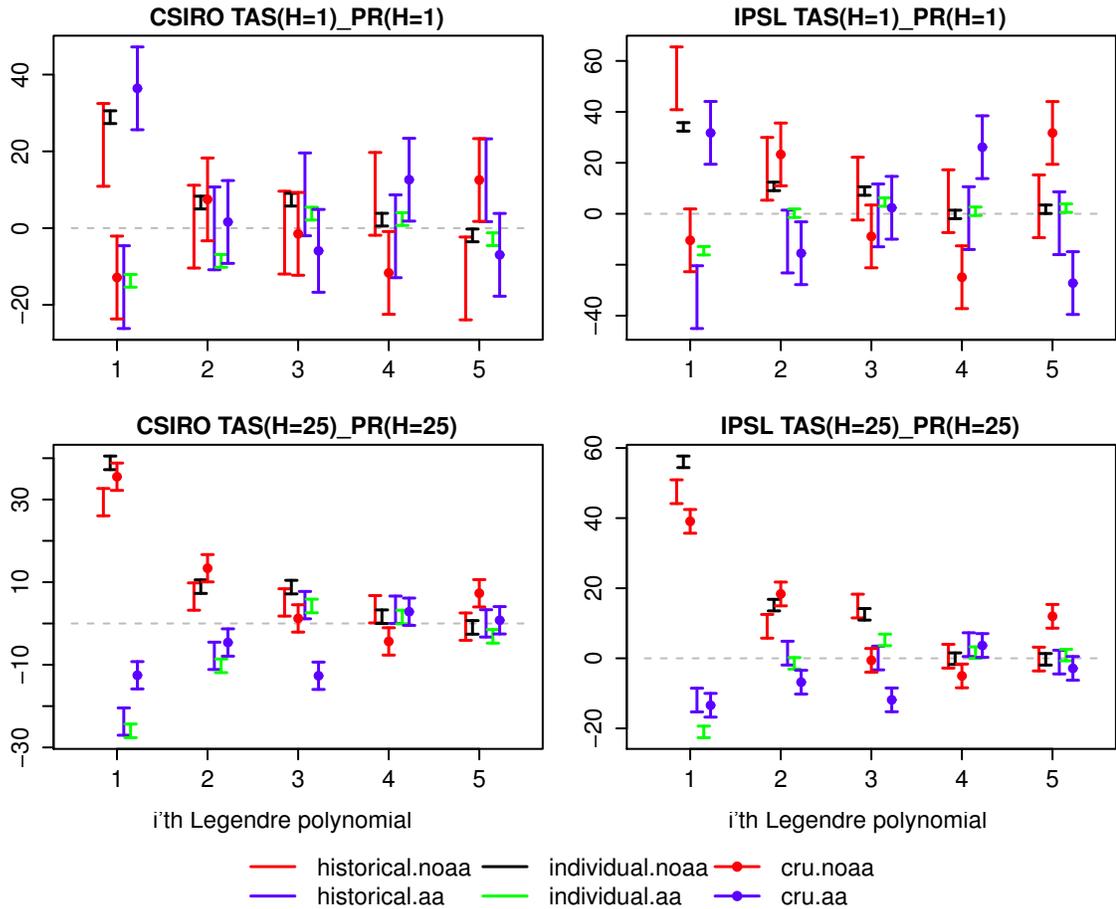


Figure 5.7: Confidence intervals of estimated coefficients of Legendre polynomials 1 to 5 based on the first (top panels) and the first twenty-five (bottom panels) Laplacian eigenfunctions, separately for noAA (warm colors) and AA (cold colors) of single-forcing runs, historical simulations and CRU observations. Estimates are made based on the joint fields of surface air temperature and precipitation.

Legendre polynomial 1, corresponding to a linear function of time (i.e., the “trend”). The confidence intervals shrink as smaller spatial scales (i.e., more Laplacian eigenvectors) are taken into account. For IPSL, the intervals shrink abruptly across 4 Laplacian eigenvectors (not shown) because of the abrupt change in multicollinearity (see fig. 5.4). The results using only 1 Laplacian eigenfunction (top row) are equivalent to fingerprinting using only land-average quantities. In a perfect model world, both forcings can be detected using only land-average quantities. However, the sign of the trend coefficients for observations are opposite to that derived from models for both noAA and AA response patterns. This

result shows that the land-average temperature-precipitation co-variability differs significantly between models and observations, consistent with the results of DelSole et al. (2015) for global-average variables.

Detection of both forcings in models and observations occurs for twenty-five Laplacian eigenvectors (bottom row of fig. 5.7). Most coefficients for models and observations are consistent, in the sense that the confidence intervals for the same forcing overlap. The main exception is a tendency for fingerprinting to underestimate the forced response compared to single-forcing runs. This bias is presumably related to random errors in the forced response patterns, which have been neglected in our fingerprinting method, and which tend to bias the amplitudes toward zero. This bias is an artifact of least squares that can be avoided by an Error-in-Variables (EIV) model (Allen and Stott 2003). Another interesting exception is that observations reveal a detectable AA response in the 3rd polynomial and a detectable noAA response in the 5th polynomial, using either model response vector, but these responses are not consistent with models. Legendre polynomials 1-2 can fit an exponential whereas polynomials 3-5 might be identified with multi-decadal variability. Thus, this discrepancy may imply that models are missing multi-decadal responses over land.

### 5.5.2 Temperature Fingerprints

Most detection and attribution studies use temperature data only. For comparison, then, we show in fig. 5.8 confidence intervals for the coefficients of Legendre polynomials using only temperature data. The top row shows the case using two Laplacian eigenvectors (the case of one Laplacian is not shown because it is impossible to separate two responses using one spatial pattern). For CSIRO, no detection is possible in a perfect model world. Despite this, the forced response is detectable in observations, but in a direction contrary to expectations (i.e., the aerosols warm the land while greenhouse gases cool the land). To understand this inconsistency, we consider scatter plots for the first two Laplacian eigenvectors over land for observations and model simulations, shown in fig. 5.9. In the case of CSIRO, individual

simulations suggest a negative correlation between the two Laplacians, whereas observations suggest a positive correlation. Thus, in order for fingerprinting to fit observations, the sign of the forced responses need to be flipped. In the case of IPSL, the slope of the relation is nearly identical for the two forcings, suggesting collinearity problems. Indeed, the correlation parameter  $\rho$  is 0.9981 and 0.9999 for CSIRO and IPSL, respectively. Recall that our fingerprint method ignores uncertainty in the response vectors. This assumption is problematic when the vectors are nearly collinear. It is likely, then, that the confidence intervals are significantly underestimated, which probably explains why the trend in the forced response for IPSL is not estimated accurately even in a perfect model world (i.e., the confidence intervals for the single-forcing runs do not overlap with those of the historical runs). Given this inconsistency in a perfect model world for IPSL, fingerprinting is not expected to work in observations. These results reveal a stark contrast between models, and between models and observations.

The results using 25 Laplacian eigenvectors of annual-mean temperature are shown in the bottom row of fig. 5.7. In general, most coefficients for models and observations are consistent (i.e., the confidence intervals for models and observations overlap), with some exceptions due to the tendency for the response to be underestimated compared to the single-forcing runs, as found earlier. Compared to the joint TAS-PR case, fewer coefficients are detectable using TAS-only, owing to the larger confidence intervals. Nevertheless, the overall consistency between the joint TAS-PR case and the TAS-only case suggests that much of the signal is contained in the temperature data. As found in the TAS-PR case, there is a detectable AA response in the 3rd polynomial and detectable noAA response in the 5th polynomial in observations, but not in models. This result again implies that models are missing a multi-decadal response over land.

## 5.6 Time Series

It proves instructive to translate some of the above results into the time domain. To do this, we multiply the estimated coefficients by the corresponding Legendre polynomial. The time

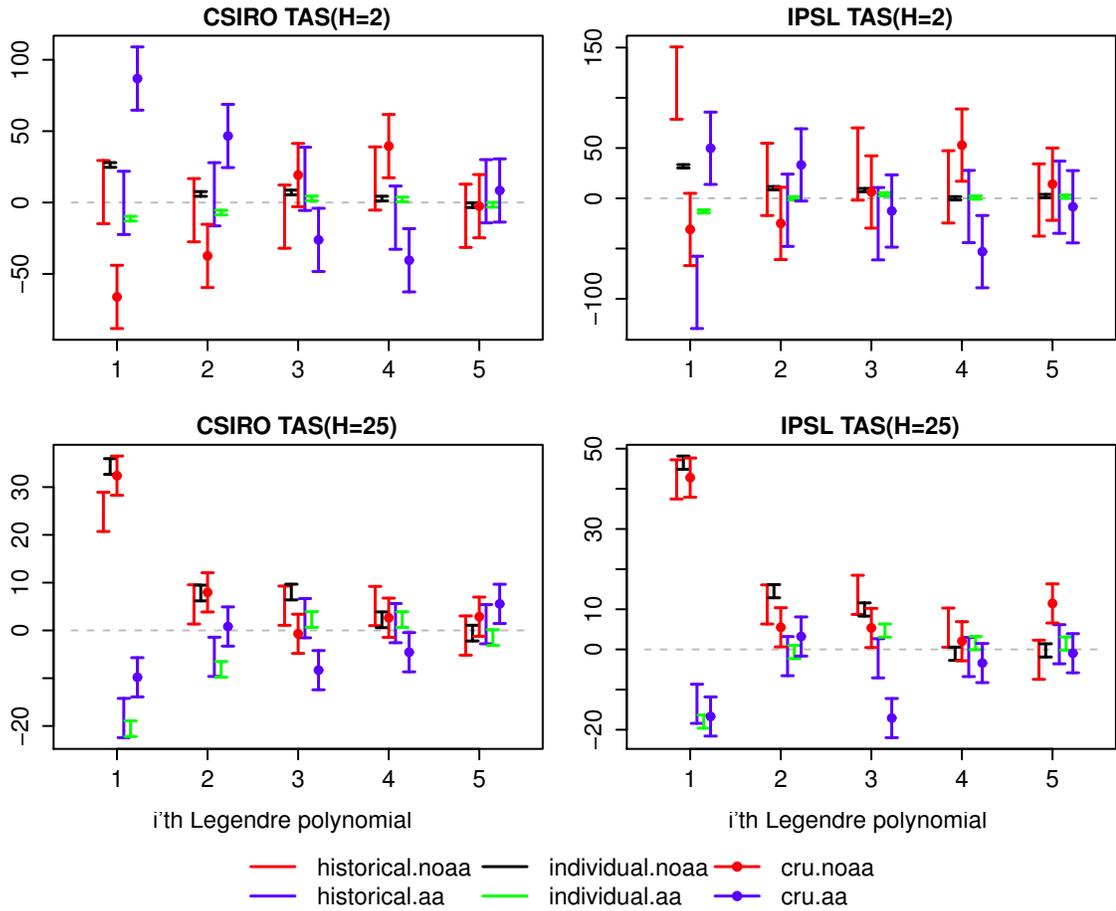


Figure 5.8: Similar to fig. 5.7 but estimates are made based on only surface air temperature.

series for the AA and noAA response based on representing TAS-PR and TAS by the first twenty-five Laplacian eigenfunctions over land are shown in fig. 5.10. Curves for different forcings have been offset relative to each other for clarity. Different curves of the same color show results for different ensemble members of the historical run. Time series estimated separately from the AA and noAA simulations (without temporal smoothing, using (A.25)) are shown in thick green and black, and time series estimated from the CRU observational dataset are shown in the thick red and blue.

Time series estimated from historical runs (thin red and blue) look essentially like smoothed versions of the time series obtained from the AA and noAA runs (think green and

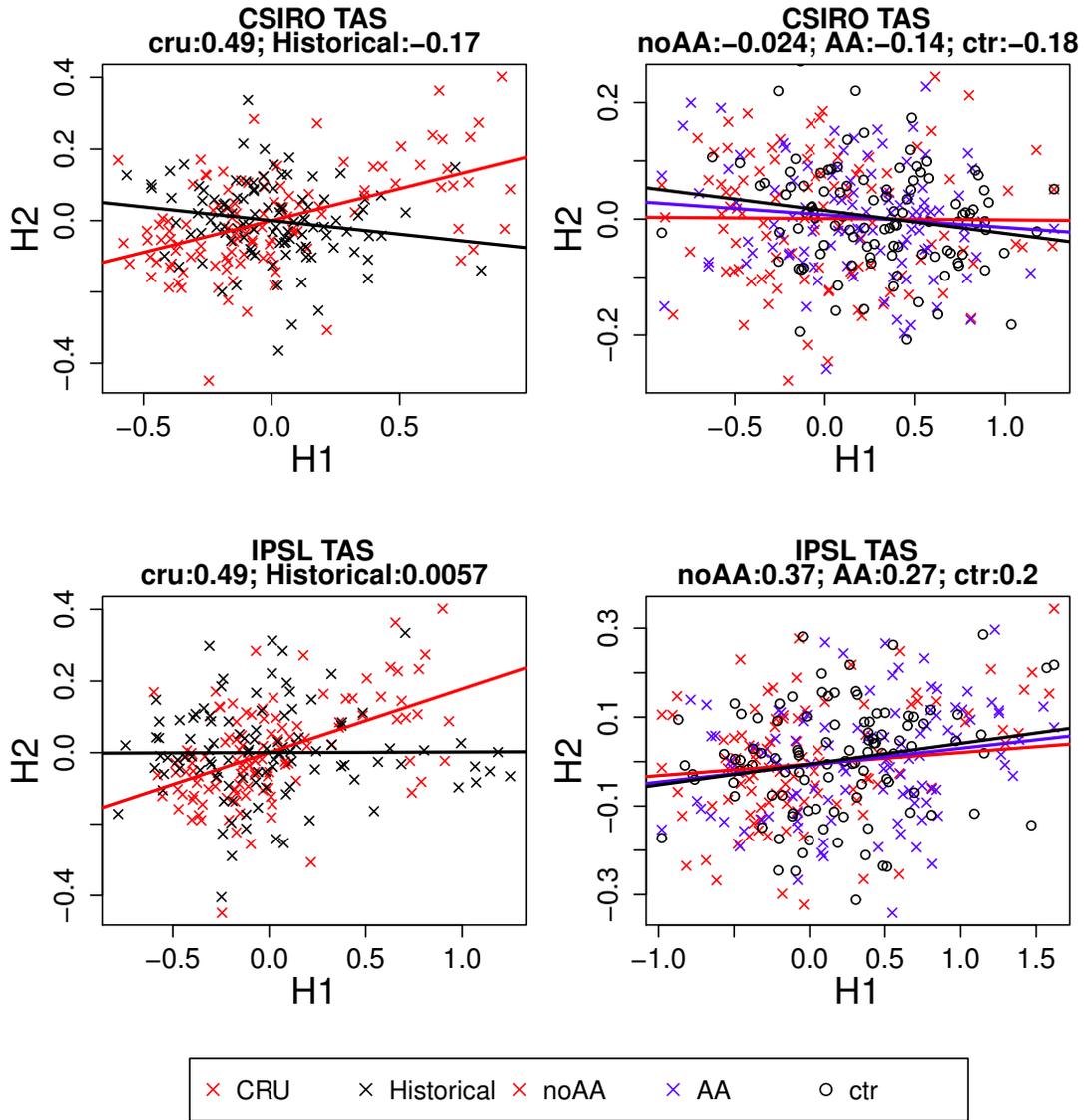


Figure 5.9: Scatter plots of the first (H1) and second (H2) Laplacian time series for surface air temperature over global land, for CSIRO (top) and IPSL (bottom), for historical simulations and observations (left), and AA, noAA and pre-industrial control simulations (right). Different simulations are shown with different colors, as indicated in the legend, and the colored lines show the least squares fit line. The title of each panel shows the correlation of the scatter plots for the respective simulations and observations. The correlation threshold for a 5% significance level is 0.2.

black), which is impressive given that the different curves were estimated separately. Moreover, time series for different ensemble members are relatively close to each other, indicating that the uncertainty due to sampling errors is relatively small compared to the overall secular changes. Time series estimated from CRU observations lie mostly within the spread of the model historical simulations, implying significant consistency between model response and observed changes. The time series from CRU observations exhibit more multi-decadal oscillations than model time series, consistent with the fact that the confidence intervals of forced responses for CRU observations exclude zero while those based on models include zero. Note that the time series are almost indistinguishable between the TAS-only and TAS-PR response vectors, strongly suggesting that most of the forced response signal is contained in the temperature data. Including the precipitation data mostly reduces the uncertainty. This again confirms conclusions obtained based on estimates of confidence intervals.

## 5.7 Forced Responses During 1948-2013

We repeated the above analysis except using the shorter period from 1948 to 2013. This allows us to assess the impact of smaller sample size and different data sets (CRU and GHCN-CAMS). The corresponding confidence intervals are displayed in fig. 5.11 for the joint TAS-PR response vector and in fig. 5.12 for the TAS-only response vector, in the same format as fig. 5.7. To obtain these results, forced response vectors are re-calculated from historical simulations spanning from 1948 to 2004. Though only half of the sample size is used, the derived forced response vectors (not shown) are very similar to those in fig. 5.6.

The top row of fig. 5.11 reveals that neither AA nor noAA is detectable in CSIRO, whereas both forcings are detectable in IPSL, using land average TAS-PR data in a perfect model world. In the case of IPSL, the predicted response is consistent with the response estimated from CRU observations, but underestimates the response estimated from GHCN-CAMS observations, especially in the linear (i.e., trend) component. The sign of the trend

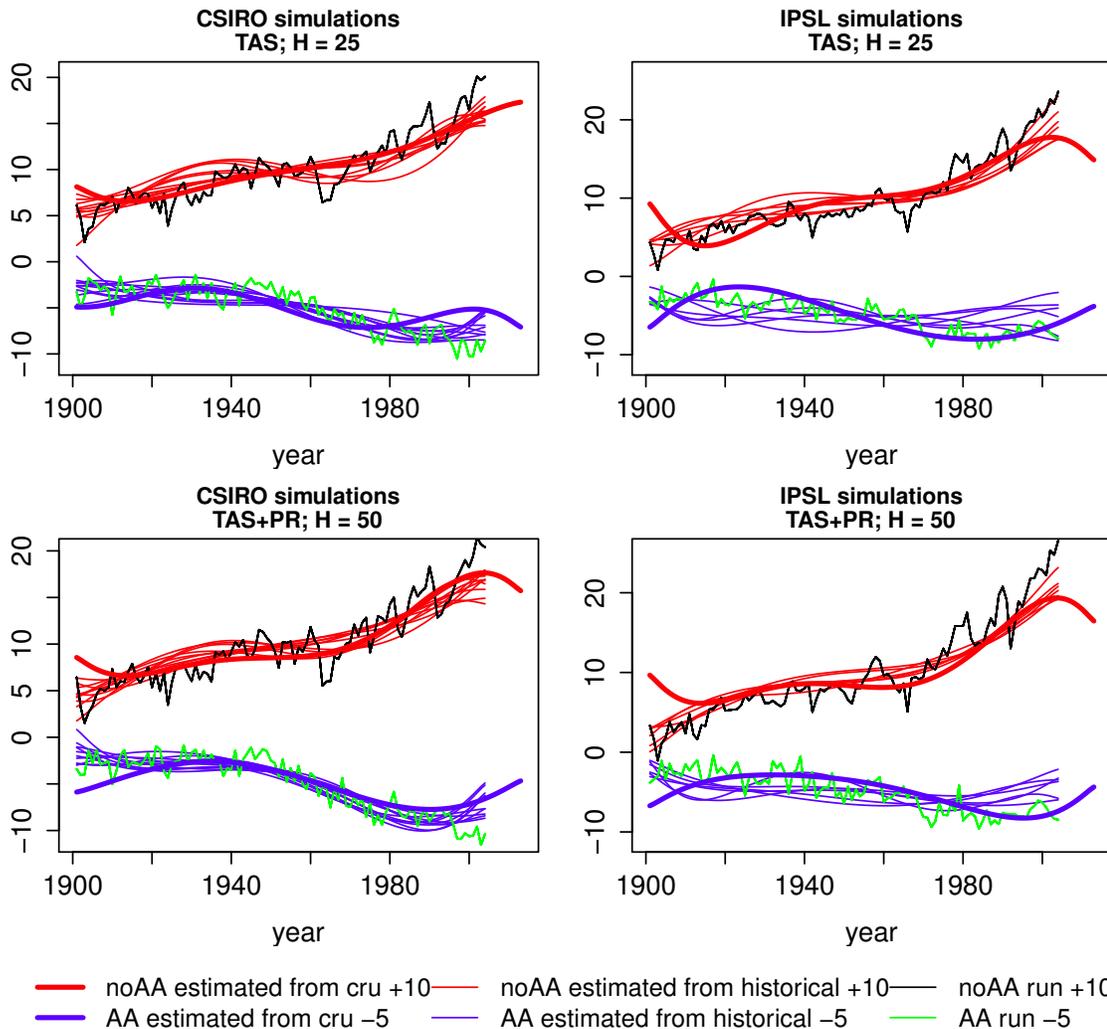


Figure 5.10: Response time series derived from historical simulations (blue and red), derived from single-forcing AA runs (green) and noAA runs (black) and from CRU observations (thick red and thick blue) using the temperature response vector (top rows) and the joint temperature-precipitation response vector (bottom rows) associated with the first 25 Laplacian eigenfunctions and Legendre polynomials 1-5. Different columns correspond to different models used for the historical simulation and response vector, as illustrated in the title. The response vector is derived from only one ensemble member of the single-forcing runs and then used to estimate response time series for each ensemble member from the historical run and observations. Time series from historical simulations are based on using five Legendre polynomials, whereas time series based on single-forcing runs are based on year-by-year estimation.

estimated in IPSL is consistent between models and observations, suggesting that the inconsistency in the sign of the trend seen in fig. 5.7 arises from inconsistencies in the first

half of the 20th century.

In the case of twenty-five Laplacian eigenvectors (bottom row of fig. 5.11), an interesting inconsistency occurs when detecting the trend in AA. For CSIRO fingerprints, the trend is detectable and attributable in the CRU data, but not for the GHCN-CAMS data. In contrast, IPSL fingerprints detect the trend in AA, but *underestimate* the trend in GHCN-CAMS. These discrepancies reveal significant sensitivity in estimating anthropogenic aerosol forcing to observational temperature-precipitation data sets and to model fingerprints on sub-centennial time scales.

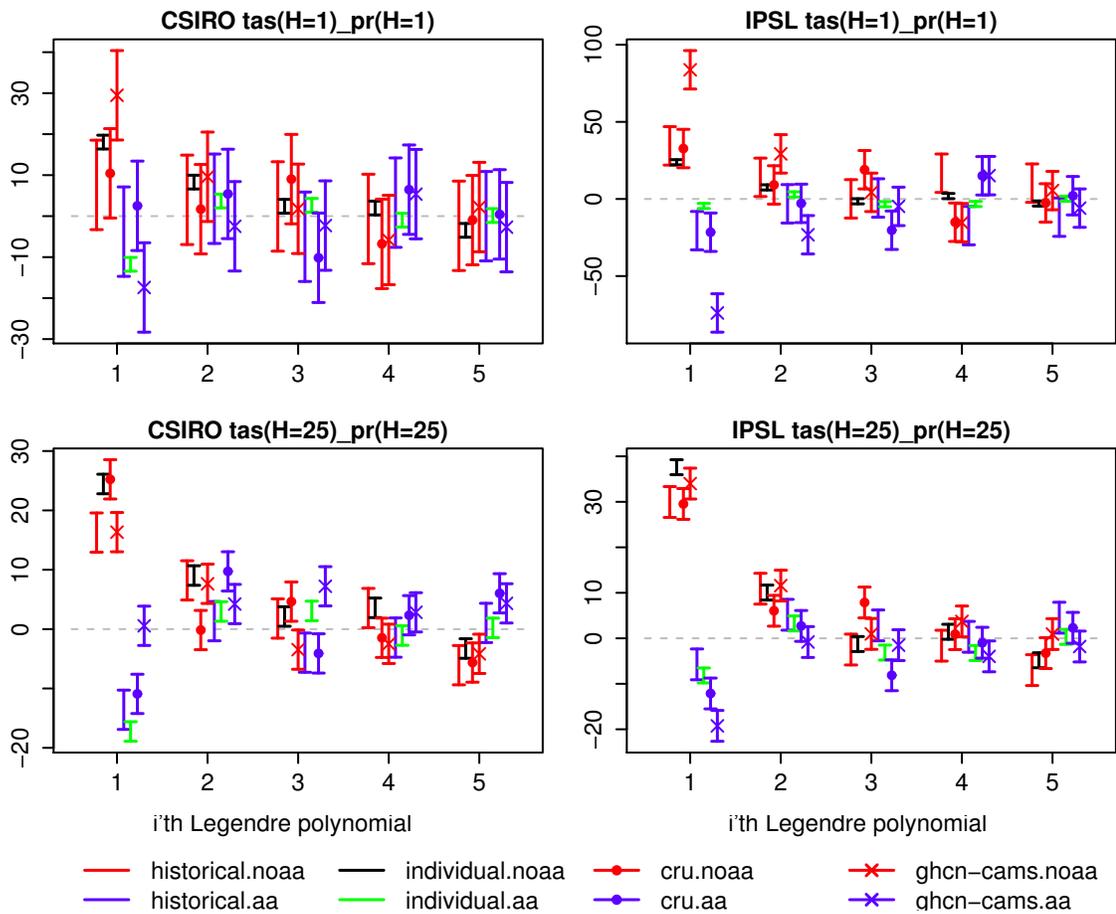


Figure 5.11: Similar to fig. 5.7 except that the response vectors of AA and noAA are derived from single-forcing runs from 1948 to 2004, and observations are CRU and GHCN-CAMS data from 1948 to 2013.

In the case of using two Laplacian eigenvectors of TAS-only over land (top row of

fig. 5.12), the forced response cannot be detected in CSIRO, and gives results that are inconsistent with single-forcing runs in IPSL in a perfect model sense. These results are similar to those using the longer time period shown in fig. 5.9 and can be explained by the same reasons. In the case of twenty-five Laplacian eigenvectors (bottom row of fig. 5.12), forced responses to noAA forcing can be detected and attributed consistently for all observations. An inconsistency occurs again when detecting the trend to AA forcing. For CSIRO fingerprints, the trend is not detectable for either observational data set. In contrast, IPSL fingerprints detect the trend in AA for both CRU and GHCN-CAMS. Since the two observational data sets are consistent and agree with the IPSL responses, we conclude that the AA response in CSIRO is not consistent with observations.

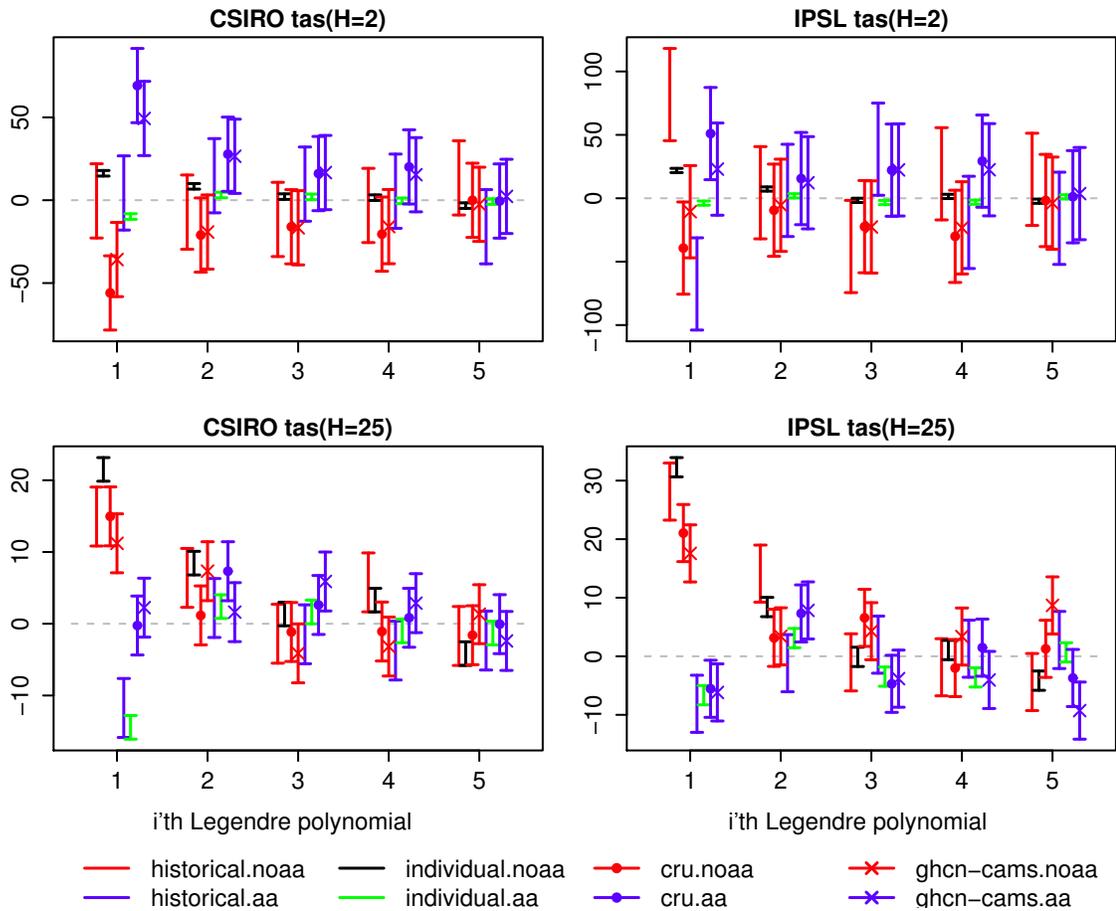


Figure 5.12: Similar to fig. 5.11 except that the response vectors of AA and noAA are derived from annual mean TAS only.

## 5.8 Summary and Discussion

This chapter showed that the time evolution of the response to anthropogenic aerosols can be estimated using annual-mean land temperature over centennial time scales. This conclusion goes beyond traditional fingerprinting studies by using only the spatial structure of the forced response to detect anthropogenic aerosol forcing, leaving the multi-decadal time evolution of the response to be inferred from data. Including annual-mean land precipitation with land temperature only modestly reduces uncertainty over centennial time scales, in contrast to the case of global (land + ocean) data in a perfect model world (DelSole et al. 2015). In addition, multi-decadal variability, as captured by Legendre polynomials 3-5 shown in fig. 5.2, are detected in observations but not in models, suggesting that models are missing significant multidecadal variability in both the anthropogenic aerosol response and the non-anthropogenic aerosol response.

When observations are restricted to sub-centennial time scales (1948-2013), temperature-only fingerprints based on IPSL can detect a trend in AA and noAA responses in observations, but fingerprints based on CSIRO can detect only the noAA response in observations (for sufficiently large number of Laplacian eigenvectors). These results, and the consistency between two different observational data sets, suggest that CSIRO simulates a temperature trend due to AA that is not consistent with observations over this period, whereas IPSL simulates trends in both AA and noAA that are consistent with observations. Using joint land temperature-precipitation information improves the consistency of AA response between CSIRO and CRU, but not with GHCN-CAMS, suggesting sensitivity to the observational data set. No consistent detection of multi-decadal variability occurs over this period using temperature-only fingerprints.

The first two Laplacian eigenvectors over land correspond to the global-land-average and the east-west gradient. Although observations reveal a significant positive correlation between these two components, CSIRO suggests a negative correlation. In addition, IPSL suggests a positive correlation between these two components, but this relation is nearly identical for control and forced simulations, making detection difficult.

This study was motivated by the results of the last chapter, which showed that, in a perfect model world, joint temperature-precipitation information over a global domain provides more accurate estimates of aerosol forced responses than do any other combination of temperature, precipitation, or sea level pressure. Unfortunately, that study also revealed significant inconsistencies in global mean precipitation between models and observations, and even between different observational data sets. Therefore, the present study examined detectability of anthropogenic aerosols using only land data, where satellite observations of precipitation can be calibrated against gauge observations. For a sufficiently large number of Laplacian eigenvectors, the present study finds that joint temperature-precipitation information over land only modestly improves the detectability of anthropogenic aerosols relative to temperature-only information.

## Chapter 6: Potential Detectability of Other Variables and Domains

The previous chapters explored the possibility of improving estimates of aerosol cooling by combining temperature data with precipitation and/or sea level pressure. That analysis is far from exhaustive. In this chapter, we present results for a much wider range of variables for estimating anthropogenic aerosols. To summarize these results efficiently, we fix the dimension and show potential detectability only for eight-dimensional response vectors. The variables, domains, and averaging windows considered here are summarized in table 6. A key addition is the inclusion of ocean data and surface radiation data.

Recall that only model data is needed to evaluate potential detectability. We use the CMIP5 model data described in sec. 2.4. Also, we compute potential detectability when only spatial structure of the forced response is specified, using five Legendre polynomials for temporal structure (fig. 4.2). The potential detectability of forced response to anthropogenic aerosol forcing in CSIRO is shown in fig. 6.1. In total, 36 response vectors are summarized in this figure and plotted in order of potential detectability values (marked by the numbers in square brackets). Red marks variables that have been analyzed in above sections and black marks variables that need to be analysed in future.

The summary figure shows that potential detectability varies greatly depending on different variables, domains and averaging windows. As we saw in chapter 4, among single variables that we have analyzed, annual-mean, global precipitation (“pr\_GLO.ann”) shows the highest ability to detect forced responses to anthropogenic aerosol forcing, followed by annual-mean global near-surface temperature (“tas\_GLO.ann”). Other single-variable such as sea level pressure, ocean temperatures over the North Atlantic and North Pacific, have much lower potential detectability of anthropogenic aerosol forcing.

When ocean surface temperature (“tos”) is combined with other oceanic variables, such

tas	near-surface temperature
tos	ocean surface temperature
pr	precipitation
psl	sea level pressure
rsds	downward shortwave surface radiation
evspsbl	surface evaporation
sos	surface salinity
oh	ocean heat content in upper 700m
LAND	land domain
OCEAN	ocean domain
NASST	North Atlantic SST
NPSST	North Pacific SST
GLO	global domain
ASIALAND	Asia
NHLAND	northern hemisphere land
NALAND	North America
ann	annual mean
djfmamjjason	concatenation of seasonal means

Table 6.1: The variables, domains, and time averaging windows used to define forced response vectors, and the associated shorthand nomenclature for each.

as heat content or salinity, potential detectability is hardly improved. But when ocean surface temperature is combined with downward shortwave radiation at the surface (“rsds”), potential detectability is greatly improved, particularly when the shortwave is available over the northern hemisphere land or Asia land regions. The highest potential detectability is achieved by using sea surface temperatures over the North Pacific and downward shortwave radiation over northern hemisphere land. This potential detectability is much larger than the potential detectability obtained in previous chapters, namely annual mean temperature and precipitation (which ranks only 6th). Combining sea surface temperatures with shortwave radiation over other land areas consistently gives greater potential detectability than other combinations involving sea surface temperatures.

Held and Soden (2006) showed that a robust response to global warming is that the net water flux into the surface—precipitation minus evaporation over the ocean—scales with changes in surface temperature. This result suggests that evaporation may be a useful variable for detection purposes. As can be seen from fig. 6.1, detectability based on combined

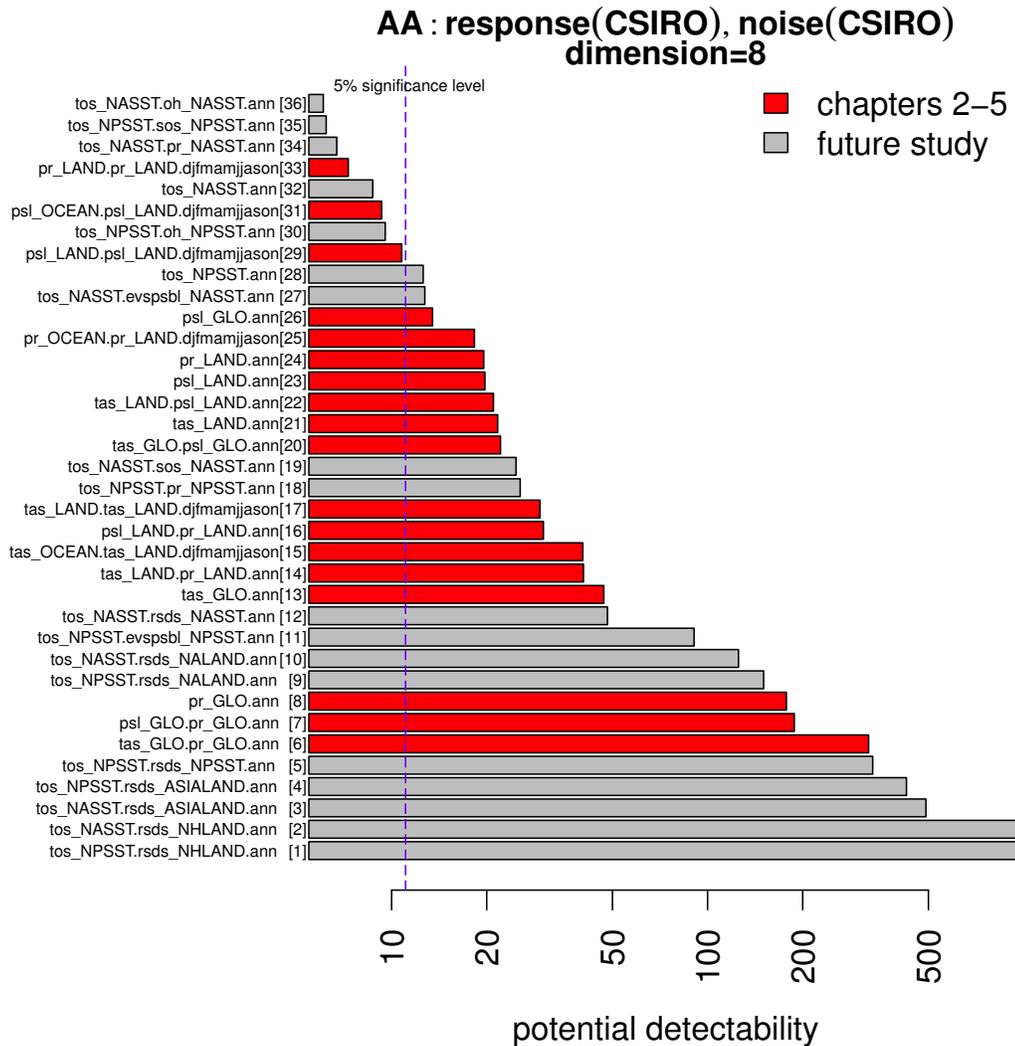


Figure 6.1: Potential detectability of anthropogenic aerosol forcing of different variables, domains, and time averaging using eight-dimensional response vectors from CSIRO. The response vector is defined along the y-axis using the nomenclature in table 6. For instance, ‘tos\_NPSST.rds\_ASIALAND.ann’ shows the potential detectability of joint sea surface temperature over North Pacific and surface downward shortwave over Asia land area. Previously studied variables are marked in red, variables to be analyzed in more depth in the future are marked in grey.

evaporation and sea surface temperature is domain dependent: higher over North Pacific than over North Atlantic. The best detectability achieved with evaporation ranks 11th.

Among all response vectors considered, no dramatic improvement of potential detectability is found by taking into account land-sea contrast or seasonal variations.

To gain an idea of how potential detectability differs for noAA, we show in the top left panel of fig. 6.2 a scatter plot of AA and noAA potential detectability in CSIRO. The number-label corresponds to the ranked response vector in fig. 6.1, and the red line shows  $x = y$  for reference. Fig. 6.2a shows a good linear relationship between the potential detectability of AA and noAA forcing. In general, detectability of noAA tends to larger than that of AA. A similar scatter plot for IPSL is shown in fig. 6.2b, which confirms the near-linear relation between detectability of the two forcings and shows an even stronger tendency for noAA to more detectable than AA.

A comparison of potential detectability of a given forcing between models is shown on the bottom row of fig. 6.2. Interestingly, the potential detectability of noAA is nearly the same for both models (as indicated by the fact that most points in fig. 6.2d lie close to the  $x = y$  line). In contrast, the potential detectability of AA tends to be larger in CSIRO than IPSL. Also, the *ranks* tend to be preserved between the models for AA and for noAA.

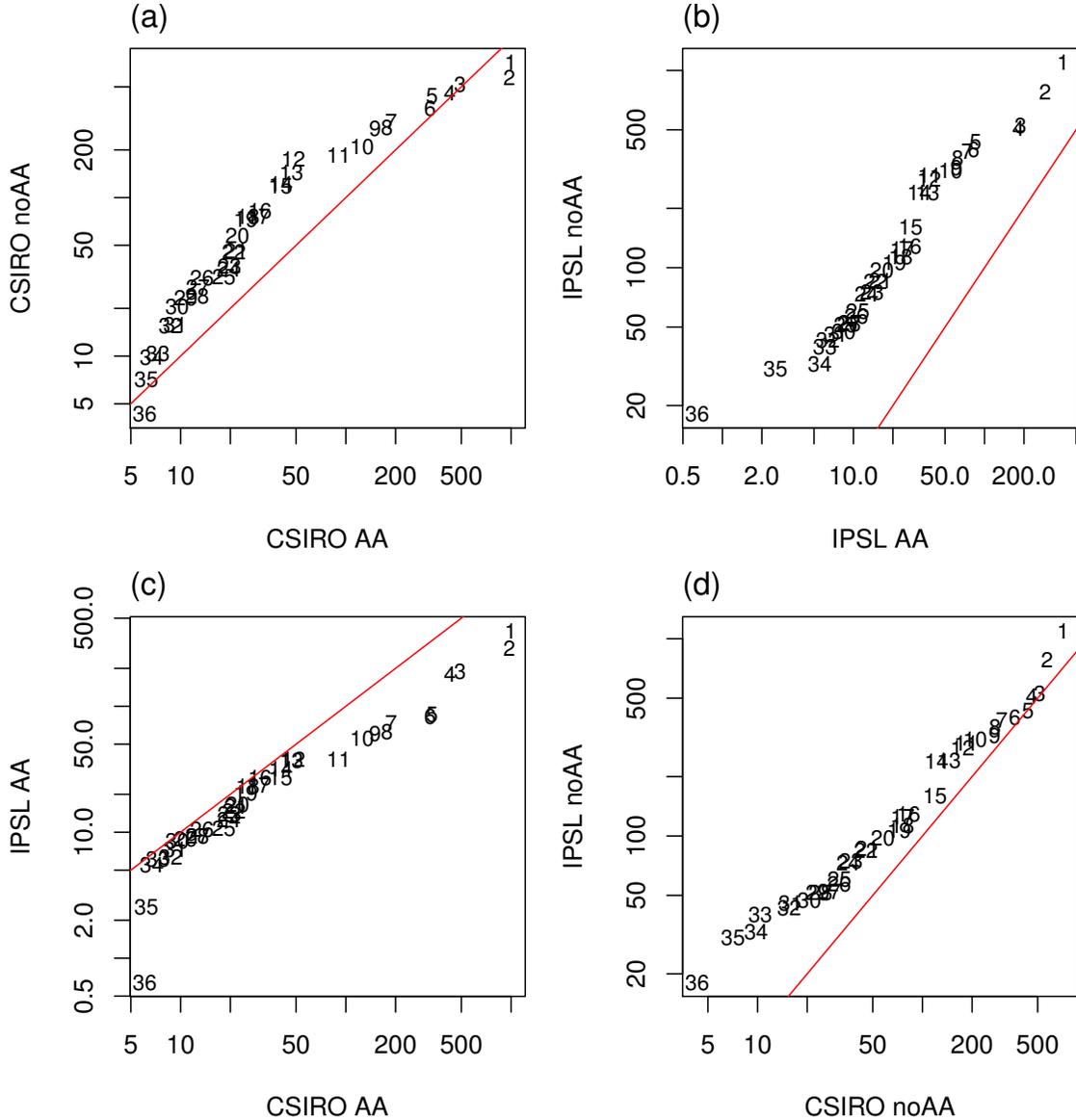


Figure 6.2: Scatter plots of potential detectability to AA and to noAA, in simulations from CSIRO (top left) and IPSL (top right). And scatter plot of potential detectability to AA in simulations from CSIRO and IPSL (bottom left), and scatter plot of potential detectability to noAA in simulations of CSIRO and IPSL (bottom right). All potential detectability are calculated for variables (combinations) shown in fig. 6.1 and the correspondence is shown by the rank numbers. Line of  $x = y$  is displayed in each panel by the red line.

## Chapter 7: Summary

The purpose of this thesis was to investigate ways to improve estimates of the response to anthropogenic aerosols. Our main conclusions are as follows.

We have shown that the response to anthropogenic aerosols and other forcings can be detected using only the spatial structure of the response. While using only spatial structure does not reduce the uncertainties, the fact that time series can be inferred from observations, and that these time series are consistent with model simulations, seems impressive and improves our confidence in the ability of the model to simulate the forced responses. However, while there was strong agreement on the trend, significant multi-decadal variations in the forced response were detected, but models did not simulate these variations. This result suggests that models may be missing important multi-decadal variations in the response to anthropogenic aerosol forcing and in other forcings.

We proposed a new measure, called *potential detectability*, that measures the degree to which a specific forcing can be detected in a perfect model world. Consequently, potential detectability allows the detectability of a forced response in a model to be compared across a variety of variables so that the best choice of variables for detecting a given forced response can be identified prior to analyzing observations. When the forcing is partitioned into two response vectors (e.g., AA and noAA), potential detectability is proportional to the product of the total-to-noise (TNR) ratio of the forced response, times a measure of multicollinearity between the two response vectors. That is,

$$\text{potential detectability} = (\text{total-to-noise-ratio}) (1 - \rho^2), \quad (7.1)$$

where  $\rho$  is the cosine of the angle between the two response vectors (in a certain space).

To our knowledge, this equation has not appeared in the literature, despite its simplicity and fundamental importance. This result substantially clarifies the relation between detectability, signal-to-noise ratio, and multicollinearity.

We showed that in a “perfect model world,” precipitation yields the most accurate estimates of anthropogenic aerosol-induced changes over a global domain than any other single variable analyzed in this study. Even better estimates could be obtained using joint temperature-precipitation information. In the real world, however, observational errors in global precipitation data are too large to permit estimation of aerosol-induced climate changes. Repeating this analysis using only land-data, where reliable rain gauge data are available, succeeds in estimating aerosol cooling. However, adding precipitation does not significantly improve the estimation relative to estimates obtained from temperature alone, as it did in the global domain analysis.

Previous studies found that hemispheric gradients and equator-to-pole gradients were important for estimating aerosol forcing, but we show that this is a model dependent feature. Previous studies also suggested that seasonal variations and land-ocean contrasts can be used to discriminate aerosol-forced responses from other forced responses, but we find that vectors that include this kind of information produce less detectability than vectors of the same dimension but which include only annual means of large-scale basis vectors (i.e., Laplacian eigenvectors). We find that sea-level pressure provides no significant gain in detectability beyond that already available in the surface temperature data from which sea-level pressure is derived. This result calls into question the usefulness of surface pressure data for estimating climate changes.

Note that we investigated two specific types of single-forcing simulations in this study because we were interested primarily in the response to anthropogenic aerosols. However, the methodology presented in this study is not limited to the particular single-forcings studied in this work. In general, our proposed methodology can be applied for any problem that can be formulated as optimal fingerprinting (or equivalently, as generalized least squares).

We examined a large number of variables, domains, and time averaging windows for

estimating anthropogenic aerosol-forced response. Surprisingly, ordering the combinations according to the rank of their potential detectability produces nearly the same order for different forcings and models. We find that the highest potential detectability of anthropogenic aerosol-induced changes is obtained using joint sea surface temperature and downward shortwave radiation over the land surface. This result has motivated us to collect observational records of shortwave radiation and investigate the use of this data for improving estimates of anthropogenic aerosol cooling, and thereby providing a new, independent, and hopefully improved estimate of climate sensitivity.

## Chapter A: Generalization of Optimal Fingerprinting

This appendix shows how optimal fingerprinting can be generalized to estimate time series for forced responses and formally derives the potential detectability measure. We begin by first summarizing generalized least squares (GLS), which is a standard framework in linear regression, and then discuss how optimal fingerprinting relates to GLS. This material is standard Allen and Tett (1999); Hegerl et al. (2007); Hegerl and Zwiers (2011). Then, we show how our modified regression problem can be framed as a GLS problem and then derive potential detectability.

Generalized least squares assumes that an observed  $K$ -dimensional vector  $\mathbf{y}$  can be modeled as

$$\begin{array}{ccccccc} \mathbf{y} & = & \mathbf{X} & \boldsymbol{\beta} & + & \boldsymbol{\epsilon} & \\ K \times 1 & & K \times M & M \times 1 & & K \times 1 & \end{array}, \quad (\text{A.1})$$

where  $\mathbf{X}$  is a matrix of known “explanatory variables,”  $\boldsymbol{\beta}$  is a vector of unknown regression coefficients to be estimated from data, and  $\boldsymbol{\epsilon}$  is a random noise term. Ordinary Least Squares (OLS) gives estimates of  $\boldsymbol{\beta}$  when elements of the noise term  $\boldsymbol{\epsilon}$  are independent and identically distributed. In contrast, GLS gives estimates of the regression coefficients when elements of the noise term are correlated, or more precisely, have covariance matrix  $\tilde{\Sigma}_\epsilon$ ; i.e.,

$$\text{cov}[\boldsymbol{\epsilon}] = \tilde{\Sigma}_\epsilon. \quad (\text{A.2})$$

A standard fact is that a variable transformation can adjust the covariance matrix so that OLS can be used. The end result is that the GLS estimates of the regression coefficients is

$$\hat{\boldsymbol{\beta}} = \left( \mathbf{X}^T \tilde{\Sigma}_\epsilon^{-1} \mathbf{X} \right)^{-1} \mathbf{X}^T \tilde{\Sigma}_\epsilon^{-1} \mathbf{y}, \quad (\text{A.3})$$

where the caret  $\hat{\cdot}$  denotes a sample estimate. The estimate (A.3) is unbiased, in the sense that  $E[\hat{\boldsymbol{\beta}}] = \boldsymbol{\beta}$ , and has covariance matrix

$$\text{cov}[\hat{\boldsymbol{\beta}}] = \left( \mathbf{X}^T \tilde{\Sigma}_\epsilon^{-1} \mathbf{X} \right)^{-1}. \quad (\text{A.4})$$

If the noise is normally distributed, then  $\hat{\boldsymbol{\beta}}$  also is normally distributed, and a 95% confidence interval for the  $m$ 'th regression coefficient is

$$\left( \hat{\beta}_m - 1.96 \left( \left( \mathbf{X}^T \tilde{\Sigma}_\epsilon^{-1} \mathbf{X} \right)^{-1} \right)_{mm}^{1/2}, \hat{\beta}_m + 1.96 \left( \left( \mathbf{X}^T \tilde{\Sigma}_\epsilon^{-1} \mathbf{X} \right)^{-1} \right)_{mm}^{1/2} \right). \quad (\text{A.5})$$

Optimal fingerprinting assumes a model of the form (2.7), repeated below:

$$\mathbf{o} = \sum_{m=1}^M \mathbf{f}_m a_m + \eta. \quad (\text{A.6})$$

In this study, we assume that the forced response  $\mathbf{f}_m$  is represented as (2.8), where the response pattern  $\mathbf{p}_m$  has been derived by maximizing the signal-to-noise ratio (i.e., by solving (2.13)), and use GLS to estimate the time series  $\mathbf{r}_m$  (which subsumes the scaling factor  $a_m$ ). In addition, we assume the response time series  $\mathbf{r}_m$  can be represented by a low-order polynomial in time (2.9). Under these assumptions, the forced response vector  $\mathbf{f}_m$  at the  $s$ 'th spatial location and  $n$ 'th time step is

$$(\mathbf{f}_m)_{s,n} = p_{s,m} r_{n,m} = \sum_{j=1}^J p_{s,m} c_{m,j} L_{n,j}. \quad (\text{A.7})$$

To express the final regression model in the form (A.1), it is convenient to “stack” the

spatial locations first and then loop over time, leading to the following identifications:

$$y_{s+S(n-1)} = o_{s,n} \quad (\text{A.8})$$

$$X_{s+S(n-1),m+M(j-1)} = p_{s,m} L_{n,j} \quad (\text{A.9})$$

$$\beta_{m+M(j-1)} = c_{m,j} \quad (\text{A.10})$$

$$\epsilon_{s+S(n-1)} = \eta_{s,n}. \quad (\text{A.11})$$

As discussed in the main text, we assume internal variability is Gaussian white noise. Under the white noise assumption, the covariance matrix of the noise term has the block diagonal structure

$$\tilde{\Sigma}_\epsilon = \begin{pmatrix} \Sigma_\eta & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \Sigma_\eta & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \Sigma_\eta \end{pmatrix}, \quad (\text{A.12})$$

where  $\Sigma_\eta$  is the single-time,  $S \times S$  covariance matrix of internal variability. The covariance matrix of the estimated coefficients is the inverse of the following matrix:

$$\left( \mathbf{X}^T \tilde{\Sigma}_\epsilon^{-1} \mathbf{X} \right)_{m+M(j-1),m'+M(j'-1)} = p_{s,m} L_{n,j} (\Sigma_\eta^{-1})_{s,s'} \delta_{n,n'} p_{s',m'} L_{n',j'} \quad (\text{A.13})$$

$$= (\mathbf{L}^T \mathbf{L})_{j,j'} (\mathbf{p}_m^T \Sigma_\eta^{-1} \mathbf{p}_m)_{m,m'}, \quad (\text{A.14})$$

where we employ the familiar summation convention in which indices that appear twice are summed over all values of that index. Because Legendre polynomials are orthogonal,

$$\mathbf{L}^T \mathbf{L} = \mathbf{I}. \quad (\text{A.15})$$

This orthogonality relation implies that estimates for different coefficients are independent

of each other, and that the matrix (A.13) has the block diagonal form

$$\mathbf{X}^T \tilde{\Sigma}_\epsilon^{-1} \mathbf{X} = \begin{pmatrix} \mathbf{P}^T \Sigma_\eta^{-1} \mathbf{P} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{P}^T \Sigma_\eta^{-1} \mathbf{P} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{P}^T \Sigma_\eta^{-1} \mathbf{P} \end{pmatrix}, \quad (\text{A.16})$$

where  $\mathbf{P}$  is the  $S \times M$  matrix containing the forced response patterns

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \dots & \mathbf{p}_M \end{pmatrix}. \quad (\text{A.17})$$

Similar considerations yield

$$\left( \mathbf{X}^T \tilde{\Sigma}_\epsilon^{-1} \mathbf{y} \right)_{m+M(j-1)} = L_{n,j} p_{s,m} (\Sigma_\eta^{-1})_{s,s'} \delta_{n,n'} \mathbf{O}_{s',n'} \quad (\text{A.18})$$

$$= (\mathbf{p}_m^T \Sigma_\eta^{-1} \mathbf{O} \mathbf{L})_j \quad (\text{A.19})$$

Substituting (A.19) and (A.16) into (A.3) gives the coefficient estimates

$$\hat{c}_{m,j} = \left( (\mathbf{P}^T \Sigma_\eta^{-1} \mathbf{P})^{-1} \mathbf{P}^T \Sigma_\eta^{-1} \mathbf{O} \mathbf{L} \right)_{m,j}. \quad (\text{A.20})$$

Let  $\hat{\mathbf{C}}$  denote the Legendre coefficient matrix defined above, and let  $\hat{\mathbf{c}}_j$  denote the  $M$ -dimensional vector of coefficients for the  $j$ 'th Legendre polynomial, so that

$$\hat{\mathbf{C}} = \begin{pmatrix} \hat{\mathbf{c}}_1 & \hat{\mathbf{c}}_2 & \dots & \hat{\mathbf{c}}_J \end{pmatrix}. \quad (\text{A.21})$$

Then, the covariance matrix of the “stacked” coefficients

$$\text{vec}[\hat{\mathbf{C}}] = \begin{pmatrix} \hat{\mathbf{c}}_1 \\ \hat{\mathbf{c}}_2 \\ \vdots \\ \hat{\mathbf{c}}_J \end{pmatrix} \quad (\text{A.22})$$

is (A.16). The block diagonal structure of the covariance matrix (A.16) implies that  $\hat{\mathbf{c}}_i$  and  $\hat{\mathbf{c}}_j$  are independent for  $i \neq j$ . Moreover, the block diagonals are equal, so the covariance matrix of the estimates for the  $j$ 'th Legendre polynomial are independent of  $j$  and given by

$$\text{cov}[\hat{\mathbf{c}}_j] = (\mathbf{P}^T \Sigma_\eta^{-1} \mathbf{P})^{-1}. \quad (\text{A.23})$$

Thus, a 95% confidence interval for the  $j$ 'th Legendre polynomial for the  $m$ 'th forced response is

$$\left( \hat{c}_{m,j} - 1.96 \sqrt{\left( (\mathbf{P}^T \Sigma_\eta^{-1} \mathbf{P})^{-1} \right)_{mm}}, \hat{c}_{m,j} + 1.96 \sqrt{\left( (\mathbf{P}^T \Sigma_\eta^{-1} \mathbf{P})^{-1} \right)_{mm}} \right) \quad (\text{A.24})$$

For future reference, we note that if the time series were estimated for each year independently, as we do in the case of a single forced response, then the generalized least squares estimate of the time series is

$$\hat{r}_{m,y} = \left( (\mathbf{P}_m^T \Sigma_\eta^{-1} \mathbf{P}_m)^{-1} \mathbf{P}_m^T \Sigma_\eta^{-1} \mathbf{O} \right)_{m,y}. \quad (\text{A.25})$$

## Bibliography

- Adler, R. F., G. Gu, and G. J. Huffman, 2011: Estimating climatological bias errors for the global precipitation climatology project (GPCP). *Journal of Applied Meteorology and Climatology*, **51** (1), 84–99, doi:10.1175/JAMC-D-11-052.1, URL <http://dx.doi.org/10.1175/JAMC-D-11-052.1>.
- Albrecht, B. A., 1989: Aerosols, cloud microphysics, and fractional cloudiness. *Science*, **245**, 1227–1230.
- Allen, M. R., and P. A. Stott, 2003: Estimating signal amplitudes in optimal fingerprinting, Part I: Theory. *Climate Dynamics*, **21**, 477–491.
- Allen, M. R., and S. F. B. Tett, 1999: Checking for model consistency in optimal fingerprinting. *Clim. Dyn.*, **15**, 419–434.
- Allen, R. J., J. R. Norris, and M. Kovilakam, 2014: Influence of anthropogenic aerosols and the pacific decadal oscillation on tropical belt width. *Nature Geosci*, **7**, 270–274.
- Barnett, T. P., and Coauthors, 2008: Human-induced changes in the hydrology of the Western United States. *Science*, **319**, 1080–1083, doi:10.1126/science.1152538.
- Bindoff, N. L., and et al., 2013: Detection and attribution of climate change: From global to regional. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, chap. 10, 867–952.
- Booth, B. B. B., N. J. Dunstone, P. R. Halloran, R. Andrews, and N. Bellouin, 2012:

- Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*, **484**, 228–232.
- Boucher, O., and Coauthors, 2013: Clouds and aerosols. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, chap. 7.
- Chang, C. Y., J. C. H. Chiang, M. F. Wehner, A. R. Friedman, and R. Ruedy, 2010: Sulfate aerosol control of tropical atlantic climate over the twentieth century. *Journal of Climate*, **24**, 2540–2555, doi:10.1175/2010JCLI4065.1.
- Church, J., and Coauthors, 2013: Sea level change. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, chap. 13, 1137–1216.
- Collins, M., and Coauthors, 2013: Long-term climate change: Projections, commitments, and irreversibility. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, chap. 12, 1029–1136.
- DelSole, T., and M. K. Tippett, 2007: Predictability: Recent insights from information theory. *Rev. Geophys.*, **45**, RG4002, doi:10.1029/2006RG000202.
- DelSole, T., and M. K. Tippett, 2015: Laplacian eigenfunctions for climate analysis. *Journal of Climate*, **28**, 7420–7436, doi:10.1175/JCLI-D-15-0049.1.

- DelSole, T., X. Yan, and M. K. Tippett, 2015: Inferring aerosol-forced variability from hydrological sensitivity. *J. Climate*, **submitted**.
- Dong, L., and T. Zhou, 2014: The indian ocean sea surface temperature warming simulated by cmip5 models during the twentieth century: Competing forcing roles of ghgs and anthropogenic aerosols. *J. Climate*, **27**, 3348–3362.
- Dufresne, J.-L., and Coauthors, 2013: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Climate Dynamics*, **40 (9-10)**, 2123–2165, URL <http://dx.doi.org/10.1007/s00382-012-1636-1>.
- Fan, Y., and H. van den Dool, 2008: A global monthly land surface air temperature analysis for 1948–present. *Journal of Geophysical Research: Atmospheres*, **113 (D1)**, n/a–n/a, doi:10.1029/2007JD008470, URL <http://dx.doi.org/10.1029/2007JD008470>.
- Flato, G., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, 741–866.
- Forest, C. E., P. H. Stone, A. P. Sokolov, M. R. Allen, and M. D. Webster, 2002: Quantifying uncertainties in climate system properties with the use of recent climate observations. *Science*, **295 (5552)**, 113–117, doi:10.1126/science.1064419, URL <http://www.sciencemag.org/content/295/5552/113.abstract>, <http://www.sciencemag.org/content/295/5552/113.full.pdf>.
- Forster, P., and Coauthors, 2007: Changes in atmospheric constituents and in radiative forcing. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, Eds., Cambridge University Press, chap. 2, 129–234.

- Gettelman, A., D. Shindell, and J. Lamarque, 2015: Impact of aerosol radiative effects on 2000–2010 surface temperatures. *Climate Dynamics*, 1–15.
- Gillett, N. P., V. K. Arora, D. Matthews, and M. R. Allen, 2013: Constraining the ratio of global warming to cumulative co2 emissions using cmip5 simulations. *Journal of Climate*, **26** (18), 6844–6858, doi:10.1175/JCLI-D-12-00476.1, URL <http://dx.doi.org/10.1175/JCLI-D-12-00476.1>, <http://dx.doi.org/10.1175/JCLI-D-12-00476.1>.
- Gregory, J. M., R. J. Stouffer, S. C. B. Raper, P. A. Stott, and N. A. Rayner, 2002: An observationally based estimate of the climate sensitivity. *Journal of Climate*, **15** (22), 3117–3121, doi:10.1175/1520-0442(2002)015<3117:AOBEOT>2.0.CO;2, URL [http://dx.doi.org/10.1175/1520-0442\(2002\)015<3117:AOBEOT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(2002)015<3117:AOBEOT>2.0.CO;2).
- Harris, I., P. Jones, T. Osborn, and D. Lister, 2014: Updated high-resolution grids of monthly climatic observations – the CRU ts3.10 dataset. *International Journal of Climatology*, **34** (3), 623–642, doi:10.1002/joc.3711.
- Hegerl, G., and F. Zwiers, 2011: Use of models in detection and attribution of climate change. *WIREs Clim Change*, **2**, 570–591.
- Hegerl, G. C., and Coauthors, 2007: Understanding and attributing climate change. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, 663–745.
- Held, I. M., and B. J. Soden, 2006: Robust responses of the hydrological cycle to global warming. *Journal of Climate*, **19** (21), 5686–5699, doi:10.1175/JCLI3990.1, URL <http://dx.doi.org/10.1175/JCLI3990.1>.
- Iles, C. E., G. C. Hegerl, A. P. Schurer, and X. Zhang, 2013: The effect of volcanic eruptions on global precipitation. *Journal of Geophysical Research: Atmospheres*, **118**, 8770–8786, doi:10.1002/jgrd.50678.

- IPCC, 2012: Managing the risks of extreme events and disasters to advance climate change adaptation. *A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*, C. Field, V. Barros, T. Stocker, D. Qin, D. Dokken, K. Ebi, M. Masstrandrea, K. Mach, G.-K. Plattner, S. Allen, M. Tignor, and P. Midgley, Eds., Cambridge University Press, 582.
- IPCC, 2014: Climate change 2014: Synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change. 151 pp.
- Jia, L., and T. DelSole, 2012: Optimal determination of time-varying climate change signals. *J. Climate*, **25**, 7122–7137.
- Jones, G. S., P. A. Stott, and N. Christidis, 2013: Attribution of observed historical nearsurface temperature variations to anthropogenic and natural causes using cmip5 simulations. *Journal of Geophysical Research: Atmospheres*, **118**, 4001–4024, doi:10.1002/jgrd.50239.
- Jones, G. S., S. F. B. Tett, and P. A. Stott, 2003: Causes of atmospheric temperature change 1960–2000: A combined attribution analysis. *Geophysical Research Letters*, **30**, n/a–n/a, doi:10.1029/2002GL016377.
- Knutti, R., and G. C. Hegerl, 2008: The equilibrium sensitivity of the earth’s temperature to radiation changes. *Nature Geosci*, **1** (11), 735–743, URL <http://dx.doi.org/10.1038/ngeo337>.
- Knutti, R., T. F. Stocker, F. Joos, and G.-K. Plattner, 2002: Constraints on radiative forcing and future climate change from observations and climate model ensembles. *Nature*, **416** (6882), 719–723, URL <http://dx.doi.org/10.1038/416719a>.
- Koren, I., Y. J. Kaufman, L. A. Remer, and J. V. Martins, 2004: Measurement of the effect of amazon smoke on inhibition of cloud formation. *Science*, **303**, 1342–1345.
- Lamarque, J.-F., and Coauthors, 2010: Historical (1850–2000) gridded anthropogenic and

- biomass burning emissions of reactive gases and aerosols: methodology and application. *Atmospheric Chemistry and Physics*, **10**, 7017–7039, doi:10.5194/acp-10-7017-2010.
- Lawley, D. N., 1956: Tests of significance for the latent roots of covariance and correlation matrices. *Biometrika*, **43**, 128–136.
- Lean, J. L., and D. H. Rind, 2009: How will Earth’s surface temperature change in future decades? *Geophys. Res. Lett.*, **36**, L15 708, doi:10.1029/2009GL038 932.
- Li, J., R. Yu, W. Yuan, H. Chen, W. Sun, and Y. Zhang, 2015: Precipitation over East Asia simulated by NCAR CAM5 at different horizontal resolutions. *Journal of Advances in Modeling Earth Systems*, **7** (2), 774–790, doi:10.1002/2014MS000414.
- Liepert, B. G., and M. Previdi, 2009: Do models and observations disagree on the rainfall response to global warming? *Journal of Climate*, **22** (11), 3156–3166, doi:10.1175/2008JCLI2472.1, URL <http://dx.doi.org/10.1175/2008JCLI2472.1>.
- Lohmann, U., 2006: Space sci. rev. *Aerosol effects on clouds and climate*, **125**, 129–137.
- Ma, P.-L., and Coauthors, 2015: How does increasing horizontal resolution in a global climate model improve the simulation of aerosol-cloud interactions? *Geophysical Research Letters*, **42** (12), 5058–5065, doi:10.1002/2015GL064183, URL <http://dx.doi.org/10.1002/2015GL064183>, 2015GL064183.
- Marvel, K., and C. Bonfils, 2013: Identifying external influences on global precipitation. *Proc. Natl Acad. Sci.*, **110**, 19 301–6.
- Masson-Delmotte, V., and Coauthors, 2013: Information from paleoclimate archives. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press.

- Molina, M., and Coauthors, 2014: *What We Know: The reality, risks, and response to climate change*. American Association for the Advancement of Science, 28 pp.
- Myhre, G., and Coauthors, 2013: Anthropogenic and natural radiative forcing. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, S. T.F., D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, chap. 8.
- Pierce, D. W., P. J. Gleckler, T. P. Barnett, B. D. Santer, and P. J. Durack, 2012: The fingerprint of human-induced changes in the ocean’s salinity and temperature fields. *Geophysical Research Letters*, **39**, n/a–n/a, doi:10.1029/2012GL053389.
- Ribes, A., S. Planton, and L. Terray, 2013: Application of regularised optimal fingerprinting to attribution. Part I: method, properties and idealised analysis. *Climate Dynamics*, **41 (11-12)**, 2817–2836.
- Ridley, H. E., and Coauthors, 2015: Aerosol forcing of the position of the intertropical convergence zone since AD 1550. *Nature Geosci*, **8**, 195–200.
- Rotstayn, L. D., S. J. Jeffrey, M. A. Collier, S. M. Dravitzki, A. C. Hirst, J. I. Syktus, and K. K. Wong, 2012: Aerosol- and greenhouse gas-induced changes in summer rainfall and circulation in the australasian region: a study using single-forcing climate simulations. *Atmospheric Chemistry and Physics*, **12**, 6377–6404, doi:10.5194/acp-12-6377-2012.
- Rotstayn, L. D., and U. Lohmann, 2002: Tropical rainfall trends and the indirect aerosol effect. *Journal of Climate*, **15**, 2103–2116, doi:10.1175/1520-0442(2002)015<2103:TRTATI>2.0.CO;2.
- Schnur, R., and K. Hasselmann, 2005: Optimal filtering for Bayesian detection and attribution of climate change. *Climate Dynamics*, **24**, 45–55.
- Shindell, D. T., V. A., F. G., and G. Milly, 2012: Precipitation response to regional radiative forcing. *Atmos. Chem. Phys.*, **12**, 6969–6982.

- Stern, D. I., and R. K. Kaufmann, 2000: Detecting a global warming signal in hemispheric temperature series: A structural time series analysis. *Climatic Change*, **47**, 411–438, doi:10.1023/A:1005672231474.
- Stott, P. A., and G. S. Jones, 2009: Variability of high latitude amplification of anthropogenic warming. *Geophysical Research Letters*, **36**, n/a–n/a, doi:10.1029/2009GL037698.
- Stott, P. A., J. F. B. Mitchell, M. R. Allen, T. L. Delworth, J. M. Gregory, G. A. Meehl, and B. D. Santer, 2006: Observational constraints on past attributable warming and predictions of future global warming. *J. Climate*, **19**, 3055–3069.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of cmip5 and the experiment design. *Bull. Amer. Met. Soc.*, **93**, 485–498.
- Twomey, S., 1977: The influence of pollution on the shortwave albedo of clouds. *J. Atmos. Sci.*, **34**, 1149–1152.
- Vaughan, D., and Coauthors, 2013: Observations: Cryosphere. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, chap. 4, 317–382.
- Wehner, M. F., and Coauthors, 2014: The effect of horizontal resolution on simulation quality in the community atmospheric model, cam5.1. *Journal of Advances in Modeling Earth Systems*, **6** (4), 980–997, doi:10.1002/2013MS000276, URL <http://dx.doi.org/10.1002/2013MS000276>.
- Wentz, F. J., L. Ricciardulli, K. Hilburn, and C. Mears, 2007: How much more rain will global warming bring? *Science*, **317** (5835), 233–235, URL <http://www.sciencemag.org/content/317/5835/233.abstract>.
- Xie, S.-P., B. Lu, and B. Xiang, 2013: Similar spatial patterns of climate responses to aerosol and greenhouse gas changes. *Nature Geosci*, **6**, 828–832.

Zhou, J., and K.-K. Tung, 2012: Deducing multidecadal anthropogenic global warming trends using multiple regression analysis. *Journal of the Atmospheric Sciences*, **70**, 3–8, doi:10.1175/JAS-D-12-0208.1.

## Curriculum Vitae

Xiaoqin Yan was born in Sichuan, Southwest of China. She graduated from University of China in 2007 with a Bachelor of Science majoring in Atmospheric Science. She received her Master of Environmental Science in 2011 from University of North British Columbia, Canada. In fall of 2011, she came to George Mason University to pursue a Ph.D. degree in Climate Dynamics with guidance of Prof. Timothy DelSole.