



January 2014

Climate Change Detection And Attribution Using GPS Radio Occultation And CMIP5 GCM Data

Sergey Molodtsov

Follow this and additional works at: <https://commons.und.edu/theses>

Recommended Citation

Molodtsov, Sergey, "Climate Change Detection And Attribution Using GPS Radio Occultation And CMIP5 GCM Data" (2014).
Theses and Dissertations. 1573.
<https://commons.und.edu/theses/1573>

This Thesis is brought to you for free and open access by the Theses, Dissertations, and Senior Projects at UND Scholarly Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UND Scholarly Commons. For more information, please contact zeinebyousif@library.und.edu.

CLIMATE CHANGE DETECTION AND ATTRIBUTION USING GPS RADIO
OCCULTATION AND CMIP5 GCM DATA

by

Sergey Molodtsov

A Thesis

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Science

Grand Forks, North Dakota

May
2014

This thesis, submitted by Sergey Molodtsov in partial fulfillment of the requirements for the Degree of Master of Science from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

Andrei Kirilenko

Xiaodong Zhang

Michael J. Hill

Doug Olsen

This thesis is being submitted by the appointed advisory committee as having met all of the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

Wayne E. Swisher
Dean of the School of Graduate Studies

Date

PERMISSION

Title	Climate change detection and attribution using GPS radio occultation and CMIP5 GCM data
Department	Earth System Science and Policy
Degree	Master of Science

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Sergey Molodtsov
Date 5/8/2014

TABLE OF CONTENTS

LIST OF FIGURES	vi
LIST OF TABLES	viii
ACKNOWLEDGMENTS	ix
ABSTRACT	x
CHAPTER	
I. INTRODUCTION	1
II. LITERATURE REVIEW	4
2.1 Detection and Attribution of climate change	4
2.2 General circulation models	4
2.2.1 What is a GCM?	4
2.2.2 GCM evaluation	6
2.3 Sources of observational climate data	6
2.3.1 Ground stations	7
2.3.2 Radiosondes	7
2.3.3 Microwave Sounding Unit (MSU) satellites	8
2.3.4 Radio Occultation data	9
III. METHODOLOGY	13
3.1 Step by step optimal fingerprinting method	15
3.2 Calculation of scaling factor (β) and its uncertainty range	21
3.3 Data preparation	22

3.3.1 GCM	22
3.3.2 GPS RO	26
3.4 Geopotential height concept	28
IV. RESULTS	29
4.1 Temperature	29
4.1.1 GCM Temperature Trends	29
4.1.2 Radio occultation temperature trends	32
4.1.3 EOF truncation	35
4.1.4 Detection and Attribution Results	39
4.2 Geopotential Height	41
4.2.1 GCM geopotential height trends	41
4.2.2 Radio occultation geopotential height trends	43
4.2.3 EOF truncation	46
4.2.4 Detection and Attribution Results	50
V. DISCUSSION	52
5.1 Temperature trend patterns	52
5.2 Geopotential height trend patterns	58
5.3 Detection and attribution of climate change using GPS RO data	59
5.4 Future research	61
VI. CONCLUSIONS	63
APPENDICES	64
REFERENCES	99

LIST OF FIGURES

Figure	Page
2.1 GPS RO profiling of the atmosphere.....	11
3.1. The scheme of showing cloud of PICTRL data and its two EOFs.....	17
3.2 As in fig. 3.1 but with the red triangle representing a vector of observations or GCMs' output under forcing scenario.....	18
3.3 The transformation of observational vector (or GCMs' output under forcing scenario) from its original coordinates into EOF coordinates.....	19
3.4 The principle of optimal detection.....	21
4.1 Zonal upper air temperature trends (K/decade) for 32 CMIP5 GCMs.....	31
4.2 The average of temperature trends of the 32 CMIP5 GCMs presented on fig. 4.1.....	32
4.3 Zonal upper air temperature trends for GPS RO data.....	33
4.4 Correlation between GPS RO temperature trend pattern (fig. 4.3) and CMIP5 GCMs' temperature trend patterns (fig. 4.1, 4.2).....	34
4.5 First three EOFs of the CMIP5 32 GCM temperature PICTRL data.....	35
4.6 Rebuilt temperature pattern using 1-14 EOFs for GPS RO data (row – 1, 3) and CMIP5 GCM data (row – 2, 4).....	37
4.7 a) The total amount of variance explained by 1- 14 EOFs; b) pattern correlation between the original and rebuilt GPS RO pattern for 1- 14 EOFs; c) pattern correlation between the original and rebuilt GCM pattern for 1- 14 EOFs.....	38
4.8 Results for the uncertainty assessment in the scaling factors (red rhombus) for 1-14 retained EOFs, calculated using temperature data.....	40
4.9 Zonal upper air geopotential height trends for 32 CMIP5 GCMs.....	42

4.10 The average of geopotential height trends of the 32 CMIP5 GCMs presented on fig. 4.8.....	43
4.11 Zonal upper geopotential height trends for GPS RO data.....	44
4.12 Correlation between GPS RO geopotential height trend pattern (fig. 4.10) and CMIP5 GCMs' geopotential height trend patterns (fig. 4.8, 4.9).....	45
4.13 First three EOFs of the CMIP5 32 GCM geopotential height PICTRL data.....	46
4.14 Rebuilt geopotential height pattern using 1-14 EOFs for GPS RO data (row – 1,3) and CMIP5 GCM data (row – 2,4).....	48
4.15 a) The total amount of variance explained by 1- 14 EOFs; b) pattern correlation between the original and rebuilt GPS RO pattern for 1- 14 EOFs; c) pattern correlation between the original and rebuilt GCM pattern for 1- 14 EOFs.....	49
4.16 Results for the uncertainty assessment in the scaling factors (red rhombus) for 1-14 retained EOFs calculated using geopotential height data.....	51
5.1 Time series of GPS RO temperature in -1.5°S, 70hPa LxA ponint (blue line) and the QBO index (orange line).....	54
5.2 Time series of GPS RO temperature in 40°N, 300hPa LxA ponint (blue line) and the ENSO index (orange line).....	56
5.3 Zonal upper air temperature trends for GISS-E2-H CMIP5 GCMs.....	58

LIST OF TABLES

Table	Page
1. CMIP 5 GCMs used in study	24
2. Trend patterns used in study	27
3. Correlation between temperature of two selected LxA points with ENSO and QBO indexes with different time lags	55

ACKNOWLEDGMENTS

I wish to express my special thanks to my advisor, Andrei Kirilenko, who helped me a lot during my work. I would like to thank my committee members, Xiaodong Zhang, Doug Olsen and Michael Hill, for their valuable comments and advice. I want to thank Tatiana Molodtsova for her help and support. Also I want to thank my colleagues in ESSP for their friendship, support and help.

ABSTRACT

General circulation models (GCMs) show a distinct anthropogenic fingerprint - the thermal expansion of Hadley Cell. This response to the increase of atmospheric concentration of greenhouse gases (GHG) is evident under a variety of forcing scenarios. However the investigation of the anthropogenic signal in the real climate system is challenging because anthropogenic signal is immersed in the natural variability and it requires highly quality data to separate signal from background variability. GPS Radio Occultation (GPS RO) technique becomes close to meeting all of the quality requirements, enabling it to become the benchmark for the climate data. The analysis was implemented for upper troposphere - lower stratosphere (UTLS) region between 50°N and -50°S latitudes. Vertical profiles of temperature and geopotential heights from 2001-2006 CHAMP and 2006-2011 COSMIC missions and CMIP5 GCM data for the same variables and time period were used in this study. Whether the anthropogenic signal is distinguishable from natural variability of the climate is being investigated using optimal fingerprinting technique. Temperature trend patterns allow the detection of climate change on 90% significance level but not the attribution, while the geopotential height trend patterns show that the detection of anthropogenic climate influence is achieved on more than 99% significance level.

CHAPTER I

INTRODUCTION

Observed global temperatures show that the past three decades were much warmer than the preceding ones since 1850 (Stocker et al., 2013). The temperature increase is widespread over the globe and is greater at higher northern latitudes (Stocker et al., 2013, Trenberth et al., 2007). Land regions have warmed faster than the oceans. Sea levels have risen at a rate consistent with warming and with the melting rates of glaciers (Stocker et al., 2013, Miller et al., 2013, Rignot et al., 2011). There are numerous climatic datasets and scientific publications, evidencing the climate change. Surface temperature measurements, radiosonde and satellite measurements show warming rates that are consistent with each other and their uncertainties (Karl et al., 2006; IPCC, 2007). The evidence of climate change is summarized in the 5th IPCC report AR-5 (Stocker et al., 2013). Nevertheless, attributing the cause of these changes demands profound climatology research. As described by Le Treut et al. (2007) the attribution of climate change relies on the use and performance of the general circulation models (GCMs). The observations are compared to the GCMs output with anthropogenic forcing and without. If there is an anthropogenic climate signal, it can be distinguished from natural variability of the climate system (Hegerl et al., 2007). The extent of contribution of anthropogenic forcing to the natural climate processes in nowadays climate system is crucial for the society.

This is a challenging task, because it requires both truthful representation of the climate system by GCMs and availability of appropriate observational data. The first problem was discussed by Randall et al. (2007) and it was concluded that even though there is still some uncertainty in

model output, GCMs can give credible representation of the climate system and the better results are achieved using multi-model average.

Secondly, comparing the complex GCMs requires recording of different atmospheric parameters of high accuracy. Radio Occultation, delivering high quality observations of the atmosphere (Kursinski et al., 1997), is believed to be able to facilitate the progress in understanding how climate is changing and, in particular, provide high quality observational data which can be used in the detection and attribution (D&A) research (Leroy et al., 2006a). The focus area of this research is upper troposphere - lower stratosphere (UTLS) region, where the GPS RO data has the highest quality (Ho et al., 2012). In this region the anthropogenic forcing exhibits the distinct pattern which reflects in UT warming and LS cooling in GCMs, thus this region is well suited for D&A research. Moreover there are different GCMs' surface temperature projections for some regions (e.g. Siberia); while UTLS trends are consistent for all the GCMs (Leroy et al., 2006). The previous studies highlight the feasibility of using GPS RO data for climate change attribution research. For example, Leroy et al. (2006a) have tested the IPCC Fourth Assessment Report (IPCC AR4) models using GPS RO. Using optimal fingerprinting techniques the authors estimated how long does it take for anthropogenic signal to emerge in the simulated GPS RO data. In the study the authors suggested that models' predictions for the 21st century can be tested with 95% confidence in 7 to 13 year using GPS RO data. In 2011 Lackner et al. used 8 years of CHAMP satellite data and two month from GPS/MET experiment to investigate the response from three CMIP 3 models to the external forcing. The authors were also using optimal fingerprinting technique and found statistically significant anthropogenic signal. They found that even though 95% confidence level is achieved for temperature trends, GCMs underestimate observed trends by a factor of two.

The main goal of this study is to implement D&A of the climate change, using GPS RO data and CMIP 5 GCM output by applying optimal fingerprinting technique which will optimize the data by natural climate variability. This includes the following steps:

- Calculation of observational and modeled trend patterns using GPS RO data as observations and CMIP5 GCMs as modeled data (section 3.1 methodology; section 4.1.1, 4.2.1 results);
- Calculation of natural variability and its main modes (section 3.1, 3.2 methodology). This requires the use of CMIP5 GCMs under natural forcings only;
- Trend pattern dimension reduction (section 3.4.2 methodology; section 4.1.3, 4.2.3 results);
- Calculation of statistical confidence levels for the anthropogenic climate change signal and assessment of how well modeled data fits observations (section 3.3 methodology; section 4.1.4, 4.2.4 results).

This research generally follows the methodology first suggested by Leroy et al. (2006a) and further developed by Lackner et al. (2011). The use of new set of CMIP5 GCMs, longer time periods and much higher resolution of trend patterns, implemented in this study, improves agreement between GCMs and observations and increases statistical significance of the findings.

CHAPTER II

LITERATURE REVIEW

2.1 Detection and Attribution of climate change

Predicting future climate has a great practical value for the society. However, for the forecasting of the future climate system development, the D&A of today's climate change is necessary.

Moreover the attribution whether its causes are due to the natural variability or due to the human influence has become a center of political and media climate change debate.

Detection of climate change is the result of finding statistical evidence that the climate has changed over time. Attribution is result of providing and proving the causes of the detected change. D&A studies evolved from single time-series of the global mean surface temperature studies (e.g. Wigley and Raper, 1990) to more complex pattern analysis of the free atmosphere (e.g. Lackner et al., 2011). As D&A tasks require the truthful representation of current climate state as well as the reliable natural variability estimation by the models as well as observational data (Allen and Tett, 1999), the following sections provide a review of GCMs and its products as well as the different sources of field data available for D&A research.

2.2 General circulation models

2.2.1 What is a GCM?

A general circulation model (GCM) is a mathematical model of the general circulation of the atmosphere. The model consists of partial differential equations, describing and balancing the most significant processes in the modeled system. Due to computational constraints some processes need to be solved through parameterization and tuning of the GCMs (Randall et al.,

2007). Among pioneers in the climate modeling it is the work by Mintz and Arakawa in 1964, who simulated global climate with two atmospheric levels and accounting for geography, oceans and ice cover (Arakawa, 1970). Manabe and Bryan (1969) developed the first model designed for decade to century time scale and suitable for climate research applications. Another effort to produce GCM capable to simulate atmosphere processes was undertaken by Cecil E. Leith, who developed climate model with a five atmospheric levels (Leith, 1964). Since then a large variety of climate models of increasing complexity has been developed. GCMs are widely used in climatic research – from studies devoted to the research of past climate – e.g. the climate of “snowball Earth” (Hyde et al., 2000), to studies describing effects of aerosols on clouds’ parameters (Lohmann and Feichter, 1997). The discussion on evolution and complexity of different climate models is provided by Le Treut et al. (2007).

The Coupled Model Intercomparison Project (CMIP) is the state of the art set of climate integrations, produced with multiple GCM running a unified set of scenarios. The current implementation CMIP5 developed for the Fifth IPCC assessment report (Stocker et al., 2013), replacing previous set of CMIP3 models (Meehl et al., 2007). Among the advancements of CMIP5 are: higher spatial and vertical resolution comparing to the preceding CMIP3 set of models, accounting for land/ocean carbon cycles, improvements in the modeling of the aerosol effects, use of volcanic and solar forcing and many others (Taylor et al., 2012, Sillmann et al., 2013). A new set of forcing scenarios (Moss et al., 2010, Taylor et al., 2012) was created to replace the Special Report on Emissions scenarios (SRES) (Nakicenovic et al., 2000). Besides the conventional long term simulations, a new set of short-term decadal projections was developed (Meehl et al., 2009, Taylor et al., 2012). Knutson et al. (2013) assessed CMIP5 GCMs output and found that the models realistically represent natural climate variability.

2.2.2 GCM evaluation

The climatic research requires the truthful representation of climate system by the GCMs, which implies the models performance assessment. So, the high attention is given to the evaluation of the GCMs (e.g., Collins et al., 2006; Randall et al., 2008; Griffies et al., 2011). The GCMs' evaluation can be done on system or on component level. The component level approach is based on studying the separate components of the GCM (e.g. Gates et al., 1999). After evaluating different GCM components it is necessary to check how all the components work in a full model. Thus system evaluation of the GCMs is based on comparing the output from the full GCM against observations. While there are still uncertainties in GCM performance found by component evaluation studies (e.g. Charlton-Perez et al., 2013), on system level Sillmann et al. (2013) showed that the models are able to realistically represent the current climate and its variability which is crucial for D&A research.

2.3 Sources of observational climate data

The calculation of predicted human forced climate system response and climate natural variability by GCMs in order to investigate if the forced climate trends are statistically different from the natural variations is not the only thing required for D&A research. To see if current climate trends are consistent with predicted forced scenarios the accurate monitoring of climatic parameters is essential for such studies. The review of different available observational data sources is provided in this section.

2.3.1 Ground stations

Ground stations' measurements (e.g. integrated in CRUTEM4 (Morice et al., 2012) or Global Historical Climatology Network (GHCN) (Menne et al., 2011)) provide the longest record of observations for different atmospheric parameters (Jones et al., 1999; Brohan et al., 2006; Trenberth et al., 2007). The time period of ground stations' observations spans over 150 years. The data from ground stations are delivering the information only on the air close to the Earth's surface, whereas for the purposes of this study the information on the atmospheric parameters of the upper troposphere and lower stratosphere is required.

2.3.2 Radiosondes

Radiosondes, the sensors attached to the balloons inflated with helium/hydrogen, provide temperature, pressure and relative humidity profiles of the atmosphere up to 35 km. They resolve some of the issues inherent to ground measurements by providing the data on the atmospheric profiles of the atmosphere. They monitor troposphere and stratosphere and have a long enough for climate research record period spanning from 1940s. There have been developed different adjustment methods (Parker et al., 1997, Durre et al., 2002) for radiosonde data which led to creation of various datasets (e.g. Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC); Hadley Centre's radiosonde temperature product (HadAT) (McCarthy et al., 2008)). However the main limitation of this data is the uneven land coverage and no ocean coverage with a bias toward Northern Hemisphere mid-latitudes. Moreover the temperature trends obtained from radiosondes are prone to errors, caused by using different observational practices in different countries (Zhai and Eskridge, 1996), the changing of instruments (Gaffen et al., 2000), and solar heating of the temperature sensor (Sherwood et al., 2005).

Even though the observational period for radiosonde data is long enough for climatology study, the temperature trend studies based on such data show that radiosondes are not able to contribute to the understanding of upper air temperature trends. Thorne et al. (2005) conducted the analysis of 1958-2002 temperature trends obtained from the HadAT data, and reported that the results do not contribute to the comprehension of late 20th century free atmospheric temperature changes.

2.3.3 Microwave Sounding Unit (MSU) satellites

The dramatic improvement in climate data was brought by the use of the satellites. The first microwave sounding unit (MSU) was launched in 1978 on TIROS-N series of National Oceanic and Atmospheric Administration (NOAA) satellite. A series of eight satellites, which carried MSU sensors were launched in following years. The MSU satellites measure microwave oxygen's radiation using different frequencies, which allows examining multiple layers of the atmosphere. An early climatologic study by Spencer and Christy (1990) proved the usefulness of the MSU sensors, which provide long-term global coverage measurements of temperature. The MSU satellites have some constraints for the observations which they provide. Since different satellites use different instruments and algorithms to collect and process data, there is difficulty in getting reliable interplatform data. These constraints include: instrument body temperature effect discussed by Christy et al. (1998, 2000); the effect due to orbital decay (loss of altitude) studied by Wentz and Schabel (1998); the orbital drift (east-west movement and local diurnal cycle variations from changing earth emissions which affect the data) (Christy et al. 2000); the errors caused by merging different MSU satellites data and the calibration of sensors (Christy et al. 1998), 2000, Zou et al. 2006).

Further improvement to the satellite atmospheric data was brought by advanced microwave sounding units (AMSU) with more frequency channels, which allow measuring the atmosphere

at larger numbers of layers. Another advantage of AMSU is that its measuring footprint is smaller comparing to MSU sensors, which leads to higher spatial resolution and accuracy (Kidder et al., 2000). Different studies were done on merging MSU and AMSU data (Goldberg and Fleming, 1995; Christy et al., 2003), however AMSU sensors inherit from MSU the same difficulties related to the calibration and merging of the data between multiple satellites. The AMSU unit on AIRS satellite (Pagano et al., 2010) compared to in-situ aircraft temperature measurements showed the discrepancy of more than 1°C (Diao et al., 2013). In spite of the developing of new calibration schemes using simultaneous nadir overpasses in order to remove some errors obtained by merging different satellites' datasets (Zou et al., 2006; Zou et al., 2009) the MSU/AMSU data still have inter-satellite sources of errors which lead to uncertainties in the calculated trends (Thorne et al., 2011). It is important to mention that both MSU/ AMSU and radiosonde sensors were not designed for the climate monitoring purposes (Randel et al., 2009).

2.3.4 Radio Occultation data

The GPS RO is a relatively new source of remotely sensed atmospheric profiles. The history of RO data started with the first experimental GPS RO satellite – GPS/MET, which was launched in April 1995 to provide about 150 soundings per day (Ware et al., 1996). After the successful GPS/MET mission (Rocken et al., 1997) the German research satellite CHALLENGING Minisatellite Payload for geoscientific research (CHAMP) was launched in July 2001. It was aimed to provide multi-year RO based climatologies for a period of more than 5 years. CHAMP provides about 230 RO profiles per day (Wickert et al., 2001). In April 2006 six identical micro FORMOSAT-3/Constellation Observation System for the Meteorology, Ionosphere, and Climate (FC) (COSMIC) satellites were launched, providing 1000-2500 daily RO profiles in neutral atmosphere (Anthes et al., 2008).

The GPS signal passes the Earth's atmosphere and is received by a low orbit RO satellite (fig. 2.1). Due to atmospheric refractivity the signal bends which causes a time delay in signal receiving. These delays, which are measured with precise atomic clocks, are used to calculate the signals' bending angle. As the GPS satellite rises from behind the Earth, it provides bending angle profiles of the atmosphere, bottom to the top. Using the bending angle profiles, it is possible to calculate the refractivity profiles applying Abel transformation (Syndergaard 1998): the refractivity (N), or microwave index of refraction (n), is a function of temperature and pressure (equation 2.1):

$$N = (n - 1) * 10^6 = a \frac{p}{T} + b \frac{p_w}{T^2}, \quad (2.1)$$

where constants $a = 77.6 \text{ K hPa}^{-1}$ and $b = 3.73 * 10^5 \text{ K}^2 \text{ hPa}^{-1}$, p – pressure (hPa), p_w partial pressure of water vapor (hPa), T – temperature (K). In the equation the second “wet” component can be neglected with the assumption of dry atmosphere (above 300 hPa). Hence, there are multiple parameters that can be derived from RO bending angle assuming the other parameters are fixed, e.g.: refractivity, temperature, pressure, water vapor content and geopotential heights (Kursinski et al., 1997; Leroy, 1997), On the other hand, the possibility to extract the accurate temperature and geopotential heights from the RO measurements are altered by the water vapor present in the atmosphere, so the best available data for the temperature profiles are above 5-7 km (Ho et al., 2012). The contribution of water vapor can be assessed using additional data sources (e.g. radiosondes).

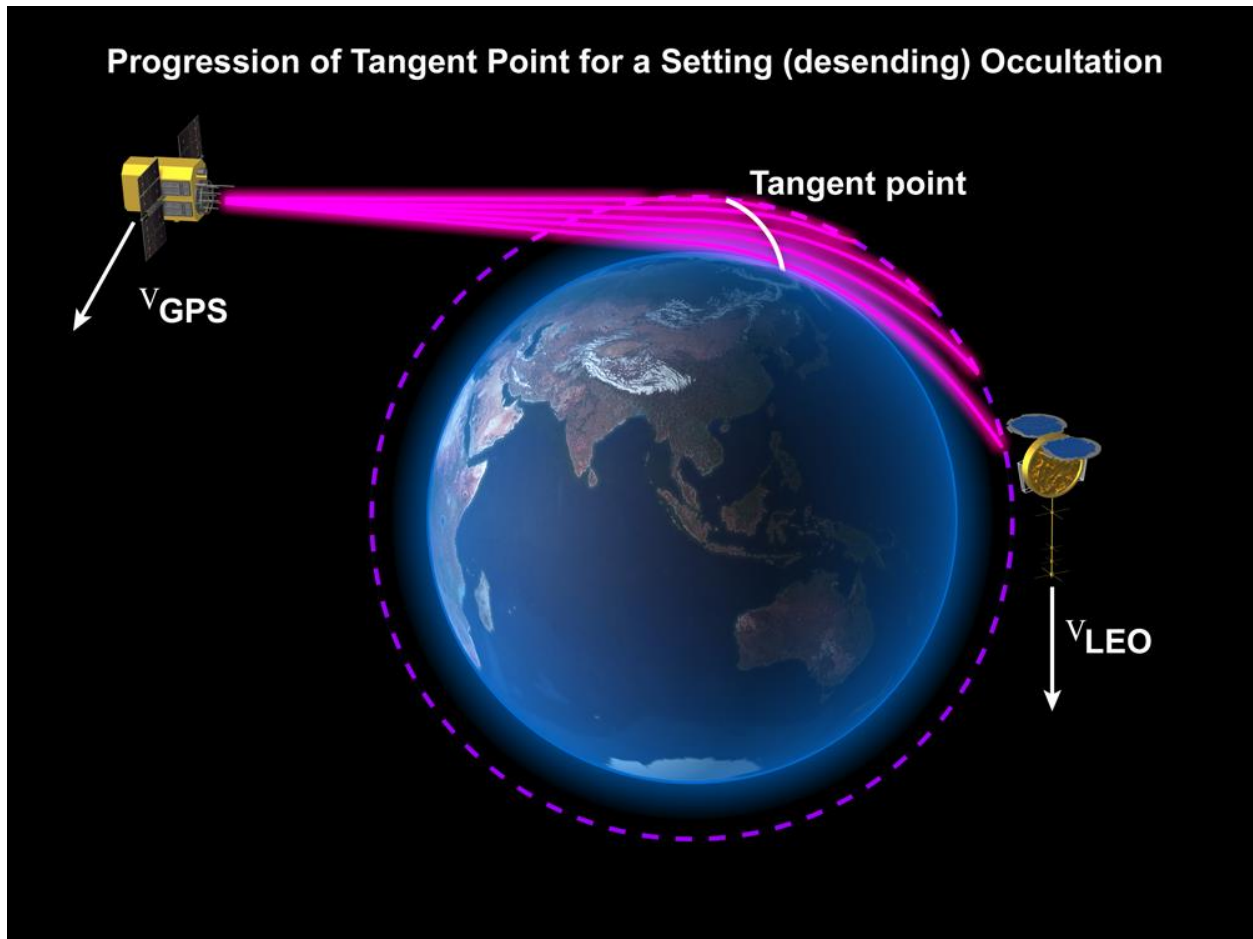


Fig. 2.1 GPS RO profiling of the atmosphere: the signal is received by RO satellite from GPS satellite. Due to refractivity of the atmosphere the signal is bended, allowing the calculation of bending angle profiles (http://www.cosmic.ucar.edu/related_papers/GPS_RO_cartoon.jpg)

Among the advantages of the RO method is weather independency, due to insensitiveness of GPS signals to clouds (Kursinski et al., 1997). Another outstanding advantage is consistency between different RO satellites. It is based on the measurement of time delays with atomic clocks which makes these measurements traceable to the international system of units (Foelshe et al., 2008). Due to its advantageous properties, the RO technique provides the unique opportunity for the climate change research (Leroy et al., 2006b). Numerous studies showed that the RO measurements are suited for the climate monitoring purposes: comparison with radiosondes (Steiner et al., 2009); comparison with MSU/AMSU satellites (Steiner et al., 2007); comparison

with Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) and Global Ozone Monitoring by Occultation of Stars (GOMOS) satellites (Gobiet et al., 2007); comparison with National Centers for Environmental Prediction (NCEP) and European Center for Medium range Weather Forecasting (ECMWF) reanalysis (Borshe et al., 2007; Gobiet et al., 2007). GPS Radio occultation (GPS RO) measurements have potential of becoming a new benchmark in data acquisition, providing new high-quality profiles, or climatologies, containing multiple atmosphere parameters with high vertical resolution (Ho et al., 2009; Steiner et al., 2009) which fits the requirements for the D&A research.

CHAPTER III

METHODOLOGY

The goal of D&A studies is to detect an anthropogenic signal in the observational data. This goal can be achieved only using GCMs in D&A research because it is impossible to separate the contribution of the anthropogenic factor in observations. However it is possible to design different climate experiments accounting for natural forcings only or adding the anthropogenic GHGs contribution and run the model under both human and natural forcings. The climate can be defined as a set of measurable parameters A in a 3-D space that change with time. Then the observations can be compared with both experiments to see if the current evolution of climate system can be represented by natural forcings only. Thus the null hypothesis that current observations of the atmosphere are due to natural variability of the climate can be formulated. If the theory is rejected it is possible to justify the detection of the climate change. However the drivers of this change still have to be found. If the experiment which incorporates known forcings (e.g. natural and anthropogenic forcings) is similar to the observations, the statistical confidence can be computed.

The simplest way to implement D&A research is to use the global mean temperature index and analyze its time-series from different GCM experiments and observations (Griggs et al., 2002). The more advanced study setup is to use spatial structure of observed trends (Knutson et al., 2000) allowing to research the regions where the trends go beyond the natural variability borders.

Different series of D&A approaches is aimed on reducing the influence of natural variability in the data. The easiest way is to run a weighted average filter over the data which allows reducing the internal variability. Bell (1982) and others worked along these lines finding the optimal filter. Another approach is called fingerprinting. A “fingerprint” can be described as a pattern of studied parameter which evolves under anthropogenic forcing against natural variability pattern (Goody et al., 1998). Let us make the following critical "additivity" assumption which allows us represent observed variability (A) as follows:

$$A = A_s + A_i \tag{3.1}$$

where A_s - is the anthropogenic forcing and A_i is the internal variability of climate. Under certain assumptions about natural climate variability, the problem can be re-formulated as a linear regression of observations on climate drivers:

$$A = \sum(\beta_i X_i) + eps, \tag{3.2}$$

where X_i is (known) values of climate drivers, B_i - coefficients (called scalings), eps - natural variability. The hypothesis then is about testing that $\beta_i = 0$ vs. $\beta_i > 0$. Rejection of null hypothesis will mean the detection of climate change. The attribution is done when β_i is close to 1, which means that the models under forcing factors represent well the observed climate change, which implies that natural variability cannot be the reason for the detected change in climate system.

The natural variability is influencing the anthropogenic fingerprint in climate system, making it even harder to detect. The other detection challenges are due to the high dimensionality of the data and the multicollinearity of the parameters. This suggests that data transformation is needed. The optimal fingerprinting allows for data transformation that maximizes signal to variability

ratio, where variability is estimated from GCM simulations prior to optimization. Optimal fingerprinting is now the most popular method of climate change D&A.

For this study the optimal fingerprinting is used as a D&A technique for the recent climate change. Optimal fingerprinting is a generalized multivariate regression adopted for D&A research (Hasselmann, 1997; Allen and Tett, 1999; Hegerl et al., 1996, 2007). There are different versions of this methodology. In this study the ordinary least squares single pattern approach (Hegerl et al., 1996) was used. Section 3.1 provides detailed description of optimal fingerprinting technique; section 3.2 is focused on scaling factor calculation and its uncertainty range calculation., section 3.3 describes the data used in this analysis;

3.1 Step by step optimal fingerprinting method

In this section the detailed step by step description of optimal fingerprinting methodology is presented. Optimal fingerprinting requires using empirical orthogonal functions (EOF) analysis (Lorenz, 1956), thus the description of EOF method is provided first. The EOF aims on finding the major modes of variability in time evolving field. Each mode of the variability is separate EOF. For example it is possible to imagine that sea surface temperature (SST) are analyzed. And EOF analysis applied to time evolving maps of SST finds two main modes of variability (two main EOFs). For example the first EOF could show seasonal cycle and the second could show El Nino pattern.

In this study the main modes of the natural variability are calculated, using the CMIP5 GCMs' output under PICTRL scenario (table 2), as PICTRL scenario refers to GCM runs only under natural forcings. To calculate the EOFs the data should be organized in the following order:

The data matrix $\mathbf{Y}^{\text{PICTRL}}$ (table 2, first part of PICTL data), presented as

$$\begin{pmatrix} n_{1,1} & \cdots & n_{1,p} \\ \vdots & \ddots & \vdots \\ n_{N,1} & \cdots & n_{N,p} \end{pmatrix},$$

where each row is a map of a field, with total amount of rows (maps) $N=160$ (section 3.3.1), and each column is a time series for one point in the field with total amount of columns (time series) $p = 160$ (section 3.3). In order to calculate EOFs the covariance matrix $\Sigma_{800*800}$ is calculated using the data \mathbf{Y}^{PICTRL} matrix using next equation:

$$\Sigma_{800*800} = \frac{1}{N-1} (\mathbf{Y}^{PICTRL})_{800*160}^T \mathbf{Y}_{160*800}^{PICTRL} \quad (3.3)$$

Each covariance matrix can be represented by its eigenvalues (λ) and corresponding eigenvectors ($f_{p \times 1}$) following the next equation:

$$\Sigma_{800*800} = \mathbf{F}_{800*800}^{PICTRL} \mathbf{\Lambda}_{800*800}^{PICTRL} (\mathbf{F}^{PICTRL})_{800*800}^T \quad (3.4)$$

where \mathbf{F} – matrix with eigenvectors f_i and $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues (λ).

Thus the eigenvectors and eigenvalues of this covariance matrix are the EOFs. Each eigenvector can be represented as a map where the mode of variability corresponded to this EOF is presented.

For simplicity the EOF calculation is represented in fig. 3.1. In this case artificial data is used.

Let's assume that our PICTRL data is presented by two dimensional data (in this simplified example each map consists only of two points)

$$\begin{pmatrix} n_{1,1} & n_{1,2} \\ \vdots & \vdots \\ n_{N,1} & n_{N,2} \end{pmatrix},$$

where N – total amount of rows containing only two values (fig. 3.1).

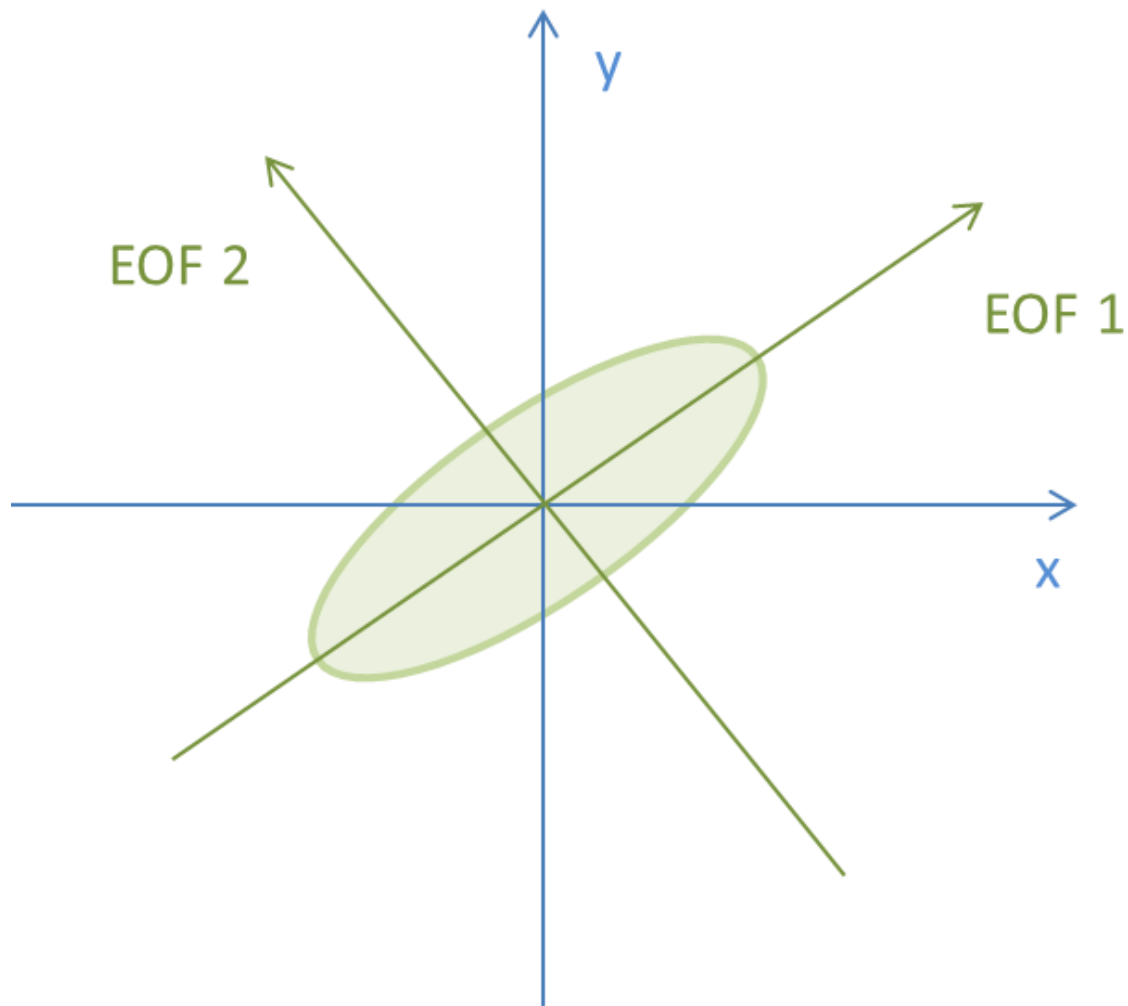


Fig. 3.1 the scheme, showing a cloud of PICTRL data and its two EOFs

So the data are represented by light green ellipse on the plot. In this case the data has the highest variance along the first green line which is the eigenvector of the first EOF calculated following equation 3.3 and equation 3.4, which describes the direction of the highest variance of the data.

The second EOF is represented by the second green line and is describing the second direction of the highest variance within the data.

In this simplified example where the data has with only two dimensions (e.g. the map field containing only two data points), the observational data point will be a red dot in fig. 3.2.

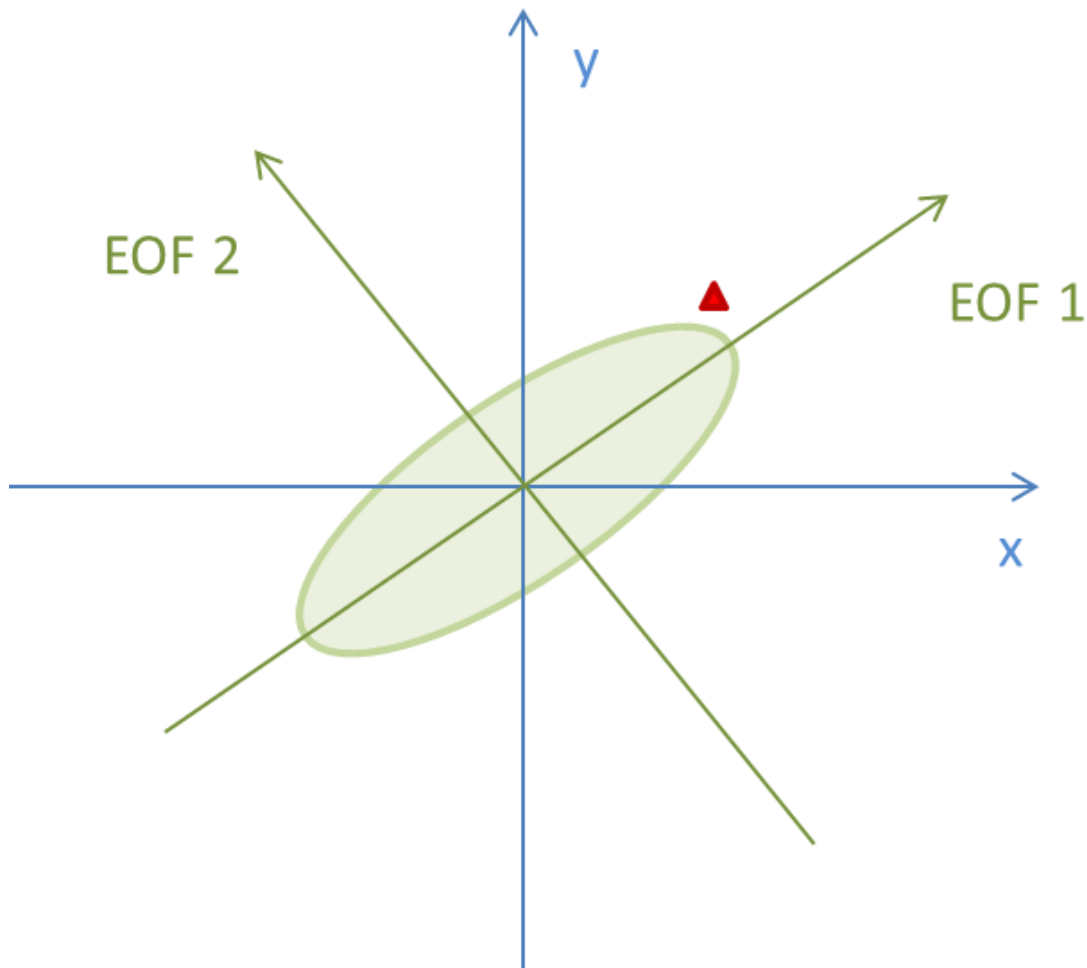


Fig. 3.2 As in fig.3.1 but with the red triangle representing a vector of observations or GCMs' output under forcing scenario

After finding the main modes of natural variability – main modes of PICTRL data, these modes are used (new vectors – two eigenvectors) to represent the observations (or forced GCM pattern). It is possible to represent the data using its original values (fig. 3.2) or transform the data onto

the main modes of variance using calculated eigenvectors (fig. 3.3).

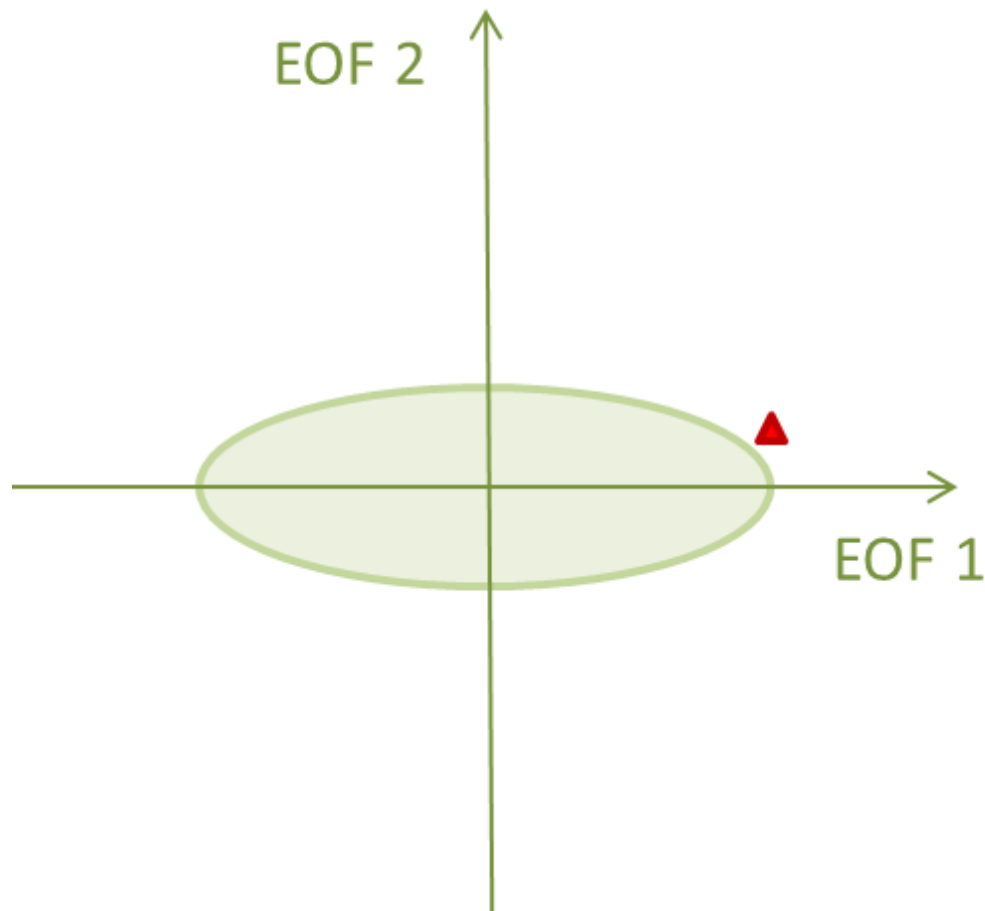


Fig. 3.3 The transformation of observational vector (or GCMs' output under forcing scenario) from its original coordinates into EOF coordinates

On this plot the abscissa line is the first EOF eigenvector and ordinate line is the second eigenvector from the previous plot. However our real data vector has 800 points (section 3.3 – LxA grid), so it has 800 dimensions. EOF analysis helps finding the main modes of variability which explain the highest direction of variance within the data. The main goal of optimal fingerprinting is to represent the data in a dimension reduced space –it can be done following the previous steps, but instead of having two dimensions, 800 dimensions are used in this study. The original multidimensional data can be represented in, as in current example, on a two-

dimensional plane which is described by two main EOFs as it could be done in fig. 3.3 if 800 original dimensions are reduced only to 2 dimensions. The dimension reduction (or transformation the data vector from original 800 dimensions to k – dimensional space) is done using the next equation:

$$b_k^{data} = \frac{(\mathbf{F}^{PICTRL})_{k*800}^T (\mathbf{Y}^{data})_{800*1}^T}{(\mathbf{F}^{PICTRL})_{k*800}^T \mathbf{F}_{800*k}^{PICTRL}} \quad (3.5)$$

where k – is selected number of EOF (a subset $k \leq 800$, as 800 is total number of EOFs), \mathbf{Y}^{data} can be vector of GPS RO data (\mathbf{Y}^{data}) or GCM under forced scenario (\mathbf{X}^{data}) (or individual trends from second part of PICTRL data ($\mathbf{Y}^{CONTROL}$) - section 3.3), $(\mathbf{F}^{PICTRL})_{k*800}^T$ is matrix of selected k amount of eigenvectors (f_{px1}) (section 3.3 and table 2 for data description). If $k = 800$ the pattern would have its original number of dimensions, so it would be in its original form, because each eigenvector (each EOF) represent a fraction of total variance, thus when all the eigenvectors are used - the total amount of variance is represented.

The optimal fingerprinting technique uses the dimension reduced and projected on selected EOFs observed and forced GCM data vectors, and uses multivariate regression algorithm (equation 3.4) in order to find β which shows how good the modeled data fits the observations. This regression algorithm also finds the direction of the least influence of natural variability (fig. 3.4).

$$\beta = \frac{((b_k^{Xdata})^T \Lambda^{-1} b_k^{Ydata})}{((b_k^{Xdata})^T \Lambda^{-1} b_k^{Xdata})} \quad (3.6)$$

where Λ is diagonal matrix with eigenvalues (λ) of the Σ , $b_k^{Ydata} / b_k^{Xdata}$ is the data trend pattern (\mathbf{Y}^{data} and \mathbf{X}^{data} , table 2) spanned into reduced EOF space.

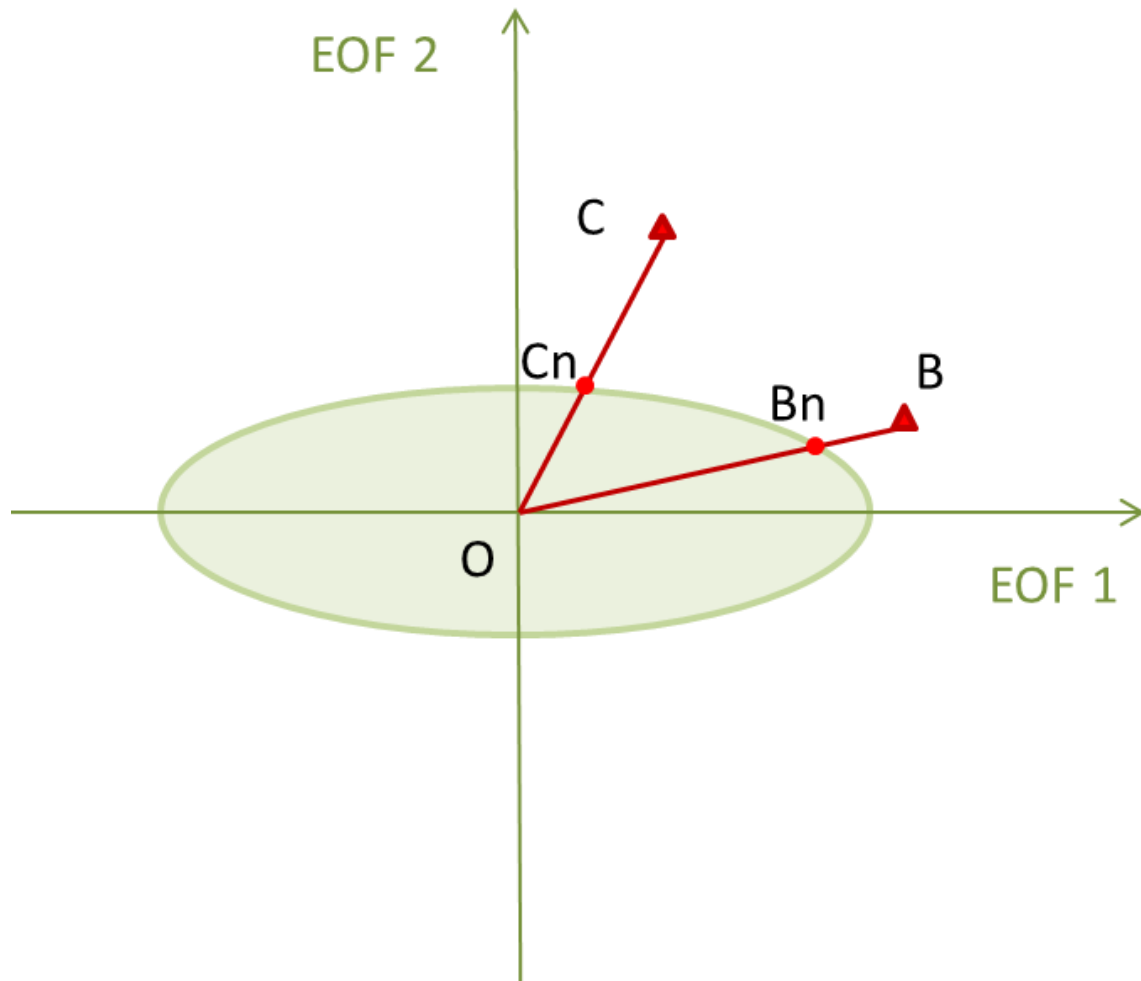


Fig. 3.4 The principle of optimal detection: finding of new direction OC which maximizes the anthropogenic signal to natural variability ratio (Mitchell et al., 2001)

Thus after transforming the signal onto two main EOFs it would lie along OB (fig. 3.4). The signal to natural variability ratio OB/OB_n would be small, because the signal lies along the direction of the main natural variability mode. The equation 3.6 finds direction OC, where the signal to natural variability ratio OC/OC_n would be maximized (Mitchell et al., 2001).

3.2 Calculation of scaling factor (β) and its uncertainty range

The main goal of optimal fingerprinting is the calculation of the scaling factor β . The β and its uncertainty range is calculated following equation 3.6. The second part of the PICTRL data - $\mathbf{Y}^{CONTROL}$ was used to calculate uncertainty ranges of the β . The individual PICTRL trends were

used instead of Y^{data} in calculation of β . As second part of PICTRL data has 165 individual trend patterns (table 2), so 165 scaling factors (β) were calculated in order to estimate the uncertainty range for β_{obs} . For the confidence level calculation Students T-test was applied to β and its uncertainty ranges. Climate change detection happens when the scaling factor (β) and its uncertainty range are positive and exclude zero. Attribution is achieved when scaling factor (β) and its uncertainty range include unity.

3.3 Data preparation

Anthropogenic GHG increase produces a distinctive pattern in the vertical profile of temperature/geopotential height which I was trying to detect. Thus for the analysis the next latitude – altitude (LxA) grid was used: latitude bands between [-49.5°S, 49.5°N] with 1° latitude interval; pressure levels - [300, 250, 200, 150, 100, 75, 50, 30 hPa]. This LxA grid was selected as the best quality of GPS RO is achieved in this region (Ho et al., 2012). In total LxA grid consists of 800 points. For each LxA grid point the temperature/geopotential height trend (K/m) was calculated using linear regression. Then to obtain temperature trends for 10-years the scaling coefficient of the regression was multiplied by 120 months. The time frame used in the study is 10 years from May 2001 till May 2011(table 2).

3.3.1 GCM

Table 1 shows the CMIP5 GCMs used in the study. A total number of 32 GCMs was used. To rescale the models on the selected latitude bands linear interpolation was used as different GCMs have different latitude resolution and pressure levels (table1). Two “business as usual” forcing scenarios were used in the study: historical scenario and RCP8.5 scenario (Moss et al., 2010; Taylor et al., 2012; table 2). Last year for the historical scenario is 2005 (few GCMs have historical scenarios till 2009), thus to be consistent with the observations it was combined with

RCP8.5 which starts in 2006 (few GCMs - 2010). The RCP8.5 scenario is considered the most realistic for the 2006-2011 time period.

Table 1 CMIP 5 GCMs used in study

Model	Institute	Resolution (Lon x Lat)	Pressure levels
ACCESS1 - 0	Commonwealth Scientific and Industrial Research Organisation and Bureau of Meteorology, Australia	192x144	17
ACCESS1 - 3		192x144	17
BCC-CSM1-1	Beijing Climate Center, China Meteorological Administration, China	128x64	17
BCC-CSM1-1(m)	Beijing Normal University, China	128x64	17
CCSM4	National Center for Atmospheric Research (NCAR), USA	288x192	17
CESM1(BGC)	National Center for Atmospheric Research (NCAR), USA	288x192	17
CESM1(CAM5)	National Center for Atmospheric Research (NCAR), USA	288x192	17
CMCC-CESM	Centro Euro-Mediterraneo per I Cambiamenti, Italy	96x48	33
CMCC-CM	Centro Euro-Mediterraneo per I Cambiamenti, Italy	480x240	17
CMCC-CMS	Centro Euro-Mediterraneo per I Cambiamenti, Italy	192x96	33
CanESM2	Canadian Centre for Climate Modelling and Analysis, Canada	128x64	22
CNRM-CM5	Centre National de Recherches Meteorologiques, Meteo-France, France	256x128	17
CSIRO-Mk3-6-0	Australian Commonwealth Scientific and Industrial Research Organization, Australia	192x96	18
FGOALS-g2	Institute of Atmospheric Physics, Chinese Academy of Sciences, China	128x60	17
FIO-ESM	The First Institute of Oceanography, SOA, China	128x64	17
GFDL-CM3	Geophysical Fluid Dynamics Laboratory, USA	144x90	23
GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory, USA	144x90	17
GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory, USA	144x90	17
GISS-E2-H	Goddard Institute for Space Studies, USA	144x90	17
GISS-E2-R	Goddard Institute for Space Studies, USA	144x90	17
HadGEM2-CC	Met Office Hadley Centre, UK	192x144	23
HadGEM2-ES	Met Office Hadley Centre, UK	192x144	17
INM-CM4	Institute for Numerical Mathematics, Russia	180x120	17
IPSL-CM5A-LR	Institut Pierre-Simon Laplace, France	96x96	17
IPSL-CM5A-MR	Institut Pierre-Simon Laplace, France	144x143	17

Table 1 cont.

IPSL-CM5B-LR	Institut Pierre-Simon Laplace, France	96x96	17
MIROC5	AORI, NIES, JAMSTEC, Japan	256x128	17
MIROC-ESM	AORI, NIES, JAMSTEC, Japan	128x64	35
MIROC-ESM-CHEM	AORI, NIES, JAMSTEC, Japan	128x64	35
MPI-ESM-LR	Max Planck Institute for Meteorology, Germany	192x96	25
MPI-ESM-MR	Max Planck Institute for Meteorology, Germany	192x96	25
MRI-CGCM3	Meteorological Research Institute, Japan	320x160	23
NorESM1-M	Norwegian Climate Centre, Norway	144x96	17

Model outputs under pre-industrial control (PICTRL) scenario were used for the calculation of natural variability without anthropogenic forcing. The CMIP5 GCM output under PICTRL scenarios provides data for several hundred years (time period differs for different GCMs). Thus for this study 110 years of data were taken from each CMIP5 GCM under PICTRL scenario. PICTRL data was calculated using the same time frame as for observation data (May of the beginning year till May + 10 years), thus 10 trend patterns from each GCM were calculated (table 2). The trends were calculated without the overlapping of the PICTRL data. The same as described above rescaling scheme was used for the calculation of the trends for the PICTRL data. As it is required by methodology (section 3.1, 3.2) the calculated PICTRL trends were separated in two parts:

1) First part (Y^{PICTRL} – table 2) was used in the optimization process. For this purpose EOF analysis was applied to this part of PICTRL trends and as a result modes of natural variability of PICTRL data were calculated. These modes of variability (EOFs) were used in the optimal fingerprinting in order to find optimal filter which would reduce the influence of natural variability on the anthropogenic signal (section 3.1).

2) The second part of PICTRL data ($Y^{CONTROL}$ – table 2) was used to calculate uncertainty range for the scaling factors (β) from optimal fingerprinting. For this purpose the individual trend patterns from $Y^{CONTROL}$ (total number - 160) were used in regression equation (equation 3.4) to calculate 160 β_{PICTRL} from PICTRL data. Then from this β_{PICTRL} distribution the uncertainty range for β were calculated.

The CMIP5 GCM data was downloaded from Earth System Grid (<http://pcmdi9.llnl.gov>) data portal.

3.3.2 GPS RO

For the study GPS RO data from two missions were used (table 2). Linear interpolation was used to rescale GPS RO data on selected pressure levels as the GPS RO data is provided for each 100 m in altitude (below 40 km). Longitude values within one latitude band were averaged for both observations and models. The GPS RO CHAMP and COSMIC Level 2 profiles were obtained from the COSMIC Data Analysis and Archive Center (<http://cdaac-www.cosmic.ucar.edu/cdaac/index.html>).

For the GCM data historical and RCP8.5 scenarios were combined to get the same time interval as GPS RO data.

As a result the next LxA trend patterns were calculated (table 2):

Table 2 Trend patterns used in study

Datasets used in study	Time Period	Number of years
GPS RO observed trend pattern		
CHAMP mission	5/2001-3/2006	11 (1 trend pattern)
COSMIC mission	5/2006-5/2011	\mathbf{Y}^{data}
CMIP5 32 GCM average trend pattern		
Average of different CMIP5 models under Historical run	5/2001-12/2005	11 (1 trend pattern)
Average of different CMIP5 models under RCP8.5 scenario	1/2006-5/2011	\mathbf{X}^{data}
PICTRL trend patterns		
PICTRL data from 32 CMIP5 models	11 years (May 1st year – May 11th year) for single trend pattern	3520 (10 trend patterns from each model) $\mathbf{Y}^{PICTRL} \mathbf{Y}^{CONTROL}$

3.4 Geopotential height concept

A geopotential height observation represents the geometrical height of the pressure surface on which the observation was taken, suggested to be the best candidate for D&A research, because it indicates the bulk atmospheric response to the anthropogenic forcing and holds more information on the dynamical structures of the atmosphere than temperature parameter (Leroy et al., 2006a). Geopotential heights are strongly related to temperature through hydrostatic equation

$$h(p) = h_s + \int_p^{p_s} \frac{RT(p')}{\mu g_0} d \ln p' \quad (3.7)$$

where $h(p)$ is geopotential height at pressure p , h_s – surface geopotential height, p_s – surface pressure, R – the ideal gas constant, $T(p')$ the temperature profile, μ – the mean air molecular mass and g_0 – WMO gravitational acceleration constant.

Geopotential heights are directly measurable from GPS RO satellites using integrated refractivity. They can be calculated using next equation.

$$p_N(h) = \frac{\mu_d g_0}{aR} \int_h^\infty N(h') dh' \quad (3.8)$$

where μ_d – molecular mass of dry air. Above 300 hPa the amount of water vapor is negligible. Assuming that the atmosphere is dry, the dry pressure can be used for calculating the geopotential heights.

The unique feature of geopotential height is that they are related to the temperature showing the bulk temperature response of the atmosphere (e.g. in case of GHG contribution – thermal expansion of the troposphere), whereas they can be calculated not relying on the information from underneath atmospheric layers (equation 3.8) which allows to use the best quality GPS RO data of the dry atmosphere.

CHAPTER IV

RESULTS

D&A research aims on using the observation and comparing the observations to the GCMs output which accounts for GHG increase. The comparison of trends from GPS RO (section 4.1.1, 4.2.1) and CMIP5 GCMs under RCP8.5 scenarios (section 4.1.2, 4.2.2) is presented first. The results from finding the main modes of natural variability and transforming (truncating) the observations and forced GCM output onto these main modes of natural variability are presented in section 4.1.3, 4.2.3. Then the truncated observations are regressed onto transformed GCMs' forced output with rotating these signals towards the direction where the natural variability has the smallest influence (section 4.1.4, 4.2.4). The following sections will describe the results of the steps described above. The analysis was done for two atmospheric parameters – temperature and geopotential height (section 4.1 and 4.2 correspondingly).

4.1 Temperature

4.1.1 GCM Temperature Trends

The individual temperature trends for each of 32 CMIP5 GCMs were calculated (fig. 4.1). The temperature trends for all the models exhibit similar patterns of UTLS temperature change. The highest warming occurs in 300 hPa, but the latitude bands differ for different models – some models show major warming region in Southern Hemisphere mid-latitudes about -40°S , other GCMs showing the warming in the equator. Another region with high warming trends occurs about 150 hPa and around 35°N . The warming trends in major warming regions vary from 0.9 K/decade (GFDL-ESM2M) to 1.8 K/decade (ACCSESS 1-3). All the models agree on cooling

trends in tropical Lower Stratosphere (LS). The amplitude of these trends varies from -1.8 K/decade (GISS-E2H) to -0.3 K/decade (FIO-ESM). Thus all the models tend to show general warming of the upper troposphere and cooling of the lower troposphere. This well-known pattern (Santer et al., 2013) can be observed in fig. 4.2 with the average of 32 CMIP5 GCMs' temperature trends (X , equation 3.1).

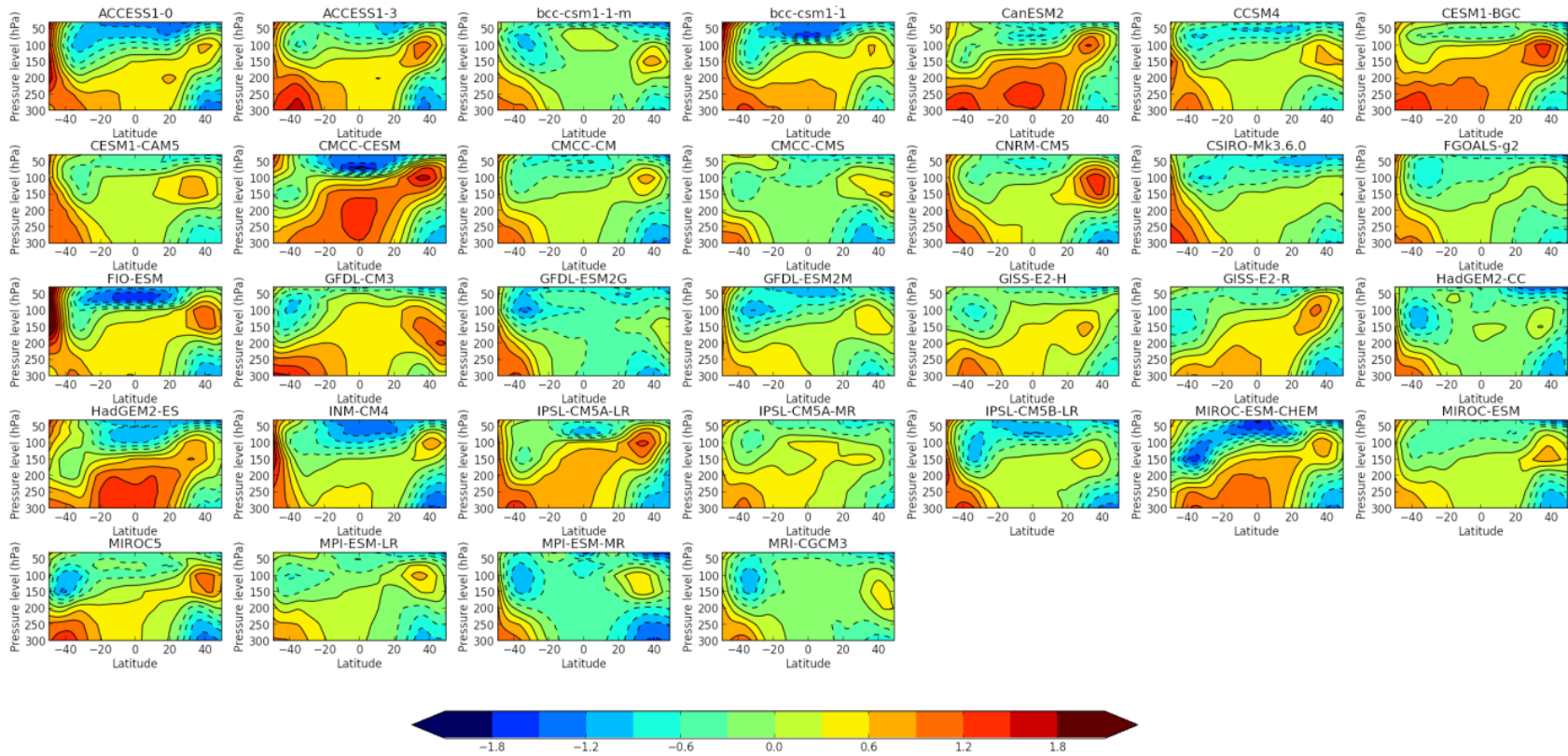


Fig. 4.1 Zonal upper air temperature trends (K/decade) for 32 CMIP5 GCMs. The trends were computed fitting 11 years of GCM temperature data (2001-2011) with linear regression, (table 2). The study region is within -50°S , 50°N and 300 and 30 hPa (section 3.3).

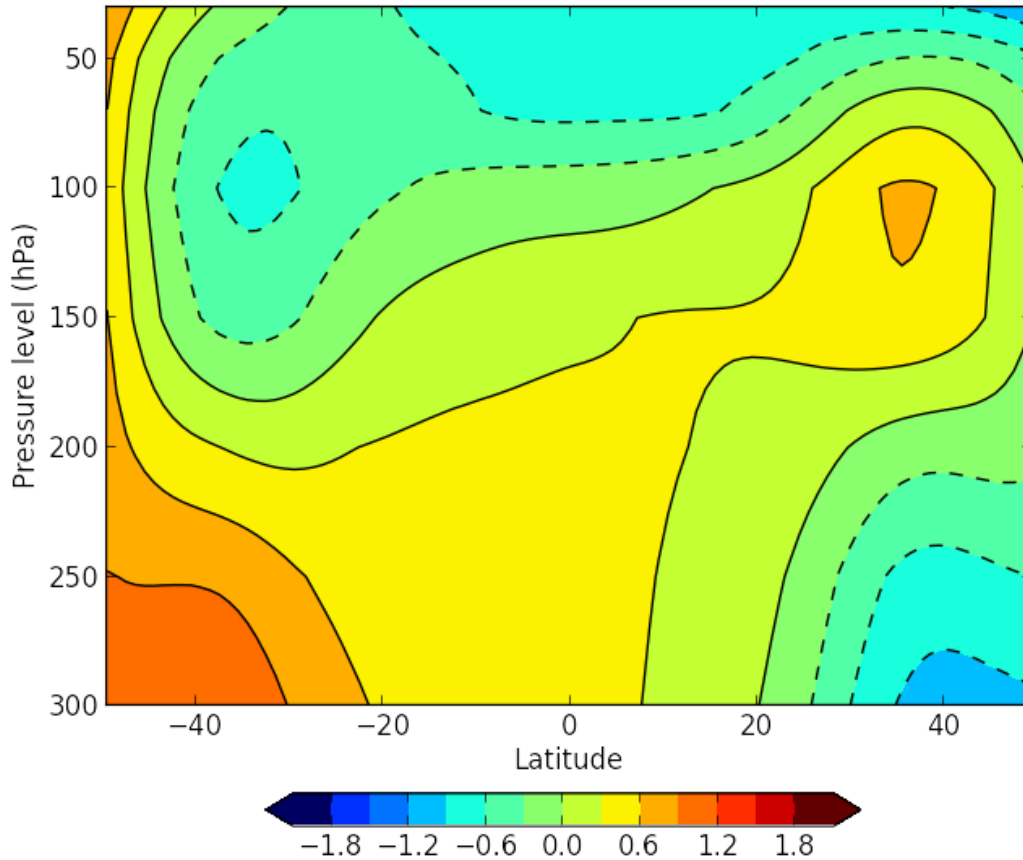


Fig. 4.2 The average of temperature trends of the 32 CMIP5 GCMs presented in fig.4.1. The units are K/decade. The study region is within -50°S , 50°N and 300 and 30 hPa (section 3.3).

4.1.2 Radio occultation temperature trends

The GPS RO temperature trends (Y , equation 3.1) computed using eleven years of data are presented in fig. 4.3. In general, the RO temperature trends show the pattern, similar to GCM-generated. There are two major warming regions which are similar to the GCMs generated trend patterns (fig. 4.1, 4.2). The correlation coefficient between GPS RO trends and average of 32 CMIP5 GCMs is 0.44. In addition GPS RO temperature pattern shows the warming in tropical region around 70 hPa. The lowest correlation with observations shows FIO-ESM GCM – 0.17; the highest correlation is – 0.6 for MIROC5 GCM. The correlation of 32 GCM average with GPS RO data is ~ 0.42 (fig. 4.4).

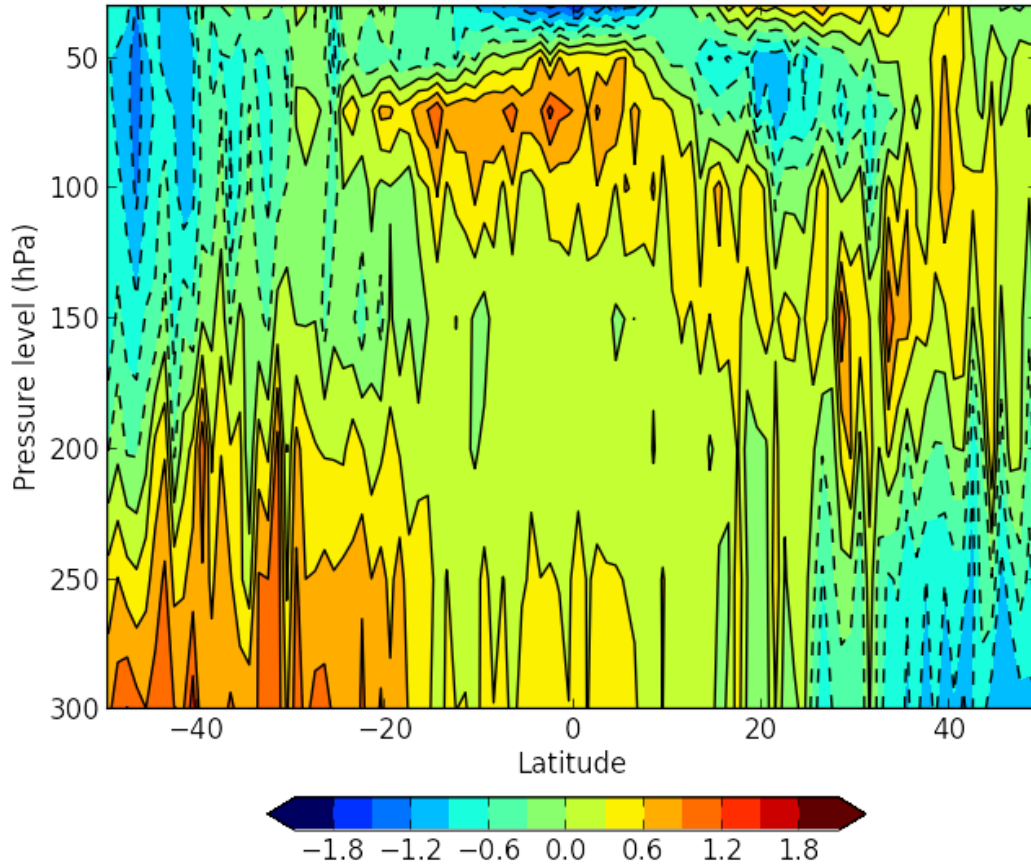


Fig. 4.3 Zonal upper air temperature trends for GPS RO data. The trends were computed using 11 years of data (2001-2011) by linear regression, using two GPS RO missions (table2). The units are K/decade. The study region is within -50°S , 50°N and 300 and 30 hPa (section 3.3).

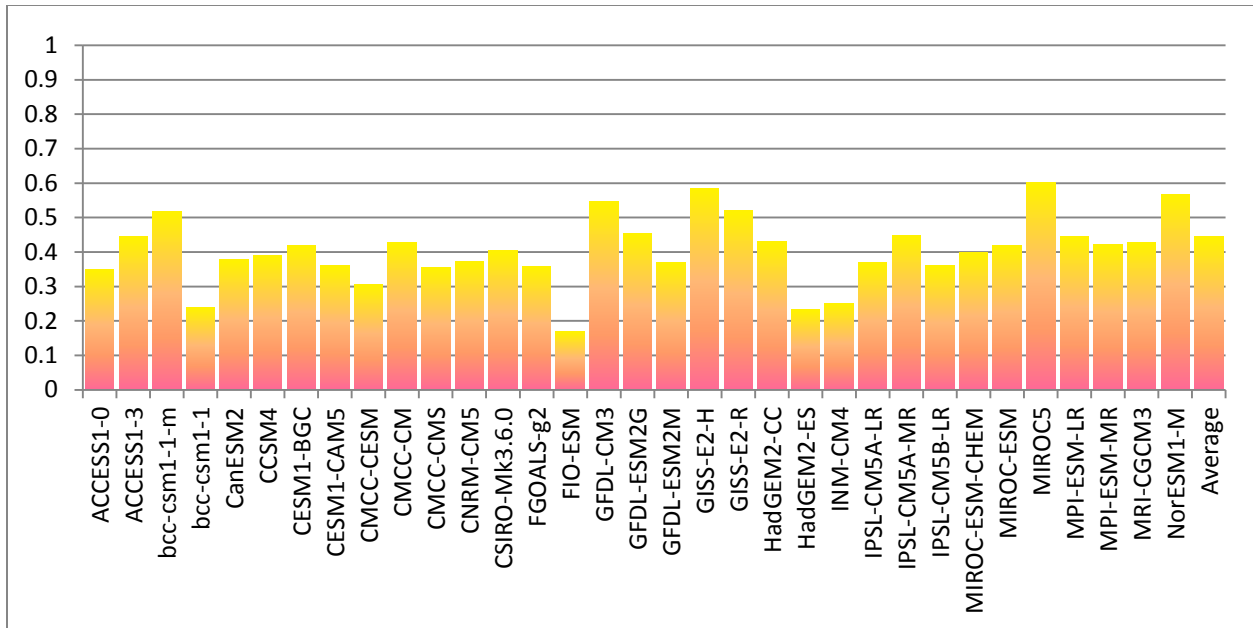


Fig. 4.4 Correlation coefficients between GPS RO temperature trend pattern (fig. 4.3) and CMIP5 GCMs' temperature trend patterns (fig. 4.1, 4.2)

4.1.3 EOF truncation

Each EOF accounts for the different mode of variability within the data. First EOF describes the fraction of total variance which shows the maximum variability within the PICTRL temperature data (fig. 4.5). This main mode of variability shows the pattern of tropical UT warming which is consistent with results by Leroy et al. (2006) and was identified as El Niño–Southern Oscillation (ENSO) mode. The second and third EOFs show subsequent main modes of variability. Clearly these modes are asymmetric and can represent the consequences of Southern Annular Mode (SAM) and Northern Annular Mode (NAM) activity correspondingly. Identifying other EOFs is a difficult task because the latitude altitude grid used to calculate trend patterns is unconventional.

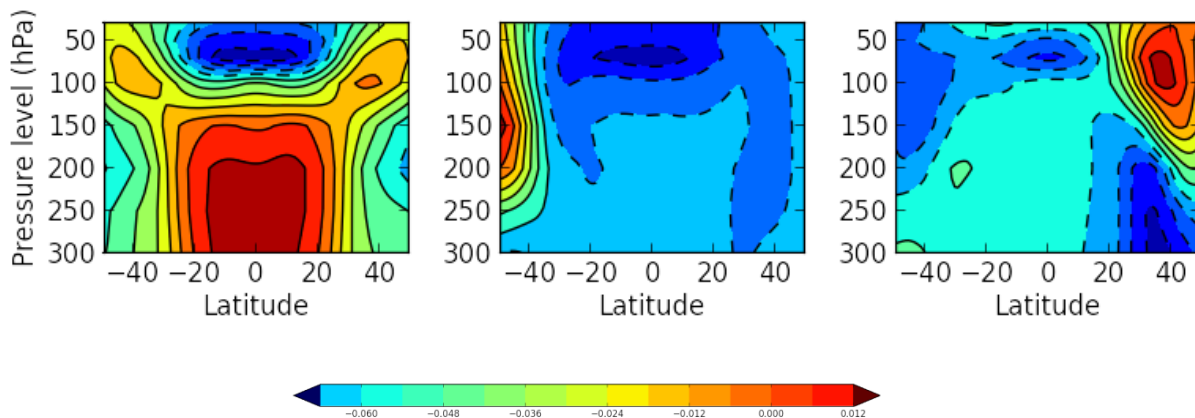


Fig. 4.5 First three EOFs of the CMIP5 32 GCM temperature PICTRL data. The units are %.

As it was described in section 3.1 the optimization is done through finding the direction of smaller natural variability's influence (as shown in fig. 3.4), by projecting the signal (Y or X) onto the main modes of natural variability (EOFs). The finding of these main modes requires computing EOFs from PICTRL data and then projecting the signal onto the main EOFs,

truncating the signal in the EOFs reduced space (when not all the EOFs are used). In fig. 4.6 the results of reconstructed from EOF space trend patterns are presented using different amounts of EOFs –the more EOFs are used, the bigger fraction of variance of original trend pattern is retained (fig. 4.6, fig. 4.7a). The first EOF describes ~ 43% of variance and first three EOFs describe ~ 65% of total variance (fig. 4.7a). The amount of explained variance reaches 90% retaining 10 EOFs. Thus the trend pattern becomes closer to the original with each subsequent EOF. It is also confirmed by the correlation coefficients between the original pattern and EOF-dimension-reduced pattern the (fig. 4.7 b, c). The correlation increases with using each subsequent EOF. The EOF-reduced GCM temperature patterns achieve the correlation of 0.8 with the original pattern retaining 10 EOFs; while EOF-reduced GPS RO temperature patterns become close to correlation of 0.8 with original pattern only retaining 12 EOFs. Observed data correlation coefficient increases much faster comparing to GCM data for the first five EOFs, however the correlation coefficient for 1 retained EOF in RO data is much lower than for GCM data (~0.15 and ~0.59 correspondingly).

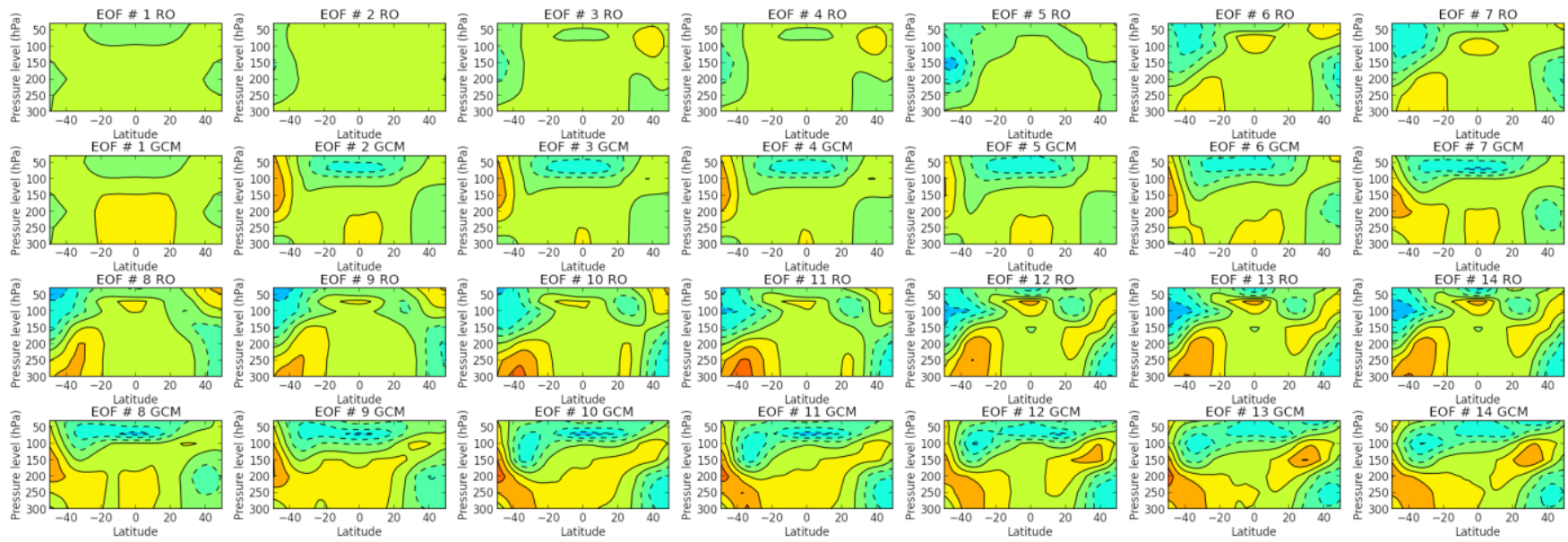


Fig. 4.6 Rebuilt temperature pattern using 1-14 EOFs for GPS RO data (row – 1,3) and CMIP5 GCM data (row – 2,4)

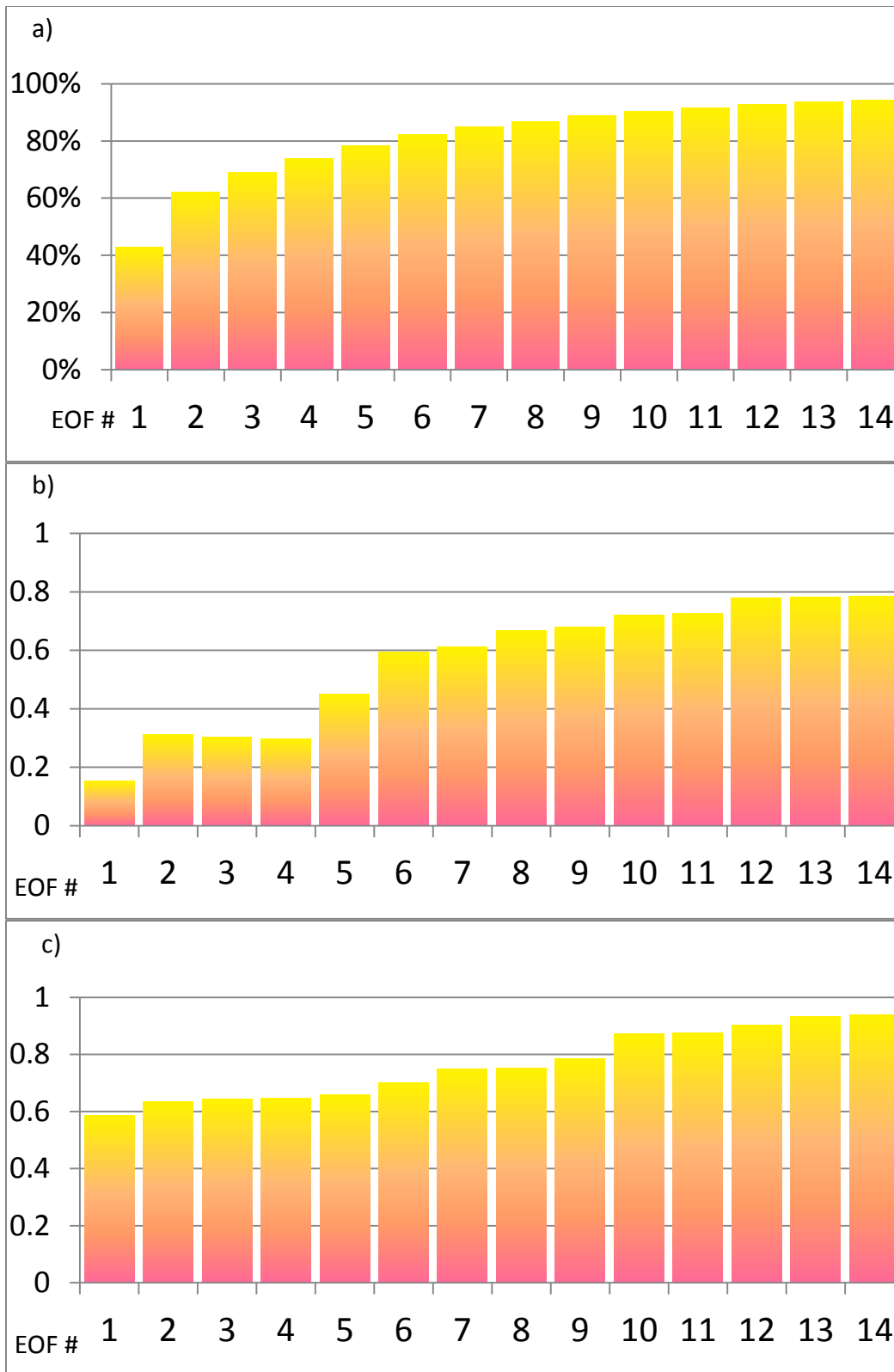


Fig. 4.7 a) The total amount of variance explained by 1- 14 EOFs; b) pattern correlation between the original and rebuilt GPS RO pattern for 1- 14 EOFs; c) pattern correlation between the original and rebuilt GCM pattern for 1- 14 EOFs.

4.1.4 Detection and Attribution Results

The optimal fingerprinting technique aims on finding if the GCMs and observational patterns can be explained by natural variability which is the testing of null hypothesis described in section 3. The patterns are optimized by natural variability and are used in a regression model before hypothesis testing. The resulted scaling factors from the regression and its uncertainty ranges show values (β) varying from ~ -0.1 to ~ 0.6 (fig. 4.8). The scaling factors show how well GCMs under anthropogenic forcing scenario predict the observations and whether these patterns can be explained by natural variability or not. Such results testify that GCMs are consistently overestimating the observational pattern. The uncertainty ranges for the scaling factors always include 0 and exclude 1 for retaining 1 to 11 EOFs. This means that for this number of EOFs the anthropogenic forcing does not have a detectable influence on the temperature pattern. The uncertainty ranges of scaling factors exclude zero, retaining 12 to 14 EOFs, which shows that these results are significantly different from natural variability and implies a detection of climate change. The level of statistical significance for the detection is 5 %. However, the GCMs are overestimating the observations by a factor of 2 ($\beta \sim 0.5$). The unity is not included into 5-95% uncertainty range of scaling factor for EOF 12-14. That means that anthropogenic forcing is not necessarily the reason for the observed changes of temperature in UTLs region for 2001-2011 time period.

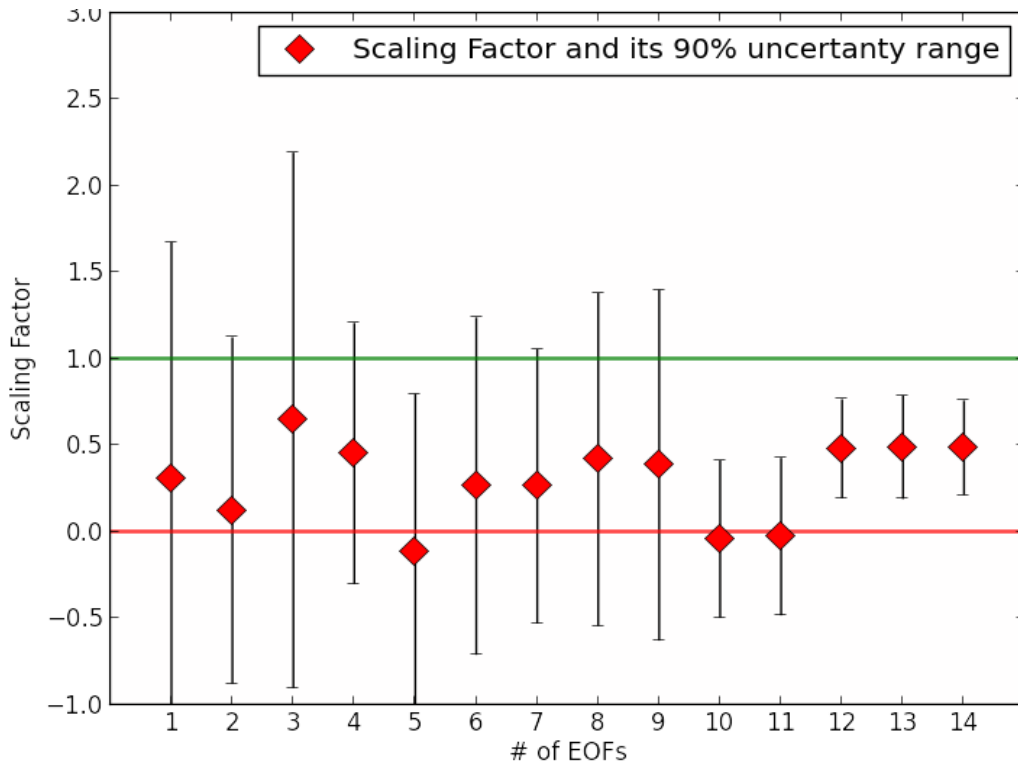


Fig. 4.8 Results for the uncertainty assessment in the scaling factors (red rhombus) for 1-14 retained EOFs, calculated using temperature data. Error bars indicate the 5% to 95% uncertainty range based on the PICTRL scaling factors.

4.2 Geopotential Height

4.2.1 GCM geopotential height trends

The individual trends for each of 32 CMIP5 GCMs were calculated for geopotential height parameter for UTLS region within -50°S , 50°N (fig. 4.9). The results of computed trends show an increase in geopotential height in Southern Hemisphere. Geopotential height trends for all the models exhibit similar patterns for UTLS. The GCMs show increase from 40 to more than 100 meters/decade mainly in the mid latitudes in the Southern Hemisphere. The Northern Hemisphere shows decreasing geopotential height trends with increasing the amplitude of these trends moving towards higher latitudes. The decreasing trend values vary from -40 to -80 m/decade around 40°N . The more similar results between the GCMs for geopotential height than temperature can be explained that geopotential height is less affected by natural variability, showing more robust trends (fig. 4.9 - 4.14). The highest warming occurs in Southern Hemisphere which can be explained by the strong warming trends in Antarctic region (Bromwich et al., 2012). The majority of the models tend to show thermal expansion of Hadley cell in upper troposphere. This well-studied pattern (Santer et al., 2013) is similarly achieved for the average of 32 CMIP5 GCMs' geopotential height trends (fig. 4.10).

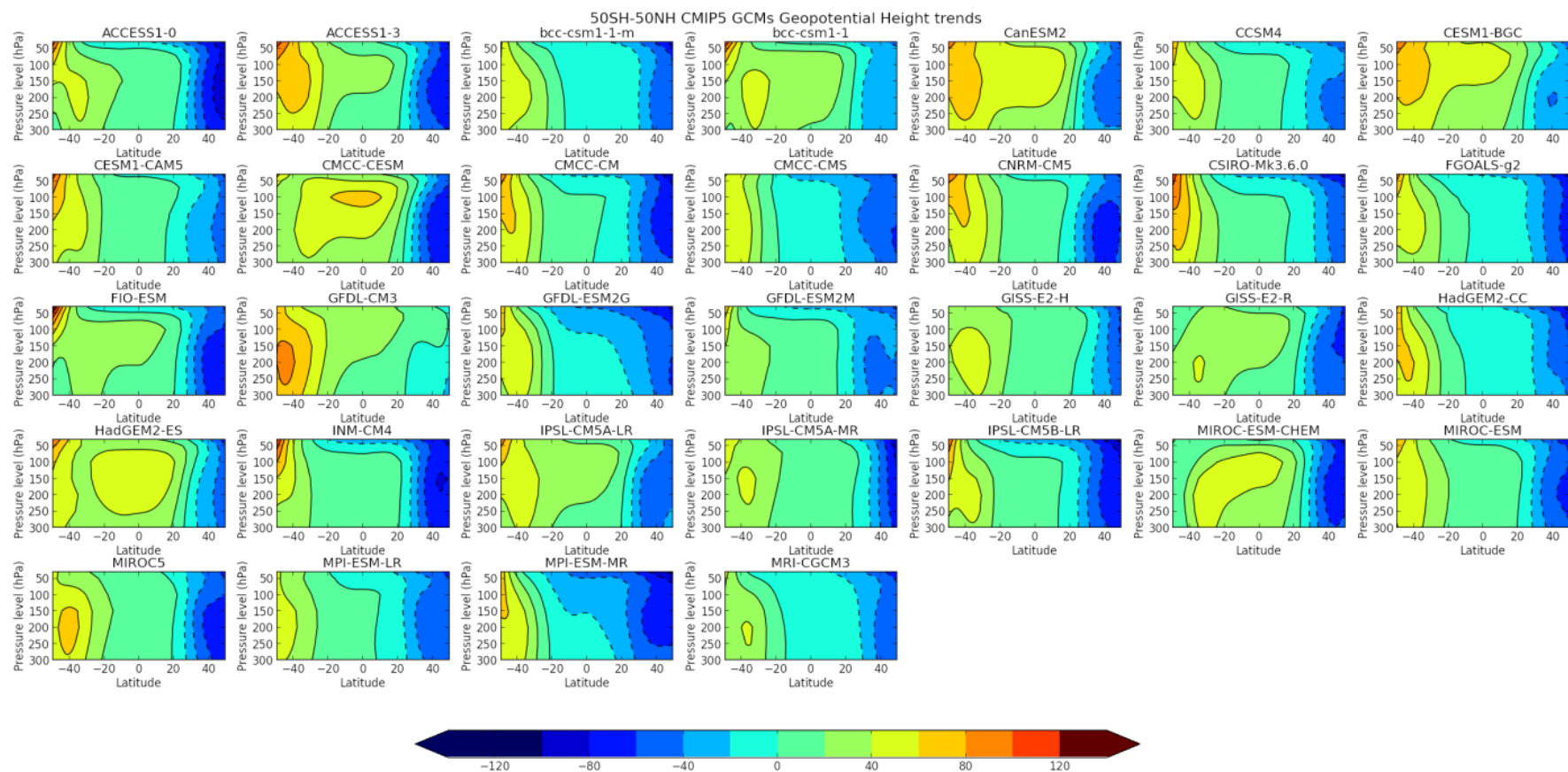


Fig. 4.9 Zonal upper air geopotential height trends for 32 CMIP5 GCMs. The trends were computed using 11 years of data (2001-2011) by linear regression, using two scenarios – historical and RCP8.5 (table 2). The units are m/decade. The study region is within -50°S , 50°N and 300 and 30 hPa (section 3.3).

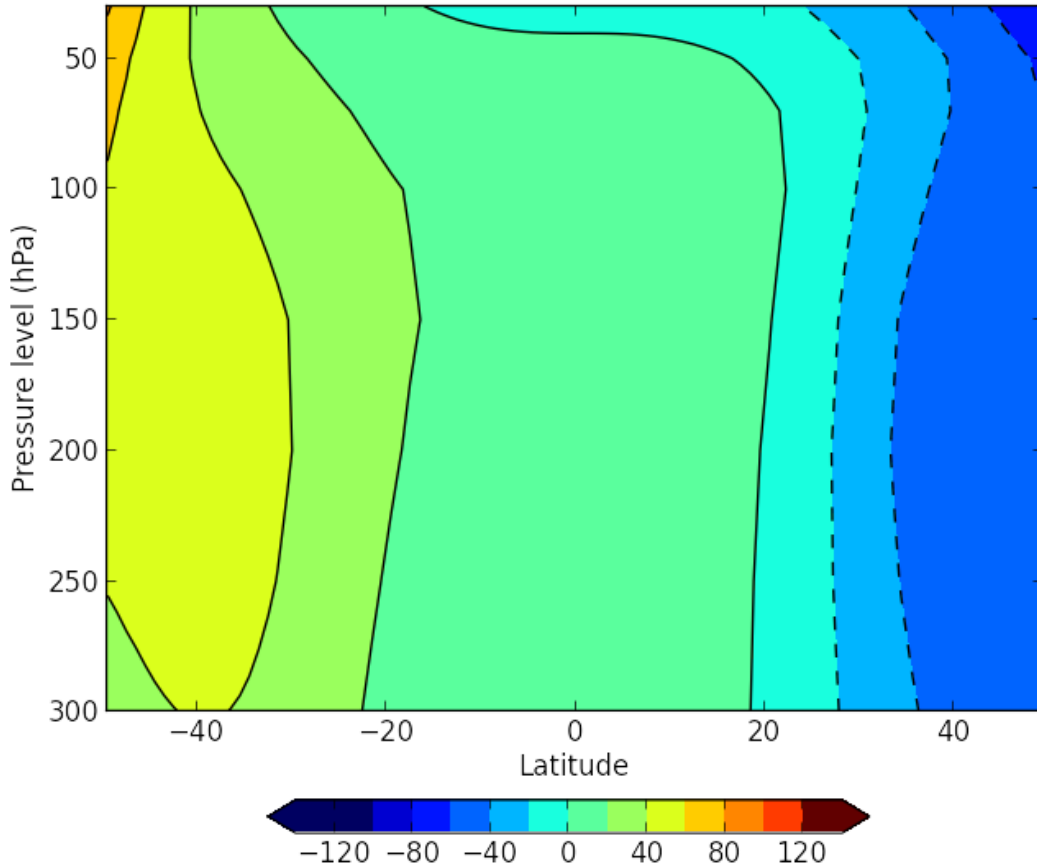


Fig. 4.10 The average of geopotential height trends of the 32 CMIP5 GCMs presented in fig.4.8. The units are m/decade. The study region is within -50°S , 50°N and 300 and 30 hPa (section 3.3).

4.2.2 Radio occultation geopotential height trends

The geopotential height trends computed using eleven years of GPS RO (fig. 4.11) show very similar results to GCM-generated patterns. The warming trends can be observed in the southern hemisphere (fig. 4.1, 4.2, 4.3), leading to thermal expansion and an increase in geopotential heights. The correlation coefficient between GPS RO geopotential height trends and average of 32 CMIP5 GCMs under RCP8.5 and historical runs is 0.92. The correlation coefficients for each particular GCM with GPS RO data (fig. 4.12) vary from the lowest one which is 0.76 and is

calculated between MIROC-ESM-CHEM GCM and GPS RO data; and the highest one – 0.95 for MIROC5 GCM and GPS RO data.

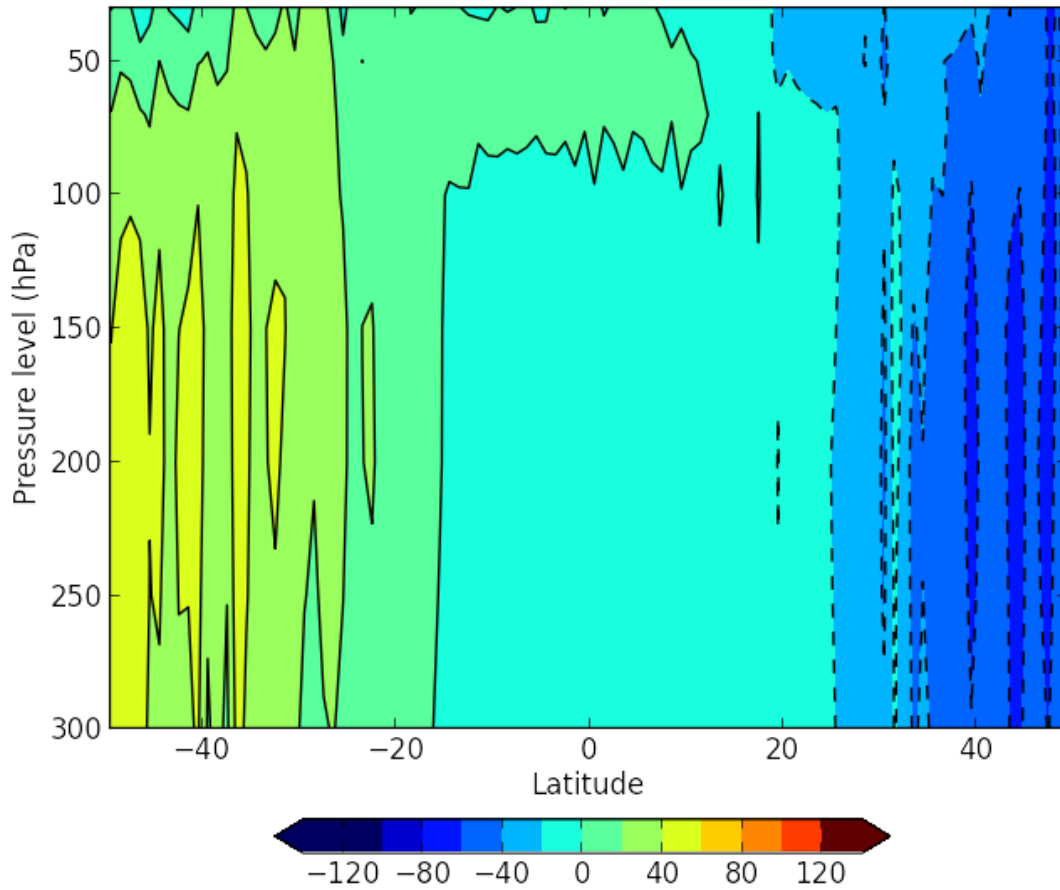


Fig. 4.11 Zonal upper geopotential height trends for GPS RO data. The trends were computed using 11 years of data (2001-2011) by linear regression, using two GPS RO missions (table 2). The units are m/decade. The study region is within -50°S , 50°N and 300 and 30 hPa (section 3.3).

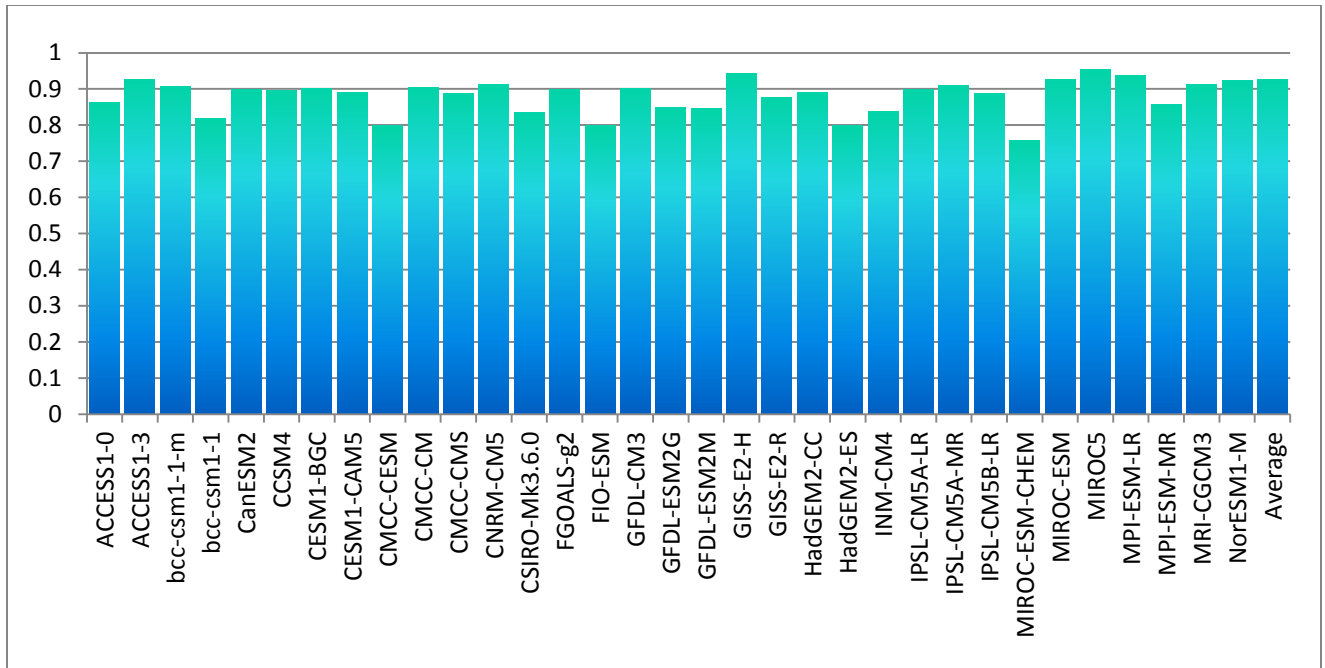


Fig. 4.12 Correlation coefficients between GPS RO geopotential height trend pattern (fig.4.10) and CMIP5 GCMs' geopotential height trend patterns (fig.4.8, 4.9).

4.2.3 EOF truncation

First three EOFs describing three main modes of variability within the PICTRL geopotential height data well defined natural variability patterns (fig. 4.13). The main mode of variability shows the pattern of tropical UT shrinking which is can related to the activity of negative phases of ENSO mode. The second and third EOFs show the second main modes of variability which demonstrates the poleward jet migration in the Southern Hemisphere, which is consistent with SAM positive phases. The third EOF is asymmetric to the second EOF and produces pattern associated with NAM positive phase, showing poleward jet migration in the Northern Hemisphere.

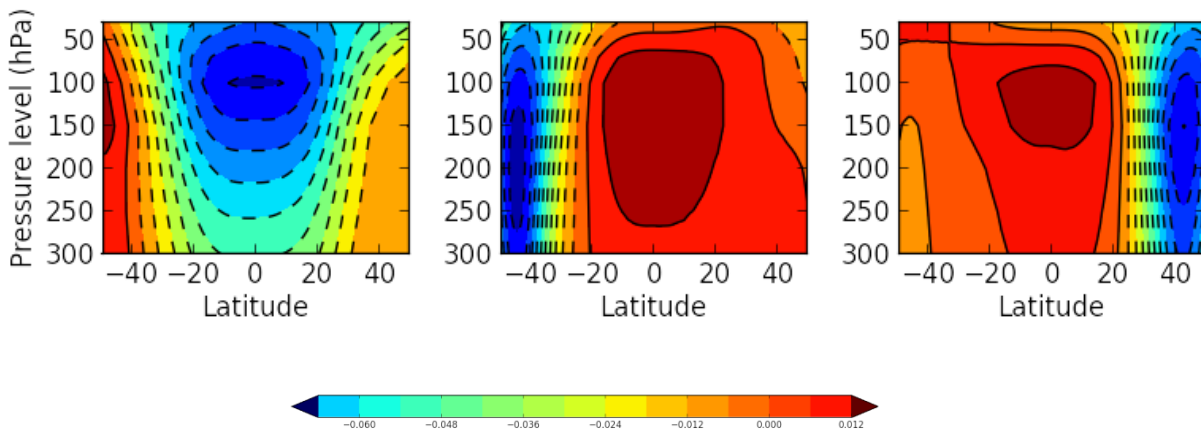


Fig. 4.13 First three EOFs of the CMIP5 32 GCM geopotential height PICTRL data. The units are %.

The geopotential height trend patterns reconstructed from EOFs space (fig. 4.14) show similar to temperature results – with the use of higher amount of EOFs in truncation the truncated pattern becomes more and more closer to the original one. It is also confirmed by the amount of variance explained by different number of EOFs used in truncation (fig. 4.15). The first EOF is responsible for the highest amount of variance and explains 67% of total variance. Comparing to

temperature there is less variability in geopotential height trends as the amount of explained variance reaches 90% retaining 4 EOFs and the first 3 EOFs are responsible for 87% amount of variance (fig. 4.15), while for temperature the EOFs reach 90% retaining variance in first 10 EOFs and first 3 EOFs explain only 69% of variance (fig. 4.6). Thus the EOF reduced geopotential height pattern becomes similar to the original ones even retaining only 3 EOFs (fig. 4.12) for both modeled and observed cases. GCM truncated pattern has a correlation coefficient of 0.96 with original one (fig. 4.13), and GPS RO trend pattern has correlation of 0.93 with the original GPS RO geopotential height pattern retaining only 3 EOFs (fig. 4.13) while the correlation for temperature patterns showed much lower increase with using higher amount of EOFs.

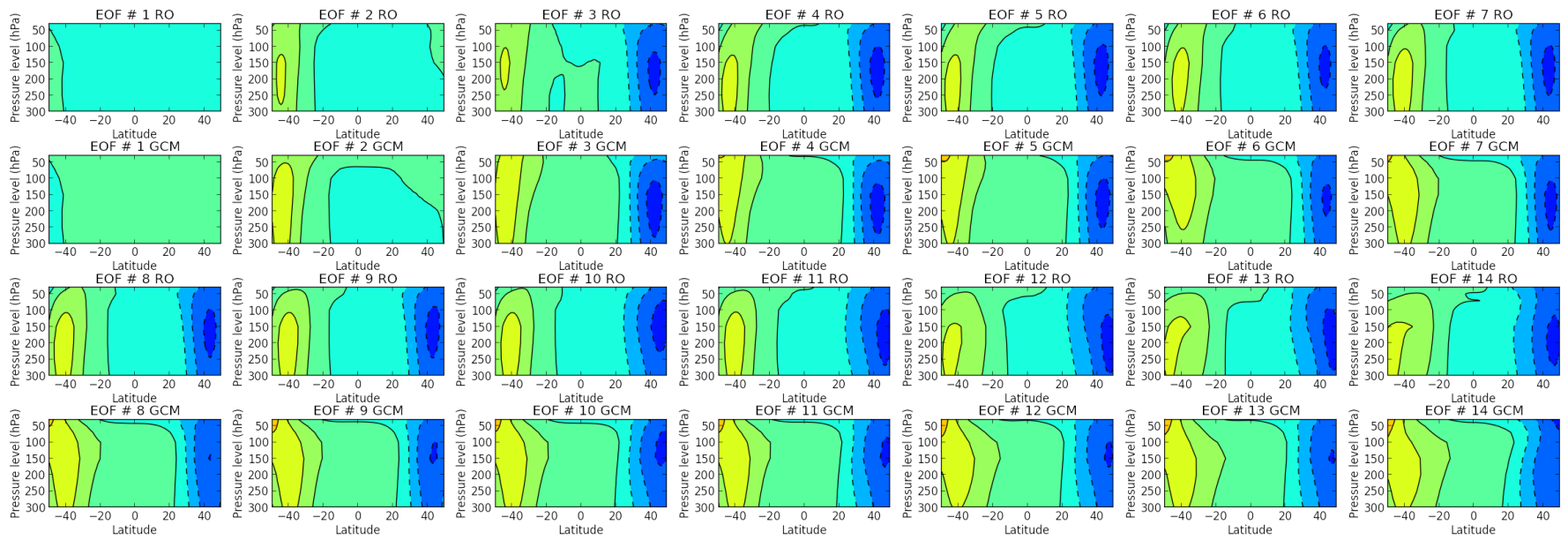


Fig. 4.14 Rebuilt geopotential height pattern using 1-14 EOFs for GPS RO data (row – 1,3) and CMIP5 GCM data (row – 2,4).

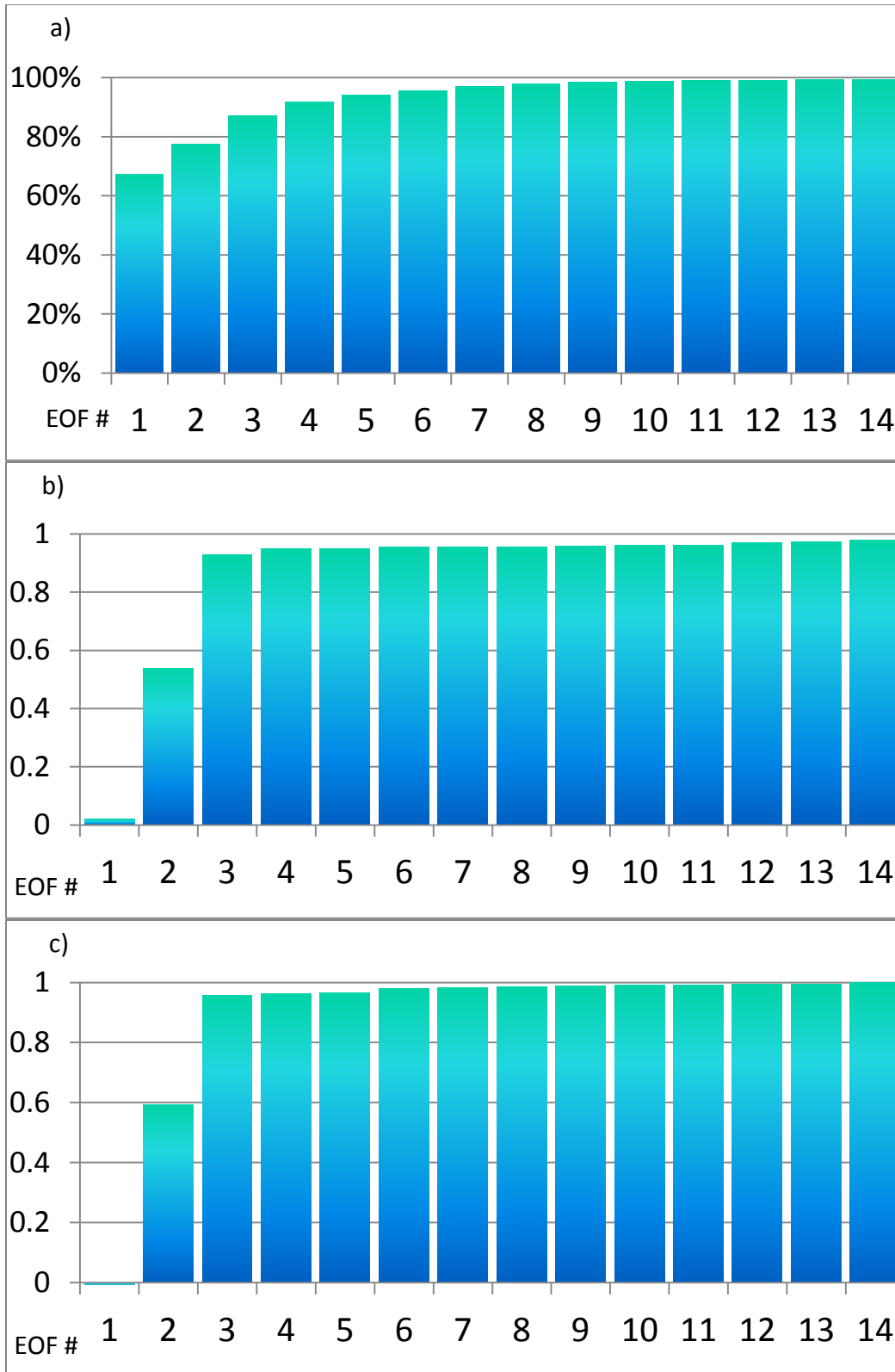


Fig. 4.15 a) The total amount of variance explained by 1- 14 EOFs; b) pattern correlation between the original and rebuilt GPS RO pattern for 1- 14 EOFs; c) pattern correlation between the original and rebuilt GCM pattern for 1- 14 EOFs.

4.2.4 Detection and Attribution Results

Optimal fingerprinting results for geopotential heights are very different from results achieved for temperature trend patterns. The scaling factors and its uncertainty ranges (fig. 4.16) show that β values are very robust to the selection of a number of EOFs used in the truncation. For the majority of EOFs the scaling factors are close to unity (except first two EOFs and EOF # 10). That implies that GCMs under RCP8.5 scenario are well predicting the observations. The uncertainty ranges for these EOFs' scaling factors always include unity and exclude 0 which testifies that the anthropogenic forcing has a detectable influence on the geopotential height pattern and these results are significantly different from natural variability. In case of geopotential heights I used 99% uncertainty range (+- 3 standard deviations). Thus the D&A is achieved on very high statistical significance level (P value ~ 0.005%).

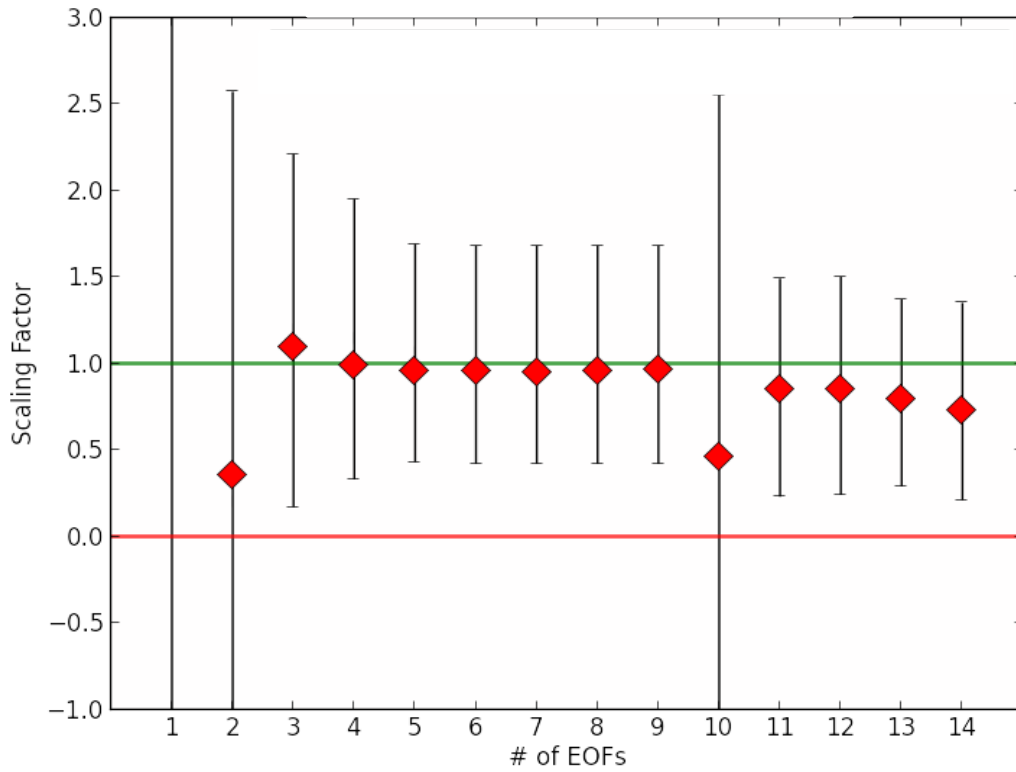


Fig. 4.16 Results for the uncertainty assessment in the scaling factors (red rhombus) for 1-14 retained EOFs calculated using geopotential height data. Error bars indicate the 0.05% to 99.95% uncertainty range based on the PICTRL scaling factors.

CHAPTER V

DISCUSSION

5.1 Temperature trend patterns

The observed and modeled temperature trend patterns (fig. 4.3 and fig. 4.1) show some discrepancies even if the main pattern is captured by the models ($r \sim 0.44$, fig. 4.4). Temperature trends between GCMs and GPS RO showed the highest disagreement in the stratosphere. The main difference between GCM and GPS RO trend patterns arises in LS around 70 hPa around the equator, where observations show distinct warming pattern, while models show the cooling of the LS. One of the causes of such mismatch can be linked to the Quasi-Biennial Oscillation (QBO) activity. The QBO is the downward migration of the equatorial stratospheric winds, which phases are associated with changing direction of winds from westerlies to easterlies (Baldwin et al., 2001). It has a large influence on the stratospheric temperature (Randel et al., 1994). This gap between GCMs and observations can be due to fact that the QBO is underestimated by the CMIP5 GCMs (Charlton-Perez et al., 2013). The time series of GPS RO temperatures in 70hPa at the equator and the time series of the QBO index (<http://www.cpc.ncep.noaa.gov/data/indices/>) show quite similar behavior (fig. 5.1). The correlation coefficient is 0.32, testifying that this warming pattern can be related to the influence of the QBO. Different time lags did not yield stronger correlation results (table 3).

Another contradictory region is in the mid-latitudes in Northern Hemisphere, where observations show overall warming LS trends, while the GCM ensemble shows cooling trends. It can be linked to the weak sudden stratosphere warming (SSW) simulation. SSW is caused due to the

weakening of the polar vortex (Matsuno, 1971). Charlton-Perez et al. (2013) showed that CMIP5 GCMs twice underestimate observed SSW events from reanalysis data. The SSW events, in turn, are strongly affected by the QBO (Gray et al., 2004) which also adds to SSW underestimation by GCMs.

Finally, another GCM – GPS RO discrepancy region is mid-latitudes in Southern Hemisphere. Here GCMs show LS warming while GPS RO shows distinctive cooling. This lack in representation of the LS trends is demonstrating that GCM ensemble fails to provide a realistic representation of the temperature trends on the decadal scale. These results are confirmed by Charlton-Perez et al. (2013). The authors concluded that the CMIP5 GCMs have weak representation of decadal stratosphere variability, which is in focus of this research, whereas the historical longer trends of stratospheric dynamics are well depicted by the CMIP5 GCMs. Also GCMs have better agreement in simulations of the mean stratospheric climate and trends, but have different results in simulating stratospheric variability (Hardiman et al., 2012). This shows that the time period of this study is rather short and is strongly affected by natural variability of the climate which is not well represented by GCMs in the stratosphere, which in turn yields significant disparateness in the temperature trend patterns between GCM ensemble and GPS RO data.

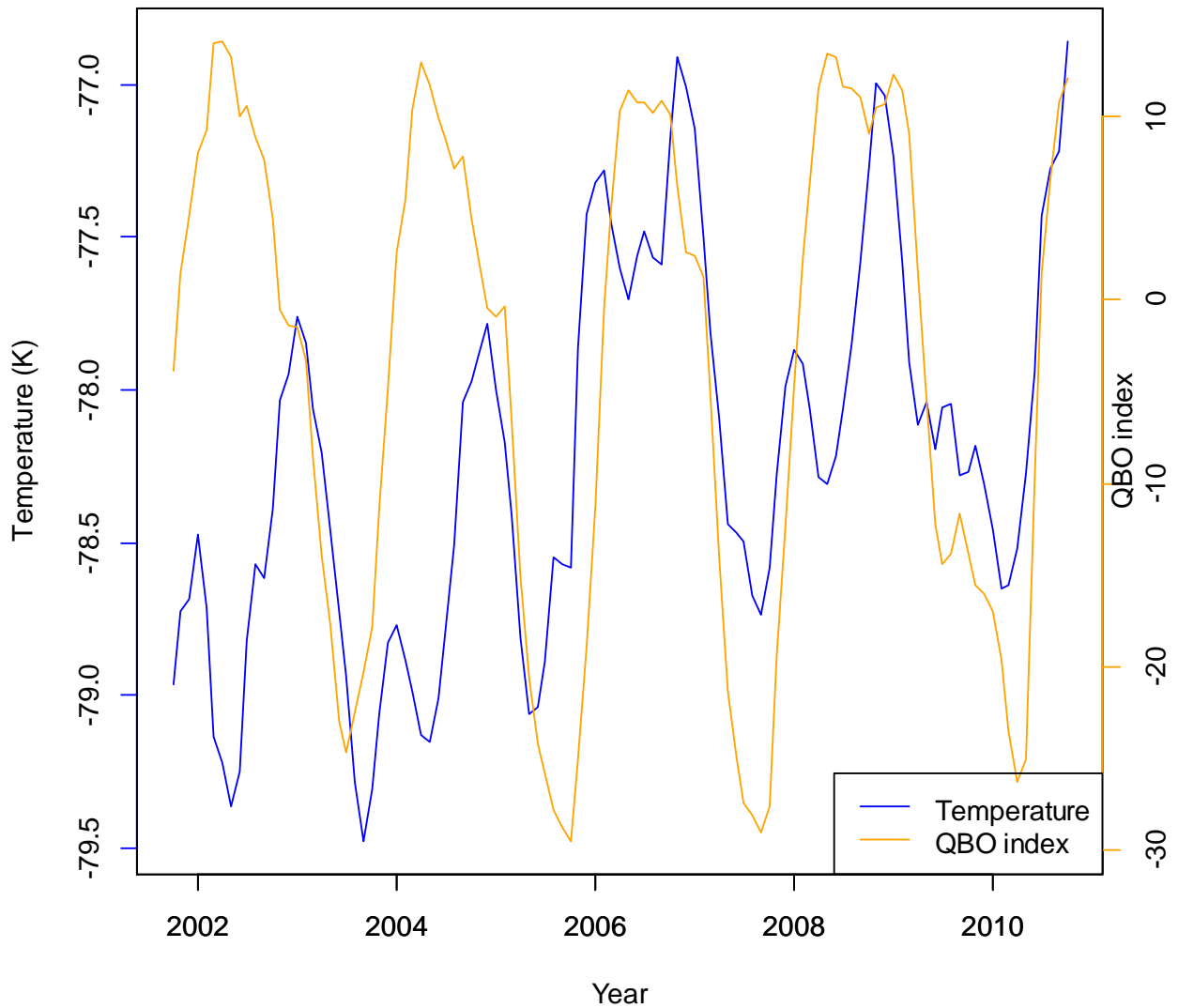


Fig. 5.1 Time series of GPS RO temperature in 1.5°S, 70 hPa LxA point (blue line) and the QBO index (orange line). The temperature time-series was smoothed using 13-month running-average filter.

All 32 CMIP5 GCMs mimic the temperature pattern calculated from GPS RO data in UT. The difference between the models arises in the magnitude of the simulated pattern. Both GPS RO and CMIP5 GCMs' patterns show warming of the troposphere in southern hemisphere, which

coincides with studies by Bromwich et al. (2012), who found the dramatic temperature increase in the Antarctic region. Also GCMs and observations agree on cooling trends in troposphere around 40°N, which is contrary to the overall tropospheric warming (Leroy et al., 2006a), which is expected to evolve under anthropogenic GHG forcing. One possible explanation of such cooling trend can be the 2011 la Nina phase. However time series of 300 hPa and 40°N point with the Oceanic Niño Index (fig.5.2) (Oceanic Niño Index, http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml) show no relation with the correlation coefficient of -0.14. The time lag of 3 months showed the highest negative correlation of -0.3 (table 3). Thus the explanation of negative temperature trends in this region cannot be explained with 2011 la Nina phase.

Table 3 Correlation between temperature from two selected LxA points with ENSO and QBO indexes correspondingly, with different time lags

Monthly lag	0	1	2	3	4	5	6
Correlation between GPS RO temperature in 40°N, 300hPa LxA point and ENSO index							
	-0.14	-0.23	-0.28	-0.30	-0.29	-0.25	-0.21
Correlation between GPS RO temperature in 1.5°S, 70 hPa LxA point and the QBO index							
	0.32	0.31	0.27	0.18	0.05	-0.07	-0.19

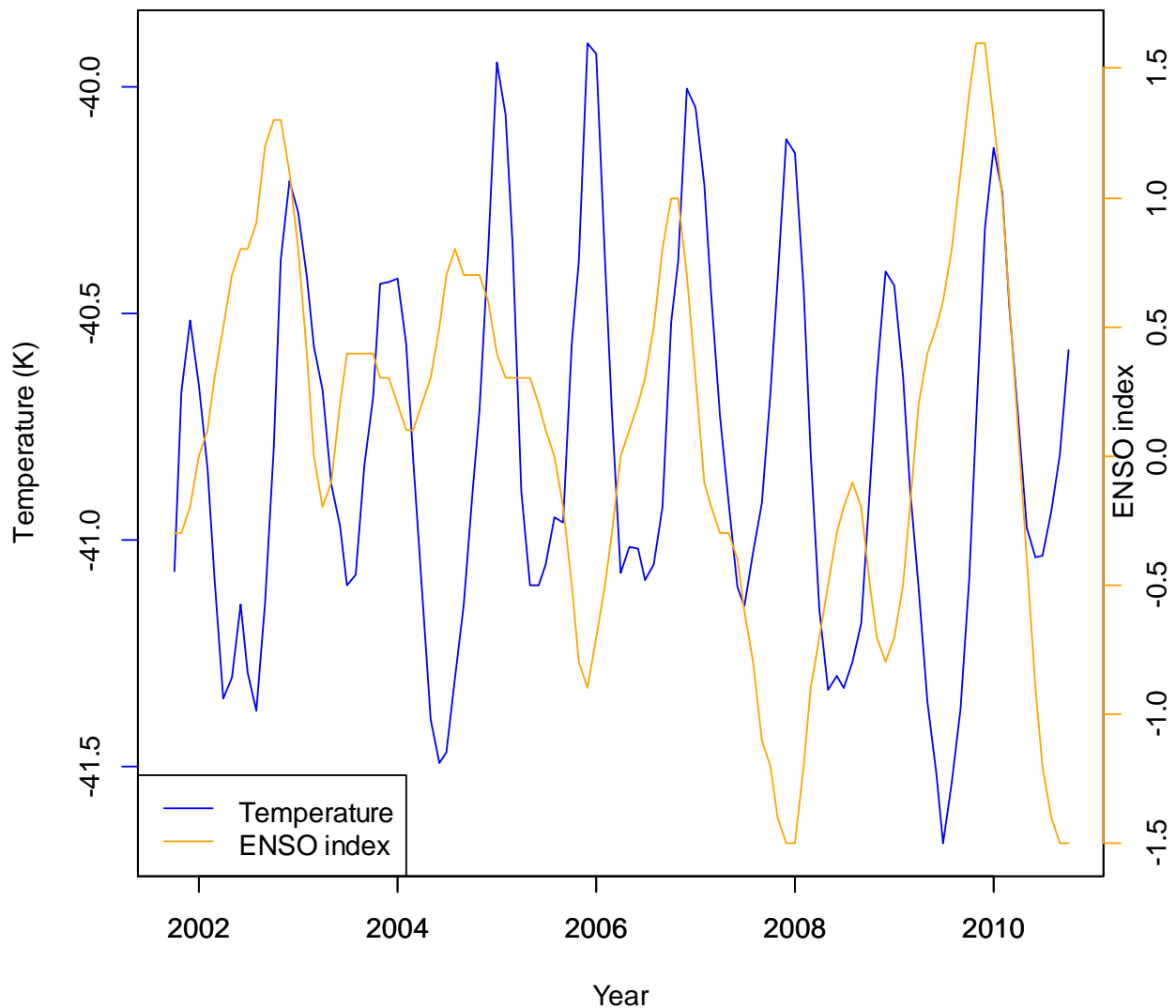


Fig. 5.2 Time series of GPS RO temperature in 40°N, 300hPa LxA point (blue line) and the ENSO index (orange line). The temperature time-series was smoothed using 13-month running-average filter.

It is noticeable that not all the models agree between each other. Some models show a major warming region centered at 40°S which is consistent with Antarctic warming; other GCMs produce a warming trends at equator, showing the thermal expansion of Hadley cell (fig. 4.1).

The latter pattern is also well described in the scientific literature and also can be monitored by observing the tropopause height in the tropics. The recent studies confirm the Hadley cell expansion, showing the increasing trends in tropical tropopause height (Steiner et al., 2011). This discrepancy in the models between warming trends has to be studied in future research.

The temperature trend patterns are more consistent between GCMs if longer time intervals are used for trend computation. The trends computed using 50 years of GISS-E2-H GCM under RCP8.5 scenario (fig. 5.3) develop very robust pattern of the tropospheric warming and stratospheric cooling in the tropics. The similar temperature trends were found by Leroy et al. (2006a). The authors computed trends for 12 CMIP3 GCMs, which were very consistent with this trend pattern (fig. 5.3). The similar well defined temperature pattern was found by Santer et al. (2013) in 34 years of MSU data. So it can be assumed that other models might agree much better between the temperature trends if longer time periods are taken into account. This will reduce the impact of natural climate variability and will yield trends which show much robust tropical thermal expansion and stratospheric cooling, leading to much better agreement between different GCMs (fig. 5.3; Leroy et. al., 2006a). This is also the case for GPS RO data because the available data record is not long enough to produce well defined UTLS temperature pattern.

Leroy et al. (2006a) discussed that surface temperature trends are not the best candidate for D&A research because there is a big uncertainty in many regions. Current study highlights the uncertainty for UTLS temperature trend patterns. The poor stratospheric temperature variability representation by the GCMs, their mismatch with observations and high influence of natural variability affects optimal fingerprinting results (section 4.1.4). The impact of natural variability in temperature trends, which are calculated for such time period, is very high, yielding complex

trend pattern in the study region. Thus the temperature parameter is not the best fit for the D&A purposes, especially on such time scale.

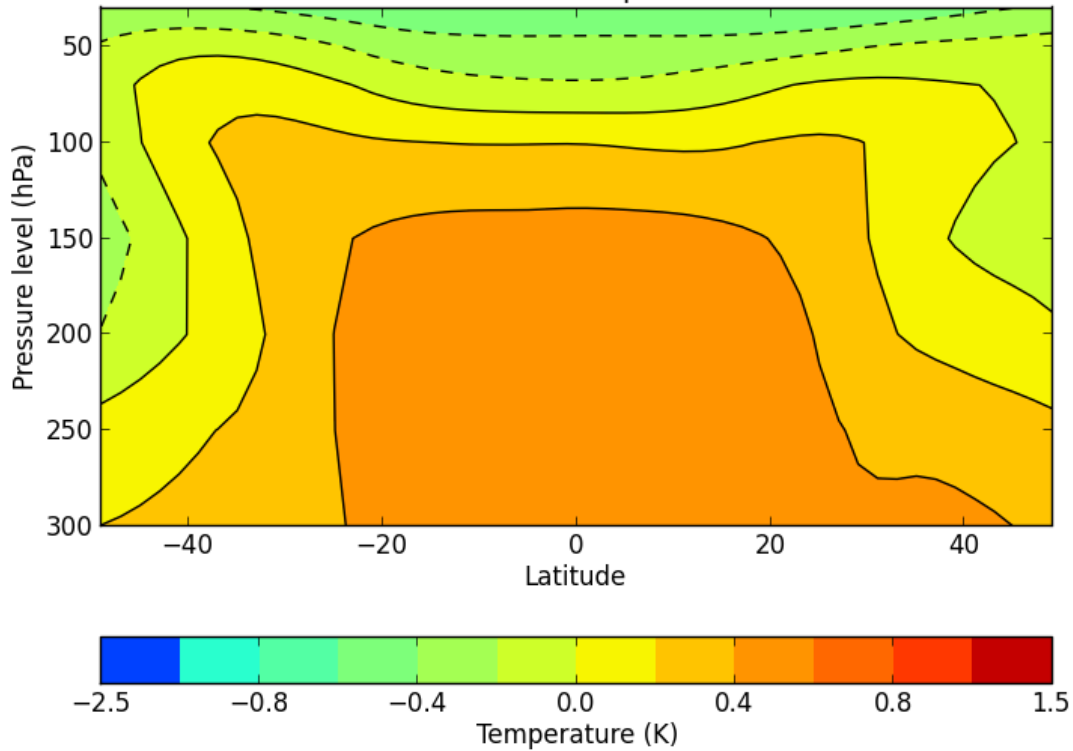


Fig. 5.3 Zonal UTLS temperature trends for GISS-E2-H CMIP5 GCMs. The trends were computed using 45 years of data (2006-2050) by linear regression, using RCP8.5 scenario. The units are K/decade. The study region is within -50°S , 50°N and 300 and 30 hPa (section 3.3).

5.2 Geopotential height trend patterns

There is much better agreement between the CMIP5 GCMs and GPS RO geopotential height trend patterns. Moreover individual GCM simulations (fig. 4.9) show more consistent trend patterns between each other, comparing with individual GCM temperature patterns (fig. 4.2).

This analysis demonstrated that temperature is more affected by natural climate variability than geopotential heights. It is confirmed by looking at the amount of total variance explained by

different number of EOFs (fig. 4.7, 4.15). Much less amount of EOFs is needed to make the truncated pattern very close to the original one for geopotential height parameter, comparing to temperature (fig. 4.15a vs fig. 4.7a). Also the correlation between optimized and original geopotential height patterns (fig. 4.7bc, 4.15bc) increases at a much higher rate. The same three EOFs already give very high correlation (~ 0.9) between optimized and original pattern, which demonstrates less variability in geopotential height trends comparing to temperature trends. Another confirmation of the stability of geopotential height trends to climate natural variations is high correlation between GPS RO and GCM geopotential height trend patterns (fig. 4.4) for using only 11 years of data. This shows that GCMs much better resolve geopotential height trends because there is less influence of natural climate variability. Thus geopotential heights can be a better candidate for the D&A research.

5.3 Detection and attribution of climate change using GPS RO data

The D&A of climate change can be done if the observed trend patterns cannot be explained just by the natural variability of the climate. According to the methodology (section 3.1) it happens when scaling factors and its uncertainty ranges exclude 0 and include 1.

In this study the GPS RO temperature trends did not show the anthropogenic signal predicted by CMIP5 GCMs. It can be due to rather short record of GPS RO data and/or due to that the temperature trends are significantly affected by natural variability of the climate system. The scaling factors show distinct overestimation of the temperature pattern by the models. It is interesting that previous research by Lackner et al. (2011), where authors applied optimal fingerprinting technique to the GPS RO data and CMIP3 GCMs data, demonstrated a detectable anthropogenic signal on 95% confidence level. The data record in Lackner et al. (2011) was shorter than in this study, thus probably the anthropogenic signal was smoothed by strong natural

variability events in the recent data records (e.g. fig. 5.1, 5.2 – the QBO and ENSO activity). Also the authors used only 40 points for trend computation, while in this research 800 points were used. Thus the trend patterns in Lackner et al. (2011) are more generalized and as an example do not show strong warming pattern in tropical LS region in GPS RO records (fig. 4.3). Geopotential height data shows very different from temperature results. The detection of anthropogenic signal is very robust in GPS RO geopotential height pattern. Furthermore the uncertainty range of scaling factors was calculated using ± 3 std. dev. which allowed achieving more than 99% confidence level. It is important that the scaling factors are close to unity which also shows good GCMs skill in simulating geopotential height pattern. Lackner et al. (2011) were also using geopotential height in their research. However, even though the authors' geopotential height scaling factors were similar to that I have found in this study (around unity), the uncertainty range of scaling factors showed the significance of 90% (± 1.3 std. dev. for 90% uncertainty range). Lackner et al. (2011) results were less significant than in this research. Such improvement in statistical significance of the anthropogenic signal detection can be explained that in this study I were using higher resolution, which is provided by new set of CMIP5 GCMs; the number of GCMs used here is 32 comparing to 3 GMS in Lackner et al. (2011); the resolution in this study is 800 LxA points comparing to 40 in Lackner et al. (2011). All these differences can yield better representation of natural variability which leads to more efficient optimization of the signal to noise ratio (fig. 3.1).

5.4 Future research

The evolving atmospheric trends are very affected by natural variability, when the time period is not long enough. It was discussed in previous sections that the temperature trends, for both GCMs and observations in UTLS, form a very complex pattern different from the expected tropical UT warming and LS cooling. The geopotential height trends also did not show this pattern. Thus an optimal fingerprinting technique which reduces the impact of natural variability on evolving atmospheric trends has to be used. However with obtaining longer records of GPS RO data it would be possible to detect anthropogenic signal in the GPS RO data without applying optimization filter. It will give an advantage for more robust and easy explainable results, because in optimization the specific amount of EOFs which is a subjective decision has to be chosen.

Another way to strengthen the detection of anthropogenic influence is to artificially maximize the impact of natural variability on the anthropogenic signal. If the human influence on atmospheric trends can be detected relative to additionally increased natural variability, it will give more statistical confidence in the D&A of climate change. However it will require long term record of observational data.

The research by Santer et al. (2013) was aimed on using remotely sensed atmospheric data in D&A study with maximization of natural variability. The authors were using 34 years of MSU temperature data in order to detect anthropogenic signal. The authors were using D&A methodology without optimization of the signal to noise ratio but with the larger additional natural variability. The temperature trends were analyzed and compared with CMIP5 GCM data. As a result the authors were able to demonstrate a detectable human influence on the recent temperature trends in the free atmosphere.

Leroy et al. (2006a) research highlighted that the temperature trends are not the best fit for the detection of the anthropogenic signal. The much better candidate for such purpose is geopotential height parameter because it indicates the bulk atmospheric response to the anthropogenic forcing (Leroy et al., 2006a). Moreover the geopotential heights can be directly measured by the GPS RO (Leroy, 1997). Previous research by Leroy et al. (2006), Lackner et al. (2011) and a current one give a good presentation of the usefulness and robustness of geopotential height trends. Even though the time scale is quite small for GPS RO data comparing to the MSU data, the geopotential height trends show strong potential for the detection of anthropogenic influence without optimization of the signal, but with the increase of the total amount of internal variability. Thus the future research can be aimed on the use of the methodology developed by Santer et al. (2013).

The future improvements in GCM resolution will provide opportunity for more precise D&A research. It will require higher spatial coverage by GPS RO satellites. The increase of GPS RO satellite numbers (such as it is proposed by GeoOptics <http://geooptics.com/>) will also allow testing higher resolution GCMs and RCMs.

CHAPTER VI

CONCLUSIONS

The optimal fingerprinting technique was applied in order to detect anthropogenic climate signal in GPS RO geopotential height and temperature data using CMIP5 GCMs. The detection is achieved by optimizing the temperature trend patterns using 12-14 EOFs with 90% significance level. The attribution of the climate change isn't achieved for temperature trend patterns. The scaling factors are showing values ~ 0.65 and lower, which shows overestimation of trends by GCMs. Geopotential height trend patterns show very significant results for D&A of climate change. The 99% significance level is achieved for geopotential height data. The scaling factors are very close to unity for retaining 3 to 9 EOFs, which demonstrates very good representation of geopotential height trends by CMIP5 GCMs.

APPENDIX A

GPS RO python scripts

Script 1.1 First script creates a list with GPS RO profiles for each year, indicating date and longitude/latitude of each GPS RO profile.

```
import numpy as N
import pylab as pl
import netCDF4
import os
import sys
from netCDF4 import Dataset
#Set parameters
dayStart=1
dayEnd=366
yearStart=2006
yearEnd=2006
dataSet=0 #0 for CHAMP, 1 for COSMIC
dataSets=['CHAMP','COSMIC']
Path=['D:\\CHAMP\\', 'D:\\DATA_COSMIC_WETPRF\\'] #path to COSMIC and CHAMP data
PathOut='D:\\Occultation\\WET_AIRS\\'

#open output file
str_fo=PathOut+dataSets[dataSet]+str(yearStart)+".txt"
fo=open(str_fo,'w')
fo.write("pathFull,file,year,start_month,start_day,start_hour,start_Latitude,end_Latitude,start_Lo
ngitude,rootgroup.nd_Longitude\n")

for year in range(yearStart,yearEnd+1): #start year cycle
    print(str(year))
    if dataSet==0: #'CHAMP':
        for day in range(dayStart,dayEnd+1): #number of the day
            pathFull=Path[dataSet]+str(year).zfill(4)+". "+str(day).zfill(3)+"\\"
            if os.path.lexists(pathFull): #try to open the folder with data for this year, this date
                files = os.listdir(pathFull) #get a list of files in this folder

                for file_ in files: #go file by file
                    try:
                        rootgroup=Dataset(pathFull+file_) #set the netCDF object

fo.write(pathFull+" "+str(file_)+", "+str(year)+", "+str(rootgroup.month)+", "+str(rootgroup.day)+
```

```

", "+str(rootgroup.hour)+", "+str(rootgroup.lat)+", "+str(rootgroup.lat)+", "+str(rootgroup.lon)+", "+
str(rootgroup.lon)+"\n')
        rootgroup.close()
    except:
        continue

elif dataSet==1: #'COSMIC':
    for day in range(dayStart,dayEnd+1): #number of the day
        pathFull=Path[dataSet]+str(year).zfill(4)+". "+str(day).zfill(3)+"\'
        if os.path.lexists(pathFull): #try to open the folder with data for this year, this date
            files = os.listdir(pathFull) #get a list of files in this folder

            for file_ in files: #go file by file
                try:
                    rootgroup=Dataset(pathFull+file_) #set the netCDF object

fo.write(pathFull+", "+str(file_)+", "+str(year)+", "+str(rootgroup.month)+", "+str(rootgroup.day)+
", "+str(rootgroup.hour)+", "+str(rootgroup.lat)+", "+str(rootgroup.lat)+", "+str(rootgroup.lon)+", "+
str(rootgroup.lon)+"\n')
                    rootgroup.close()
                except:
                    continue
fo.close()
#here are parametrs which are used for comparison

#for netcdf files
#lat
#lon
#month
#day
#hour
#minute

```

Script 1.2 This script is used to deal with a problem within GPS RO *netcdf* files. The problem is that missing values are written as a string instead of float numbers. This script rewrites missing values as floats.

```

import numpy as N
import netCDF4
import os
import sys
import glob

from netCDF4 import Dataset

Path2='D:\\gropotential height\\temp2\\'

```

```

c=0
a=0
for infile in glob.glob( os.path.join(Path2,'*txt') ):
    fi=open(infile,'r')

    Lines=fi.readlines()
    fi.close

    a+=1
    print a
    for line in Lines:
        #print Lines
        c+=1
        words=line.split(',')
        #print words[0]+words[1]
        try:
            #files=words[0]+words[1]
            fin=Dataset(words[0]+words[1],'r+')
            Vp = fin.variables['MSL_alt'] # for cosmic variables

            x=Vp.missing_value
            Vp.delncattr('missing_value')
            Vp.setncattr('missing_value',float(x))

            Vp1 = fin.variables['Temp'] # for cosmic variables

            x1=Vp1.missing_value
            Vp1.delncattr('missing_value')
            Vp1.setncattr('missing_value',float(x1))

            Vp2 = fin.variables['Pres'] # for cosmic variables

            x2=Vp2.missing_value
            Vp2.delncattr('missing_value')
            Vp2.setncattr('missing_value',float(x2))
            fin.close()
            #print 'bingo'
        except:
            #print "Warning: no missing_value attribute"+"\n"+str(infile)
            continue

print 'Done'

```

Script 1.3 Next script is sorting GPS profiles for each latitude band and for each month

```
import numpy as N
import os
import sys
import glob

Path='G:\\gropotential height\\list1\\'# directory of list with GPS RO profiles

lat=[]
a=-89.5
while (a <= 90):#creating latitude coordinates each 1 degree
    lat.append(a)
    a+=1
print len(lat)

cosmicData=[]
row=0
for infile in glob.glob( os.path.join(Path,'*txt') ):
    #open input files
    fi=open(infile,'r')
    print infile
    cosmicLines=fi.readlines()
    fi.close

    c=0L
    for line in cosmicLines:
        c+=1L
        if (c==1):continue
        row+=1
        words=line.split(',')

        for i in range(2,6):
            words[i]=int(words[i])
        for i in range(6,len(words)):
            words[i]=float(words[i])
        cosmicData.append(words)

for year in range(2005,2006):
    for month in range(1,13):

        for x in range(len(lat)):
            print lat[x]
            str_fo='G:\\gropotential height\\latitudes\\'+str(year)+'\\'+str(month)+'\\'+str(lat[x])+'.txt'

            dHour = 1 #delta hour
```

```

#open output file
fo=open(str_fo,'w')

fo.write("pathFull,file,year,start_month,start_day,start_hour,start_Latitude,end_Latitude,start_Longitude,rootgroup.nd_Longitude\n")

for ic in range(0,row):
    #print int(cosmicData[ic][2]), int(year)
    #print int(cosmicData[ic][3]), int(month)
    #print bingo
    if not(int(cosmicData[ic][2])==int(year)):continue
    if not(int(cosmicData[ic][3])==int(month)):continue
    if not(float(lat[x])-0.5<=cosmicData[ic][6]<=float(lat[x])+0.5): continue #latitude for
tropics
    #print 'bingo'
    for i in range(len(cosmicData[ic])):
        if i==len(cosmicData[ic])-1:
            fo.write(str(cosmicData[ic][i])+"\n")
        else:
            fo.write(str(cosmicData[ic][i])+',')
    #print(str(ic))
fo.close()

```

Script 1.4 The next script calculates monthly averages of pressure and temperature fields for using lists sorted for each latitude band

```

import numpy as N
import netCDF4
import xlwt
import os
import sys
import glob
#from netCDF4 import Dataset
from scipy.io import netcdf

Path='G:\\gropotential height\\latitudes\\'

lat=[]
a=-89.5
while (a <= 90):#creating latitude coordinates each 1 degree
    lat.append(a)
    a+=1
print len(lat)

for year in range(2011,2012):

```

```

book = xlwt.Workbook()
book1 = xlwt.Workbook()
book2 = xlwt.Workbook()

for month in range(1,13):
    print month, "MONTH"
    Temperature=[[0 for x in range(400)] for x2 in range(len(lat))]
    matrix=[[0 for x1 in range(400)] for x22 in range(len(lat))]
    Pressure=[[0 for x in range(400)] for x2 in range(len(lat))]
    matrix1=[[0 for x1 in range(400)] for x22 in range(len(lat))]
    Altitude=[[0 for x in range(400)] for x2 in range(len(lat))]
    matrix2=[[0 for x1 in range(400)] for x22 in range(len(lat))]
    col=0
    sheetname=str(month).zfill(2)
    sheet1 = book.add_sheet(str(sheetname),cell_overwrite_ok=True)
    sheetname1=str(month).zfill(2)
    sheet11 = book1.add_sheet(str(sheetname1),cell_overwrite_ok=True)
    sheetname2=str(month).zfill(2)
    sheet12 = book2.add_sheet(str(sheetname2),cell_overwrite_ok=True)
    for x in range(len(lat)):
        #for x in range(4,5):
            print lat[x]
            infile=Path+str(year)+'\\'+str(month)+'\\'+str(lat[x])+'.txt'
            f=open(infile, 'rU')
            lines=f.readlines()
            f.close()
            #print infile
            a1=0
            for line in lines:
                a1+=1
                if a1==1:continue
                #print a1
                words=line.split(',')
                try:
                    fin=netcdf.netcdf_file('G'+words[0][1:]+words[1],r')
                    T = fin.variables['Temp']
                    Alt = fin.variables['MSL_alt']
                    Pres = fin.variables['Pres']
                    b=0
                    #if str(Alt[1]).rstrip()=='0.1':
                        #print '1111'
                    if str(Alt[1]).rstrip()!='0.1':
                        print str(Alt[1]).rstrip()
                        b+=1
                        print 'bingo'
                        if str(Alt[1+b]).rstrip()!='0.1':

```



```

print str(Alt[1+b]).rstrip()
b+=1
print 'bingo'
if str(Alt[1+b]).rstrip()!='0.1':
    print str(Alt[1+b]).rstrip()
    b+=1
    print 'bingo'
if str(Alt[1+b]).rstrip()!='0.1':
    print str(Alt[1+b]).rstrip()
    b+=1
    print 'bingo'
#print b
for z in range (400-b):
    if str(T[z+b])==str(['--']) or str(T[z+b])==str('--') or int(T[z+b])==-999:pass
    #if :print 'AAAAAAAAAAAAA'
    else:
        #try:
            #print len(T)
            #print z
            Temperature[x][z]+=float(T[z+b])
            matrix[x][z]+=1
        #except:
            #pass
if str(Pres[z+b])==str(['--']) or str(Pres[z+b])==str('--') or int(Pres[z+b])==-
999:pass
else:
    #try:
        Pressure[x][z]+=float(Pres[z+b])
        matrix1[x][z]+=1
    #except:
        #pass
if str(Alt[z+b])==str(['--']) or str(Alt[z+b])==str('--') or int(Alt[z+b])==-999:pass
else:
    #try:
        Altitude[x][z]+=float(Alt[z+b])
        #print float(Alt[z+b])
        matrix2[x][z]+=1
    #except:
        #pass

    fin.close()
except:continue
#try:
for x in range (len(lat)):
    #print x, "LAT"
    col+=1

```

```

sheet1.write(0, col, lat[x])
sheet11.write(0, col, lat[x])
sheet12.write(0, col, lat[x])
for z in range(0,400):
    #print z
    if matrix[x][z]==0:pass
    else:
        Temp=float(Temperature[x][z])/float(matrix[x][z])
        sheet1.write(z+1, col, float(Temp))

    if matrix1[x][z]==0:pass
    else:
        Press=float(Pressure[x][z])/float(matrix1[x][z])
        sheet11.write(z+1, col, float(Press))

    if matrix1[x][z]==0:pass
    else:
        Altit=float(Altitude[x][z])/float(matrix2[x][z])
        sheet12.write(z+1, col, float(Altit))
    #print matrix1[d][x][y], matrix[d][x][y]

#except:pass
book.save('new_T__'+str(year)+'.xls')
book1.save('new_P__'+str(year)+'.xls')
book2.save('new_A__'+str(year)+'.xls')

```

Script 1.5 The next script interpolates the data (temperature or geopotential heights) for selected pressure levels

```

import numpy as N
import xlrd
import os
import sys
import glob
from scipy import stats
from scipy import polyfit
import matplotlib.pyplot as plt

pres=[]
a=400

while (a > 300):#creating latitude coordinates each 1 degree
    pres.append(a)
    a-=10
print len(pres)

```

```

while (a > 50):
    pres.append(a)
    a-=5
print len(pres)

while (a >= 10):
    pres.append(a)
    a-=1
print len(pres)
print pres

print bingo
#pres = np.arange(5.0, 30.2, 0.1)
#lat = np.arange(-89.5, 89.5, 1)

for year in range(2001,2002):
    Path='G:\\gropotential height\\final_data\\'
    fi = xlrd.open_workbook(Path+'New_T__'+str(year)+'.xls') # change New_T to New_A to
calculate geopotential heights
    fi1 = xlrd.open_workbook(Path+'New_P__'+str(year)+'.xls')
    for month in range(1,13):
        sh = fi.sheet_by_index(month-1)
        sh1 = fi1.sheet_by_index(month-1)
        output='G:\\gropotential
height\\final_data\\txt_data\\interpol_txt\\'+str(year)+'_'+str(month)+'.txt'
        fo = open(output, 'w')
        for ry in range(1,sh.ncols):
            if ry==sh.ncols-1:
                fo.write (str(sh.row_values(0)[ry])+'\n')
            else:
                fo.write (str(sh.row_values(0)[ry])+',')

    for i in range(len(pres)):
        print pres[i]

        for ry in range(1,sh1.ncols):
            y=""
            try:
                for rx in range(10,sh1.nrows):
                    #if str(sh1.row_values(rx)[ry]).rstrip('\n')==":print 'BINGO'
                    if sh1.row_values(rx)[ry]<pres[i] and str(sh1.row_values(rx)[ry]).rstrip('\n')!=":
                        #print sh1.row_values(rx)[ry]
                        x0=float(sh1.row_values(rx)[ry])
                        x1=float(sh1.row_values(rx-1)[ry])
                        dx0=1-(pres[i]-x0)/(x1-x0)

```

```

        dx1=1-(x1-pres[i])/(x1-x0)
        y=sh.row_values(rx)[ry]*dx0 + sh.row_values(rx-1)[ry]*dx1
        if y>100:
            print y, dx0, dx1, x0, x1, sh.row_values(rx)[ry], sh.row_values(rx-1)[ry],
year, month, pres[i], sh1.row_values(0)[ry]
            #print sh1.row_values(rx)[ry], sh1.row_values(rx-1)[ry], pres[i], dx0,
dx1,sh.row_values(rx)[ry], sh.row_values(rx-1)[ry], y
            break

    if ry==sh.ncols-1:
        fo.write (str(y)+'\n')
    else:
        fo.write (str(y)+',')
except:
    if ry==sh.ncols-1:
        fo.write (str(y)+'\n')
    else:
        fo.write (str(y)+',')

    #print 'empty col'
    pass

fo.close

```

Script 1.6 Next script calculates trends for each point from LxA grid (section 3.3).

```

import numpy as N
import xlrd
import xlwt
import os
import sys
import glob
from scipy import stats
from scipy import polyfit
#from rpy_options import set_options
#set_options(RHOME='C:\Program Files (x86)\R\bin\i386\bin\')
#import rpy
#from rpy import r
import matplotlib.pyplot as plt

sheetName='Trends'
book = xlwt.Workbook()
sheet1 = book.add_sheet(sheetName,cell_overwrite_ok=True)

sheet2 = book.add_sheet(sheetName+'1',cell_overwrite_ok=True)

```

```

for lat in range(181):
    print lat
    for alt in range(1,103):
        print alt
        y=[]
        x=[]
        a=0
        for year in range(2001,2012):
            for month in range(1,13):
                Path='G:\\gropotential
height\\final_data\\txt_data\\interpol_zg\\'+str(year)+'_'+str(month)+'.txt'
                f=open(Path, 'rU')
                lines=f.readlines()
                f.close()

                try:
                    words=lines[alt].split(',')

                    y.append(float(words[lat]))
                    a+=1
                    x.append(float(a))
                    if float(words[lat])>100:
                        print year, month, lat, alt, words[lat]
                        print Path
                except:
                    a+=1
                    #reg.append("")
            #fi.close()
        try:
            print y
            print len(y)
            print x
            print len(x)
            print a
            A = N.vstack([x, N.ones(len(x))]).T

            m, c = N.linalg.lstsq(A,y)[0]
            m1, c1 = polyfit(x,y,1)
            print m, c
            print m1, c1
            sheet1.write(alt,lat,float(m)*120)
            sheet2.write(alt,lat,float(m1)*120)
        except:
            print "empty cell"
            pass

```

```

#plt.plot(x, y, 'o', label='Original data', markersize=10)
#pl=[]
#for i in range(len(x)):
#    #pl.append(m1*x[i] + c1)
#print pl
#plt.plot(x, pl, 'r', label='Fitted line')
#plt.legend()
#plt.show()

book.save('interpol_trends_from_txt_zg.xls')

```

2. GCM python scripts

2.1 PICTRL data

Script 2.1.1 First script splits *netcdf* files into smaller files because original files can be too big for scipy *netcdf* module to process them (can be used *netcdf4* module, but it is very slow)

```

import numpy as N
#import pylab as pl
import netCDF4
import os
from netCDF4 import Dataset
from scipy import stats
from scipy import polyfit
from scipy.io import netcdf
from netCDF4 import Dataset, date2num
import glob
import netcdftime as nt

Path='N:\\Kirilenko_Coburn_ModelOutput\\Pre Industrial Control Run\\'
#Path='N:\\Kirilenko_Coburn_ModelOutput\\Pre Industrial Control Run\\geopotential_height\\'
l=os.listdir(Path)
#for li in range(1,len(l)):
for li in range(26,27):
    print l[li]
    #print bingo
    index=0
    for infile in glob.glob( os.path.join(Path+str(l[li])+ '\\', '*nc' ) ):
        print infile

        #fin=Dataset(Path+'ta_Amon_IPSL-CM5A-MR_piControl_r1i1p1_180001-
189912_test.nc','r+')
        fin=Dataset(infile)
        #T = fin.variables['average_T1']
        a=[]

```

```

#print fin.variables
#print fin.dimensions
#print fin.groups

Pl=fin.dimensions['plev']
Tm=fin.dimensions['time']
Lt=fin.dimensions['lat']
Ln=fin.dimensions['lon']
#Bn=fin.dimensions['bnds']
plevs1=fin.variables['plev']
times1=fin.variables['time']
lats1=fin.variables['lat']
lons1=fin.variables['lon']
#temp1=fin.variables['zg']
temp1=fin.variables['ta']
print times1.units
#print times1.units[11:15]
#e=float(times1.units[11:15])*365
#print e

#print bingo
#print len(Tm)
#if len(Tm)>220:
#    #print float(float(len(Tm))/220), len(Tm)/220
#    #b=len(Tm)/220
#    #for num in range(b):
#        #a.append(220)
#    #if float(float(len(Tm))/220)!=len(Tm)/220:
#        #i = len(Tm)-b*220
#        #a.append(i)
#    #print a
#else:
#    #a.append(len(Tm))
if len(Tm)>22:
    print float(float(len(Tm))/22), len(Tm)/22
    b=len(Tm)/22
    for num in range(b):
        a.append(22)
    if float(float(len(Tm))/22)!=len(Tm)/22:
        i = len(Tm)-b*22
        a.append(i)
    print a
else:
    a.append(len(Tm))
#print bingo
c=0

```

```

#for i in range(1):
for i in range(len(a)):
    print a[i]
    index+=1
    #f = netcdf.netcdf_file(str(i)+'111tst_2.nc','w')
    #f = netcdf.netcdf_file(Path+'new_models\\'+str(l[li])+'_'+str(index).zfill(3)+'.nc','w')
    f = netcdf.netcdf_file(Path+'aall_model1\\'+str(l[li])+'_'+str(index).zfill(3)+'.nc','w')
    #f = Dataset(str(i)+'111tst_2.nc','w', format='NETCDF4')

    time=f.createDimension('time', a[i])
    plev=f.createDimension('plev',len(P1))
    lat=f.createDimension('lat',len(Lt))
    lon=f.createDimension('lon',len(Ln))
    #f.createDimension('bnds',len(Bn))

    times = f.createVariable('time','f4',('time',))
    plevs = f.createVariable('plev','f4',('plev',))
    lats = f.createVariable('lat','f4',('lat',))
    lons = f.createVariable('lon','f4',('lon',))
    temp = f.createVariable('ta','f4',('time','plev','lat','lon',))
    times.units =times1.units
    times[:]=times1[c:a[i]+c]
    plevs[:]=plevs1[: ]
    lats[:]=lats1[: ]
    lons[:]=lons1[: ]
    temp[:,::,::]=temp1[c:a[i]+c,::,::]

    print temp[0,6,11,17],'AND',temp1[0+c,6,11,17]
    f.close()
    c+=a[i]
    print c, 'C'
fin.close()

```

Script 2.1.2 Next script processes created *netcdf* files into *excel* files with averaging all longitude values over latitude bands. Only selected pressure levels are used and latitude bands are used (section 3.3).

```

import numpy as N
import xlrd
import xlwt
import os
import sys
import glob
from scipy import stats
from scipy import polyfit
from scipy.io import netcdf

```



```

from netCDF4 import Dataset

#Path='N:\\Kirilenko_Coburn_ModelOutput\\Pre Industrial Control Run\\aall_models\\'
Path='C:\\ta\\'

class MyWorkbook:
    """ allow access to a workbooks sheets"""
    def __init__(self,*args,**kwargs):
        self.wb = xlwt.Workbook(*args,**kwargs)
        self.sheets = []
    def add_sheet(self,sheet_name):
        self.sheets.append(self.wb.add_sheet(sheet_name,cell_overwrite_ok=True))
        return self.sheets[-1]
    def GetSheetByIndex(self,n):
        return self.sheets[n]
    def save(self,fname_or_stream):
        return self.wb.save(fname_or_stream)

#Path='G:\\GCM\\pcmdi.ipcc4.gfdl_cm2_0.sresa1b.run1.monthly.ta_A1.200101-210012.nc'
#Path='P:\\Sergey\\GISS-E2-H_piControl_r1i1p1\\'
#Path='C:\\Users\\Sergey.Molodtsov\\Downloads\\PICTRL1\\GISS-E2-R_piControl_r1i1p1\\'

#Path='N:\\Kirilenko_Coburn_ModelOutput\\Pre Industrial Control Run\\GISS-E2-
R_piControl_r1i1p1\\'
#sheet1 = book.add_sheet(sheetName,cell_overwrite_ok=True)
#sheet2 = book.add_sheet(sheetName+'1',cell_overwrite_ok=True)

#book = xlwt.Workbook()

l=os.listdir(Path)
for li in range(len(l)):
#for infile in glob.glob( os.path.join(Path,'*nc') ):
    #sheetName='Trends'

    row=1
    print l[li][-24]
    #print bingo
    #if str(l[li][-24])=='CMCC-CESM' or str(l[li][-24])=='CMCC-CMS' or str(l[li][-
24])=='GFDL-CM3' or str(l[li][-24])=='MIROC-ESM-CHEM' or str(l[li][-24])=='MIROC-
ESM':continue
    #print bingo
    fin=netcdf.netcdf_file(Path+str(l[li]),'r')
    fin1=Dataset(Path+str(l[li]),'r')
    #fin=Dataset(infile,'r+')
    m= MyWorkbook()

```

```

for lvl in range(7,17):
    sheetName=str(lvl)
    m.add_sheet(sheetName)
#print len(fin.variables['ta'])
time_gcm = fin.variables['time']
time_gcm1 = fin.variables['time']
ta = fin.variables['ta']
lat = fin.variables['lat']
pres = fin.variables['plev']
lon = fin.variables['lon']
print time_gcm1.units[11:15]
e=float(time_gcm1.units[11:15])*365
print e
fin1.close()
#print len(lat[:])
#print bingo
#for x in range(20,76):#96 lat
#for x in range(31,112):#IPSL-CM5A-MR
#for x in range(len(lat[:])):
    #for x in range(19,71):
        #print lat[x], x
#print bingo
#for lvl in range(17):
    #print pres[lvl], lvl
#for x in range(31,112):
    #print lat[x], x

#print ta[z,lvl,x,y]
#print bingo
sheet=0
#for lvl in range(8,18):#GFDL-CM3 16
for lvl in range(7,23):#MIROC-ESM-CHEM, MIROC-ESM, CMCC-CESM,9,27,28
#for lvl in range(7,22): #CMCC-CMS, 11
    if lvl==10 or lvl==12 or lvl==13 or lvl==15 or lvl==16 or lvl==21:continue# or
lvl==21:continue
    #if lvl==10 or lvl==12 or lvl==13 or lvl==15 or lvl==16:continue# or lvl==21:continue
#print len(lat[:])
#print bingo
#print lvl
#sheetName=str(pres[lvl])
#sheet1 = book.add_sheet(sheetName,cell_overwrite_ok=True)
#book.active_sheet=lvl
#sheet1==book.active_sheet(sheet)
if li==1:
    print pres[lvl]
s = m.GetSheetByIndex(sheet)
sheet+=1

```



```

import xlrd
import xlwt
import glob

class MyWorkbook:
    """ allow access to a workbooks sheets"""
    def __init__(self,*args,**kwargs):
        self.wb = xlwt.Workbook(*args,**kwargs)
        self.sheets = []
    def add_sheet(self,sheet_name):
        self.sheets.append(self.wb.add_sheet(sheet_name,cell_overwrite_ok=True))
        return self.sheets[-1]
    def GetSheetByIndex(self,n):
        return self.sheets[n]
    def save(self,fname_or_stream):
        return self.wb.save(fname_or_stream)

```

```
Path = 'G:\\GCM\\Pictrl_34GCMs\\ta\\'
```

```

lat=[]
a=-49.5
while (a <= 49.5):
    lat.append(a)
    a+=1
print lat
print len(lat)
count=0
for infile in glob.glob( os.path.join(Path,'*xls') ):
    count+=1
    #if count<23:continue
    sheet=0
    print infile[27:]
    #print bingo
    fi = xlrd.open_workbook(infile)
    m= MyWorkbook()
    for lvl in range(7,17):
        sheetName=str(lvl)
        m.add_sheet(sheetName)
    for lvl in range(7,17):
        #print lvl
        sh = fi.sheet_by_index(sheet)
        s = m.GetSheetByIndex(sheet)
        sheet+=1
        col=0
        for x in range(len(lat)):
            for ry in range(1,sh.ncols):

```

```

if float(sh.row_values(0)[ry])>float(lat[x]) and sh.row_values(0)[ry]!=str("):
    col+=1
    s.write(0,col, lat[x])
    x0=float(sh.row_values(0)[ry-1])
    x1=float(sh.row_values(0)[ry])
    #print sh.row_values(0)[ry-1],sh.row_values(0)[ry]
    #print float(x0), x1
    dx0=1-(lat[x]-x0)/(x1-x0)
    dx1=1-(x1-lat[x])/(x1-x0)
    for rx in range(1, sh.nrows):
        if col==1:
            s.write(rx,0, sh.row_values(rx)[0])

            y=sh.row_values(rx)[ry-1]*dx0 + sh.row_values(rx)[ry]*dx1
            s.write(rx,col, y)
        break
m.save('ta\\interpolated_'+str(infile[27:]))

```

Script 2.1.4 This script calculates trends from PICTRL excel files.

```

import xlrd
import xlwt
from scipy import stats
from scipy import polyfit
import numpy as N
import glob
import os
#Path='G:\\GCM\\15_GCMs_RCP+Hist.xls'
#Path='G:\\GCM\\BNU-ESM_r1i1p1_200101-201112.xls'
Path='G:\\GCM\\PICTRL_DATA\\PICTRL_1\\'

book = xlwt.Workbook()
sheet = book.add_sheet('Trends',cell_overwrite_ok=True)
col=0

for infile in glob.glob( os.path.join(Path,*xls' )):
    fi = xlrd.open_workbook(infile)
    month=0
    pres=[300, 250, 200, 150, 100, 70, 50, 30, 20, 10]
    months=132

    sh_test = fi.sheet_by_index(0)
    per=sh_test.nrows/132
    print float(per)
    print int(per)

```

```

print 'COL', col
#print bingo

#for period in range(10):#change the subsets of PICTRL data
for period in range(10,20):
    col+=1
    print period
    row=0
    sheet.write(0,col+2,period)
    for lvl in range(7,17):
        month=months
        sh = fi.sheet_by_index(lvl-7)
        for ry in range(1,sh.ncols):
            row+=1
            a=0
            y=[]
            x=[]
            for rx in range(month-131,month+1):
                a+=1
                if a<5 or a>125:continue
                else:
                    try:
                        y.append(float(sh.row_values(rx)[ry]))
                        x.append(float(a))
                    except:
                        #print lvl, month, sh.row_values(rx)[ry]
                        #print 'bingo'
                        pass
            #A = N.vstack([x, N.ones(len(x))]).T
            m, c = polyfit(x,y,1)
            #m, c = N.linalg.lstsq(A,y)[0]
            if period==0:
                sheet.write(row,0,float(pres[lvl-7]))
                sheet.write(row,1,float(sh.row_values(0)[ry]))
                sheet.write(row,col+2,float(m)*120)
            months+=132
book.save('trends_CMIP5_PICTRL_110yr_part2.xls')

```

Script 2.1.5 The next script merges all the excel files into one

```

import xlrd
import xlwt
import os
import glob

```

```

class MyWorkbook:
    """ allow access to a workbooks sheets"""
    def __init__(self,*args,**kwargs):
        self.wb = xlwt.Workbook(*args,**kwargs)
        self.sheets = []
    def add_sheet(self,sheet_name):
        self.sheets.append(self.wb.add_sheet(sheet_name,cell_overwrite_ok=True))
        return self.sheets[-1]
    def GetSheetByIndex(self,n):
        return self.sheets[n]
    def save(self,fname_or_stream):
        return self.wb.save(fname_or_stream)

m= MyWorkbook()

for lvl in range(7,17):
    sheetName=str(lvl)
    m.add_sheet(sheetName)
Path='G:\\GCM\\Pictrl_34GCMs\\intepolated\\zg3\\'

row=0

a=0

l=os.listdir(Path)
for li in range(len(l)):

#for infile in glob.glob( os.path.join(Path,'*xls') ):
    #print infile[50:-8]
    print l[li][:8]
    #print bingo
    a+=1
    fi = xlrd.open_workbook(Path+str(l[li]))
    sheet=0
    for lvl in range(7,17):
        sh = fi.sheet_by_index(lvl-7)
        s = m.GetSheetByIndex(sheet)
        sheet+=1
        if a==1:
            for ry in range(1,sh.ncols):
                s.write(0,ry,float(sh.row_values(0)[ry]))
        x=0
        for rx in range(1,sh.nrows):
            x+=1
            for ry in range(sh.ncols):
                s.write(row+rx,ry,float(sh.row_values(rx)[ry]))

```

```

print x
row+=x
print 'done', l[li]

```

```

#m.save('15_GCMs_RCP+Hist.xls')
m.save('zg_33PICTRL_p3.xls')

```

2.2 GCM historical+RCP8.5 data (section 3.3, table 2)

Script 2.2.1 As it was in 2.1 the first script splits *netcdf* files into smaller files because original files can be too big for *scipy netcdf* module to process them (can be used *netcdf4* module, but it is very slow). There are different versions of this script, because some GCMs have different amount of files which are used, other GCMs stop their historical run in the date which is different from the majority's GCMs date. Here are two scripts – first for historical scenario (section 3.3).

```

import numpy as N
#import pylab as pl
import netCDF4
import os
from netCDF4 import Dataset
from scipy import stats
from scipy import polyfit
from scipy.io import netcdf
from netCDF4 import Dataset, date2num
import glob
import netcdftime as nt

Path='N:\\Kirilenko_Coburn_ModelOutput\\D - Historical\\Monthly\\new historical
models\\2005_12_1file\\zg\\'
Path_out='N:\\Kirilenko_Coburn_ModelOutput\\D - Historical\\Monthly\\new historical
models\\'
##l=os.listdir(Path)
#print bingo
#for li in range(1,len(l)):

for infile in glob.glob(os.path.join(Path,'*nc')) :
    print infile[94:-17]
    #print bingo
    #fin=Dataset(Path+'ta_Amon_IPSL-CM5A-MR_piControl_r1i1p1_180001-
189912_test.nc','r+')
    fin=Dataset(infile)
    #T = fin.variables['average_T1']
    a=[]
    #print fin.variables

```



```

#print fin.dimensions
#print fin.groups

Pl=fin.dimensions['plev']
Tm=fin.dimensions['time']
Lt=fin.dimensions['lat']
Ln=fin.dimensions['lon']
#Bn=fin.dimensions['bnds']
plevs1=fin.variables['plev']
times1=fin.variables['time']
lats1=fin.variables['lat']
lons1=fin.variables['lon']
temp1=fin.variables['zg']
#temp1=fin.variables['ta']
print times1.units
if len(Tm)>12:
    print float(float(len(Tm))/1), len(Tm)/1
    b=5
    #b=len(Tm)/12
    for num in range(b):
        a.append(1)
    if float(float(len(Tm))/1)!=len(Tm)/1:
        i = len(Tm)-b*1
        a.append(i)
    print a
else:
    a.append(len(Tm))
#continue
year = 2001
for i in range(b):
    print i
    #print bingo
    #f = netcdf.netcdf_file(str(i)+'111tst_2.nc','w')
    #f = netcdf.netcdf_file(Path+'new_models\\'+str(l[li])+'+'+str(index).zfill(3)+'_'.nc','w')
    f = netcdf.netcdf_file(Path_out+'zg\\'+str(infile[94:-17])+'+'+str(year+i)+'01-
'+str(year+i)+'12.nc','w')
    #f = Dataset(str(i)+'111tst_2.nc','w', format='NETCDF4')

time=f.createDimension('time', 12)
plev=f.createDimension('plev',len(Pl))
lat=f.createDimension('lat',len(Lt))
lon=f.createDimension('lon',len(Ln))
#f.createDimension('bnds',len(Bn))

times = f.createVariable('time','f4',('time',))
plevs = f.createVariable('plev','f4',('plev',))

```

```

lats = f.createVariable('lat','f4',('lat',))
lons = f.createVariable('lon','f4',('lon',))
temp = f.createVariable('ta','f4',('time','plev','lat','lon',))
times.units =times1.units
times[:]=times1[len(times1)-60+(i+1)*12-12:len(times1)-60+(i+1)*12]
plevs[:]=plevs1[:]
lats[:]=lats1[:]
lons[:]=lons1[:]
temp[:,::,,:]=temp1[len(times1)-60+(i+1)*12-12:len(times1)-60+(i+1)*12,::,,:]
print times[:]
#print temp[0,6,11,17],'AND',temp1[0+c,6,11,17]
f.close()
#c+=a[i]
#print c, 'C'
fin.close()
#index+=1

```

Script 2.2.2 The second for RCP 8.5 scenario (section 3.3)

```

import numpy as N
#import pylab as pl
import netCDF4
import os
from netCDF4 import Dataset
from scipy import stats
from scipy import polyfit
from scipy.io import netcdf
from netCDF4 import Dataset, date2num
import glob
import netcdftime as nt

Path='N:\\Kirilenko_Coburn_ModelOutput\\G - RCP8.5\\Monthly\\new rc
models\\2006_1_1file\\zg\\'
Path_out='N:\\Kirilenko_Coburn_ModelOutput\\G - RCP8.5\\Monthly\\new rc
models\\'
##l=os.listdir(Path)
#print bingo
#for li in range(1,len(l)):

for infile in glob.glob(os.path.join(Path,'*nc')) :
    print infile[82:-17]
    #print bingo
    #fin=Dataset(Path+'ta_Amon_IPSL-CM5A-MR_piControl_r1i1p1_180001-
189912_test.nc','r+')
    fin=Dataset(infile)
    #T = fin.variables['average_T1']
    a=[]

```

```

#print fin.variables
#print fin.dimensions
#print fin.groups

Pl=fin.dimensions['plev']
Tm=fin.dimensions['time']
Lt=fin.dimensions['lat']
Ln=fin.dimensions['lon']
#Bn=fin.dimensions['bnds']
plevs1=fin.variables['plev']
times1=fin.variables['time']
lats1=fin.variables['lat']
lons1=fin.variables['lon']
temp1=fin.variables['zg']
#temp1=fin.variables['ta']
print times1.units

#f = netcdf.netcdf_file(str(i)+'111tst_2.nc','w')
#f = netcdf.netcdf_file(Path+'new_models\\'+str(l[li])+ '_'+str(index).zfill(3)+'_nc','w')
f = netcdf.netcdf_file(Path_out+'zg\\'+str(infile[82:-17])+ '_200601-201112.nc','w')
#f = Dataset(str(i)+'111tst_2.nc','w', format='NETCDF4')

time=f.createDimension('time', 72)
plev=f.createDimension('plev',len(Pl))
lat=f.createDimension('lat',len(Lt))
lon=f.createDimension('lon',len(Ln))
#f.createDimension('bnds',len(Bn))

times = f.createVariable('time','f4',('time',))
plevs = f.createVariable('plev','f4',('plev',))
lats = f.createVariable('lat','f4',('lat',))
lons = f.createVariable('lon','f4',('lon',))
temp = f.createVariable('ta','f4',('time','plev','lat','lon',))
times.units =times1.units
times[:]=times1[:72]
plevs[:]=plevs1[:]
lats[:]=lats1[:]
lons[:]=lons1[:]
temp[:,::,:]=temp1[:72,::,:]

#print temp[0,6,11,17],'AND',temp1[0+c,6,11,17]
f.close()
#c+=a[i]
#print c, 'C'
fin.close()

```

```
#index+=1
```

Script 2.2.3 The next script processes *netcdf* files into *excel* files, choosing selected pressure levels and averaging all longitude values over selected latitude bands

```
import numpy as N
import xlrd
import xlwt
import os
import sys
import glob
from scipy import stats
from scipy import polyfit
from scipy.io import netcdf
from netCDF4 import Dataset

Path='N:\\Kirilenko_Coburn_ModelOutput\\RCP+Historical\\zg\\'

class MyWorkbook:
    """ allow access to a workbooks sheets"""
    def __init__(self,*args,**kwargs):
        self.wb = xlwt.Workbook(*args,**kwargs)
        self.sheets = []
    def add_sheet(self,sheet_name):
        self.sheets.append(self.wb.add_sheet(sheet_name,cell_overwrite_ok=True))
        return self.sheets[-1]
    def GetSheetByIndex(self,n):
        return self.sheets[n]
    def save(self,fname_or_stream):
        return self.wb.save(fname_or_stream)

#Path='G:\\GCM\\pcmdi.ipcc4.gfdl_cm2_0.sresa1b.run1.monthly.ta_A1.200101-210012.nc'
#Path='P:\\Sergey\\GISS-E2-H_piControl_r1i1p1\\'
#Path='C:\\Users\\Sergey.Molodtsov\\Downloads\\PICTRL1\\GISS-E2-R_piControl_r1i1p1\\'

#Path='N:\\Kirilenko_Coburn_ModelOutput\\Pre Industrial Control Run\\GISS-E2-
R_piControl_r1i1p1\\'
#sheet1 = book.add_sheet(sheetName,cell_overwrite_ok=True)
#sheet2 = book.add_sheet(sheetName+'1',cell_overwrite_ok=True)

#book = xlwt.Workbook()

l=os.listdir(Path)
for li in range(16,17):
    print l[li]
    #print bingo
```

```

row=1
m= MyWorkbook()
#for lvl in range(8,18):#GFDL-CM3 16
#for lvl in range(7,23):#MIROC-ESM-CHEM, MIROC-ESM, CMCC-CESM, CMCC-CMS
10, 11, 27,28
    #if lvl==10 or lvl==12 or lvl==13 or lvl==15 or lvl==16 or lvl==21:

for lvl in range(7,17):
    sheetName=str(lvl)
    m.add_sheet(sheetName)
for infile in glob.glob( os.path.join(Path+str(l[li])+"\\",'*nc' )):
    #sheetName='Trends'

    print infile[77:]
    #print bingo
    fin=netcdf.netcdf_file(infile,'r')
    fin1=Dataset(infile)
    #fin=Dataset(infile,'r+')

    #print len(fin.variables['ta'])
    time_gcm = fin.variables['time']
    time_gcm1 = fin.variables['time']
    ta = fin.variables['ta']#ONLY ta because I created all files with such variable name
    #ta = fin.variables['zg']
    lat = fin.variables['lat']
    pres = fin.variables['plev']
    lon = fin.variables['lon']
    print time_gcm1.units[11:15]
    print len(time_gcm[:])
    e=float(time_gcm1.units[11:15])*365
    print e
    fin1.close()
    #print len(lat[:])
    #print bingo
    #for x in range(20,76):#96 lat
    #for x in range(31,112):#IPSL-CM5A-MR
    #for x in range(len(lat[:])):
        #for x in range(19,71):
            #print lat[x], x
    #print bingo
    #for lvl in range(17):
        #print pres[lvl], lvl
    #for x in range(31,112):
        #print lat[x], x

```



```

        #n+=1
        #print t
        #print len(t)
        #sheet1.write(lvl+1, x,float(m)*120)
        #sheet2.write(lvl+1, x,float(m1)*120)
        #book.save('PICTRL_'+str(infile[52:])+'.xls')

for time in range(len(time_gcm[:])):
    row+=1
    fin.close()

m.save('G:\\GCM\\New_PICTRL_HIST_11_10\\zg_'+str(l[li])+'.xls')
print 'done', infile[77:]

```

Script 2.2.4 Because different GCMs have different spatial resolution next script interpolates GCMs output over selected latitude bands (section 3.3)

```

import sys
import os
import xlrd
import xlwt
import glob

class MyWorkbook:
    """ allow access to a workbooks sheets"""
    def __init__(self,*args,**kwargs):
        self.wb = xlwt.Workbook(*args,**kwargs)
        self.sheets = []
    def add_sheet(self,sheet_name):
        self.sheets.append(self.wb.add_sheet(sheet_name,cell_overwrite_ok=True))
        return self.sheets[-1]
    def GetSheetByIndex(self,n):
        return self.sheets[n]
    def save(self,fname_or_stream):
        return self.wb.save(fname_or_stream)

Path = 'G:\\GCM\\New_PICTRL_HIST_11_10\\ro\\'

lat=[]
a=-50
while (a <= 50):
    lat.append(a)
    a+=2.5
print lat
print len(lat)

```

```

#print bingo
#lat=[]
#a=-49.5
#while (a <= 49.5):
    #lat.append(a)
    #a+=1
#print lat
#print len(lat)

for infile in glob.glob( os.path.join(Path, '*xls' )):
    sheet=0
    print infile[32:-21]
    #print bingo
    fi = xlrd.open_workbook(infile)
    m= MyWorkbook()
    for lvl in range(7,8):
        #for lvl in range(7,17):
            sheetName=str(lvl)
            m.add_sheet(sheetName)
    for lvl in range(7,8):
        #for lvl in range(7,17):
            print lvl
            sh = fi.sheet_by_index(sheet)
            s = m.GetSheetByIndex(sheet)
            sheet+=1
            col=0
            for x in range(len(lat)):
                for ry in range(1,sh.ncols):
                    if float(sh.row_values(0)[ry])>float(lat[x]) and sh.row_values(0)[ry]!=str("):
                        col+=1
                        s.write(0,col, lat[x])
                        x0=float(sh.row_values(0)[ry-1])
                        x1=float(sh.row_values(0)[ry])
                        #print sh.row_values(0)[ry-1],sh.row_values(0)[ry]
                        #print float(x0), x1
                        dx0=1-(lat[x]-x0)/(x1-x0)
                        dx1=1-(x1-lat[x])/(x1-x0)
                        for rx in range(1, sh.nrows):
                            if col==1:
                                s.write(rx,0, sh.row_values(rx)[0])

                                y=sh.row_values(rx)[ry-1]*dx0 + sh.row_values(rx)[ry]*dx1
                                s.write(rx,col, y)
                            break
    m.save(str(infile[32:-21])+'_25degree.xls')

```


Script 2.2.5 Next script merges all *excel* files into one

```
import xlrd
import xlwt
import os
import glob

class MyWorkbook:
    """ allow access to a workbooks sheets"""
    def __init__(self,*args,**kwargs):
        self.wb = xlwt.Workbook(*args,**kwargs)
        self.sheets = []
    def add_sheet(self,sheet_name):
        self.sheets.append(self.wb.add_sheet(sheet_name,cell_overwrite_ok=True))
        return self.sheets[-1]
    def GetSheetByIndex(self,n):
        return self.sheets[n]
    def save(self,fname_or_stream):
        return self.wb.save(fname_or_stream)

m= MyWorkbook()

for lvl in range(7,17):
    sheetName=str(lvl)
    m.add_sheet(sheetName)
Path='G:\\GCM\\Pictrl_34GCMs\\intepolated\\zg3\\'

row=0

a=0

l=os.listdir(Path)
for li in range(len(l)):

#for infile in glob.glob( os.path.join(Path,'*xls') ):
    #print infile[50:-8]
    print l[li][::-8]
    #print bingo
    a+=1
    fi = xlrd.open_workbook(Path+str(l[li]))
    sheet=0
    for lvl in range(7,17):
        sh = fi.sheet_by_index(lvl-7)
        s = m.GetSheetByIndex(sheet)
        sheet+=1
```

```

if a==1:
    for ry in range(1,sh.ncols):
        s.write(0,ry,float(sh.row_values(0)[ry]))
x=0
for rx in range(1,sh.nrows):
    x+=1
    for ry in range(sh.ncols):
        s.write(row+rx,ry,float(sh.row_values(rx)[ry]))

print x
row+=x
print 'done', l[li]

#m.save('15_GCMs_RCP+Hist.xls')
m.save('zg_33PICTRL_p3.xls')

```

Script 2.2.6 This script calculates trends for each particular GCM

```

import xlrd
import xlwt
from scipy import stats
from scipy import polyfit
import numpy as N
import glob
import os
Path='G:\\GCM\\ta_34_GCMs_RCP+Hist_25degree.xls'#####CHANGE HERE
ZG/TA
#Path='G:\\GCM\\BNU-ESM_r1i1p1_200101-201112.xls'

Path1 = 'G:\\GCM\\New_PICTRL_HIST_11_10\\interpolated_tables\\ta_25degree\\'
names=[]
for infile in glob.glob( os.path.join(Path1,'*xls') ):
    #print infile[55:-4]
    print infile[64:-4]
    #names.append(str(infile[55:-4]))
    names.append(str(infile[64:-4]))
#print bingo

fi = xlrd.open_workbook(Path)

book = xlwt.Workbook()
sheet = book.add_sheet('Trends',cell_overwrite_ok=True)
month=0
pres=[300, 250, 200, 150, 100, 70, 50, 30, 20, 10]
months=132

```

```

sh_test = fi.sheet_by_index(0)
per=sh_test.nrows/132
print float(per)
print int(per)
#print bingo

for period in range(per):
    print period
    row=0
    sheet.write(0,period+2,str(names[period]))
    for lvl in range(7,17):
        month=months
        sh = fi.sheet_by_index(lvl-7)
        for ry in range(1,sh.ncols):
            row+=1
            a=0
            y=[]
            x=[]
            for rx in range(month-131,month+1):
                a+=1
                #if a<5 or a>125:continue
                if a<5 or a>124:continue###TEST
                else:
                    try:
                        y.append(float(sh.row_values(rx)[ry]))
                        x.append(float(a))
                    except:
                        #print lvl, month, sh.row_values(rx)[ry]
                        #print 'bingo'
                        pass
            #A = N.vstack([x, N.ones(len(x))]).T
            m, c = polyfit(x,y,1)
            #m, c = N.linalg.lstsq(A,y)[0]
            if period==0:
                sheet.write(row,0,float(pres[lvl-7]))
                sheet.write(row,1,float(sh.row_values(0)[ry]))
                sheet.write(row,period+2,float(m)*120)
        months+=132
book.save('ta_34_GCMs_RCP+Hist_TRENDS_FIXED_25degree.xls')

```

3 Optimal fingerprinting in R (R Core Team, 2012)

#First it is necessary to read PICTR and CONTROL data which is organized as a table with 160 columns and #800 (as total there are 800 points) rows

```
t1 <- read.csv("PICTRL_5.csv")
```

```
c1 <- read.csv("CONTROL_5.csv")
```

```
trends <- t1
```

```
CONTROL <- array(c1)
```

```
# Covariance matrix calculation
```

```
C <- cov(t(trends))
```

```
# Eigenvectors and eigenvalues calculation
```

```
eig <- eigen(C)
```

```
val <- eig$values
```

```
vec <- eig$vectors
```

```
# Then read GPS RO data, which is organized as a table with one column and 800 data points
```

```
tRO <- read.csv("RO_25degree.csv")
```

```
RO <- c(tRO)
```

```
# Same for forced GCM data which is organized as a table with one column and 800 data points
```

```
tGCM <- read.csv("fGCM_33_ta.csv")
```

```
GCM <- c(tGCM)
```

```
Y <- c(RO)
```

```
# Calculation of 14 scaling factors using 1-14 EOFs
```

```
vec5 <-
```

```
cbind(vec[,1],vec[,2],vec[,3],vec[,4],vec[,5],vec[,6],vec[,7],vec[,8],vec[,9],vec[,10],vec[,11],vec[,12],vec[,13],vec[,14])
```

```

DMinverse14 <-
diag(c(cbind(val[1],val[2],val[3],val[4],val[5],val[6],val[7],val[8],val[9],val[10],val[11],val[12],
val[13],val[14])^-1))

bRO14 <- DMinverse14%%t(vec5)%%RO$RO

bGCM14 <- DMinverse14%%t(vec5)%%GCM$fGCM

bCONTROL14 <- DMinverse14%%t(vec5)%%t(t(CONTROL))

aRO <- (t(bGCM14)%%DMinverse14%%bGCM14)^-
1%%t(bGCM14)%%DMinverse14%%bRO14

aRO14 <- aRO

aCONTROL <- (t(bGCM14)%%DMinverse14%%bGCM14)^-
1%%t(bGCM14)%%DMinverse14%%bCONTROL14

aCONTR14 <- c(array(aCONTROL))

```

REFERENCES

- Allen, Myles R., and Simon FB Tett. 1999. Checking for model consistency in optimal fingerprinting. *Climate Dynamics* 15 (6): 419-34.
- Anthes, Richard A., PA Bernhardt, Y. Chen, L. Cucurull, KF Dymond, D. Ector, SB Healy, S-P Ho, DC Hunt, and Y-H Kuo. 2008. THE COSMIC/FORMOSAT-3 MISSION early results. *Bulletin of the American Meteorological Society* 89 (3).
- Arakawa, Akio. 1970. Numerical simulation of large-scale atmospheric motions. *Numerical Solution of Field Problems in Continuum Physics 2* : 24-40.
- Baldwin, MP, LJ Gray, TJ Dunkerton, K. Hamilton, PH Haynes, WJ Randel, JR Holton, MJ Alexander, I. Hirota, and T. Horinouchi. 2001. The quasi-biennial oscillation. *Reviews of Geophysics* 39 (2): 179-229.
- Baldwin, M. P., M. Dameris, and T. G. Shepherd. 2007. Atmosphere. how will the stratosphere affect climate change? *Science (New York, N.Y.)* 316 (5831) (Jun 15): 1576-7.
- Bell, Thomas L. 1982. Optimal weighting of data to detect climatic change: Application to the carbon dioxide problem. *Journal of Geophysical Research: Oceans (1978–2012)* 87 (C13): 11161-70.
- Borsche, M., G. Kirchengast, and U. Foelsche. 2007. Tropical tropopause climatology as observed with radio occultation measurements from CHAMP compared to ECMWF and NCEP analyses. *Geophysical Research Letters* 34 (3).

Brohan, Phillip, John J. Kennedy, Ian Harris, Simon FB Tett, and Phil D. Jones. 2006. Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *Journal of Geophysical Research: Atmospheres* (1984–2012) 111 (D12).

Bromwich, David H., Julien P. Nicolas, Andrew J. Monaghan, Matthew A. Lazzara, Linda M. Keller, George A. Weidner, and Aaron B. Wilson. 2013. Central west antarctica among the most rapidly warming regions on earth. *Nature Geoscience* 6 (2): 139-45.

Charlton-Perez, Andrew J., Mark P. Baldwin, Thomas Birner, Robert X. Black, Amy H. Butler, Natalia Calvo, Nicholas A. Davis, Edwin P. Gerber, Nathan Gillett, and Steven Hardiman. 2013. On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models. *Journal of Geophysical Research: Atmospheres* 118 (6): 2494-505.

———. 2013. On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models. *Journal of Geophysical Research: Atmospheres* 118 (6): 2494-505.

Christy, John R., Roy W. Spencer, and William D. Braswell. 2000. MSU tropospheric temperatures: Dataset construction and radiosonde comparisons. *Journal of Atmospheric & Oceanic Technology* 17 (9).

Christy, John R., Roy W. Spencer, and Elena S. Lobl. 1998. Analysis of the merging procedure for the MSU daily temperature time series. *Journal of Climate* 11 (8).

Christy, John R., Roy W. Spencer, William B. Norris, William D. Braswell, and David E. Parker. 2003. Error estimates of version 5.0 of MSU–AMSU bulk atmospheric temperatures. *Journal of Atmospheric & Oceanic Technology* 20 (5).

Collins, William D., Cecilia M. Bitz, Maurice L. Blackmon, Gordon B. Bonan, Christopher S. Bretherton, James A. Carton, Ping Chang, Scott C. Doney, James J. Hack, and Thomas B.

Henderson. 2006. The community climate system model version 3 (CCSM3). *Journal of Climate* 19 (11).

Diao, Minghui, Loayeh Jumbam, Justin Sheffield, Eric F. Wood, and Mark A. Zondlo. 2013. Validation of AIRS/AMSU-A water vapor and temperature data with in situ aircraft observations from the surface to UT/LS from 87° N–67° S. *Journal of Geophysical Research: Atmospheres* 118 (12): 6816-36.

Durre, Imke, Thomas C. Peterson, and Russell S. Vose. 2002. Evaluation of the effect of the Luers–Eskridge radiation adjustments on radiosonde temperature homogeneity. *Journal of Climate* 15 (11).

Durre, Imke, Russell S. Vose, and David B. Wuertz. 2006. Overview of the integrated global radiosonde archive. *Journal of Climate* 19 (1).

Foelsche, U., G. Kirchengast, M. Borsche, B. Pirscher, and AK Steiner. 2008. Creating a consistent radio occultation data base for climate studies in the upper troposphere and lower stratosphere. Paper presented at Proc. of the ECMWF GRAS-SAF Workshop on Applications of RO Measurements, .

Gaffen, Dian J., Michael A. Sargent, RE Habermann, and John R. Lanzante. 2000. Sensitivity of tropospheric and stratospheric temperature trends to radiosonde data quality. *Journal of Climate* 13 (10).

Gobiet, A., G. Kirchengast, GL Manney, M. Borsche, C. Retscher, and G. Stiller. 2007. Retrieval of temperature profiles from CHAMP for climate monitoring: Intercomparison with envisat MIPAS and GOMOS and different atmospheric analyses. *Atmospheric Chemistry and Physics* 7 (13): 3519-36.

Goldberg, Mitchell D., and Henry E. Fleming. 1995. An algorithm to generate deep-layer temperatures from microwave satellite observations for the purpose of monitoring climate change. *Journal of Climate* 8 (5): 993-1004.

Goody, Richard, James Anderson, and Gerald North. 1998. Testing climate models: An approach. *Bulletin of the American Meteorological Society* 79 (11): 2541-9.

Gray, Lesley J., Simon Crooks, Charlotte Pascoe, Sarah Sparrow, and Michael Palmer. 2004. Solar and QBO influences on the timing of stratospheric sudden warmings. *Journal of the Atmospheric Sciences* 61 (23).

Griffies, Stephen M., Michael Winton, Leo J. Donner, Larry W. Horowitz, Stephanie M. Downes, Riccardo Farneti, Anand Gnanadesikan, William J. Hurlin, Hyun-Chul Lee, and Zhi Liang. 2011. The GFDL CM3 coupled climate model: Characteristics of the ocean and sea ice simulations. *Journal of Climate* 24 (13).

Griggs, David J., and Maria Noguer. 2002. Climate change 2001: The scientific basis. contribution of working group I to the third assessment report of the intergovernmental panel on climate change. *Weather* 57 (8): 267-9.

Hardiman, SC, N. Butchart, TJ Hinton, SM Osprey, and LJ Gray. 2012. The effect of a well-resolved stratosphere on surface climate: Differences between CMIP5 simulations with high and low top versions of the met office climate model. *Journal of Climate* 25 (20).

Hasselmann, Klaus. 1997. Multi-pattern fingerprint method for detection and attribution of climate change. *Climate Dynamics* 13 (9): 601-11.

Hegerl, Gabriele C., HANs voN SToRcH, Klaus Hasselmann, Benjamin D. Santer, Ulrich Cubasch, and Philip D. Jones. 1996. Detecting greenhouse-gas-induced climate change with an optimal fingerprint method. *Journal of Climate* 9 (10): 2281-306.

Hegerl, GC, FW Zwiers, P. Braconnot, NP Gillett, Y. Luo, JA Marengo Orsini, N. Nicholls, JE Penner, and PA Stott. 2007. Understanding and attributing climate change. 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.

Ho, Shu-peng, Doug Hunt, Andrea K. Steiner, Anthony J. Mannucci, Gottfried Kirchengast, Hans Gleisner, Stefan Heise, Axel Engeln, Christian Marquardt, and Sergey Sokolovskiy. 2012. Reproducibility of GPS radio occultation data for climate monitoring: Profile-to-profile inter-comparison of CHAMP climate records 2002 to 2008 from six data centers. *Journal of Geophysical Research: Atmospheres* (1984–2012) 117 (D18).

Ho, Shu-peng, Gottfried Kirchengast, Stephen Leroy, Jens Wickert, Anthony J. Mannucci, Andrea Steiner, Doug Hunt, William Schreiner, Sergey Sokolovskiy, and Chi Ao. 2009. Estimating the uncertainty of using GPS radio occultation data for climate monitoring: Intercomparison of CHAMP refractivity climate records from 2002 to 2006 from different data centers. *Journal of Geophysical Research: Atmospheres* (1984–2012) 114 (D23).

Hyde, William T., Thomas J. Crowley, Steven K. Baum, and W. Richard Peltier. 2000. Neoproterozoic ‘snowball Earth’ simulations with a coupled climate/ice-sheet model. *Nature* 405 (6785): 425-9.

Intergovernmental Panel on Climate Change. Working Group 1. 2007. *Climate change 2007: The physical science basis: Summary for policymakers* IPCC Secretariat.

Jones, Phil D., Mark New, David E. Parker, Seelye Martin, and Ignatius G. Rigor. 1999. Surface air temperature and its changes over the past 150 years. *Reviews of Geophysics* 37 (2): 173-99.

Karl, Thomas R., Susan J. Hassol, Christopher D. Miller, and William L. Murray. 2006. *Temperature trends in the lower atmosphere. steps for understanding and reconciling differences.*

Kidder, Stanley Q., Mitchell D. Goldberg, Raymond M. Zehr, Mark DeMaria, James FW Purdom, Christopher S. Velden, Norman C. Grody, and Sheldon J. Kusselson. 2000. Satellite analysis of tropical cyclones using the advanced microwave sounding unit (AMSU). *Bulletin of the American Meteorological Society* 81 (6): 1241-59.

Knutson, Thomas R., Fanrong Zeng, and Andrew T. Wittenberg. Multi-model assessment of regional surface temperature trends: CMIP3 vs CMIP5 historical (20C3M) runs 2.

Knutson, TR, TL Delworth, KW Dixon, and RJ Stouffer. 1999. Model assessment of regional surface temperature trends (1949–1997). *Journal of Geophysical Research: Atmospheres* (1984–2012) 104 (D24): 30981-96.

Kursinski, ER, GA Hajj, JT Schofield, RP Linfield, and K. Rer Hardy. 1997. Observing earth's atmosphere with radio occultation measurements using the global positioning system. *Journal of Geophysical Research: Atmospheres* (1984–2012) 102 (D19): 23429-65.

Lackner, Bettina C., Andrea K. Steiner, Gottfried Kirchengast, and Gabriele C. Hegerl. 2011. Atmospheric climate change detection by radio occultation data using a fingerprinting method. *Journal of Climate* 24 (20).

Lambert, Steven J., and George J. Boer. 2001. CMIP1 evaluation and intercomparison of coupled climate models. *Climate Dynamics* 17 (2-3): 83-106.

Le Treut, Herve, Richard Somerville, Ulrich Cubasch, Yihui Ding, Cecilie Mauritzen, Abdalah Mokssit, Thomas Peterson, and Michael Prather. 2007. Historical overview of climate change.

Leith, Cecil E. 1964. Numerical simulation of the earth's atmosphere University of California, Lawrence Radiation Laboratory.

Leroy, SS, JG Anderson, and JA Dykema. 2006b. Testing climate models using GPS radio occultation: A sensitivity analysis. *Journal of Geophysical Research: Atmospheres* (1984–2012) 111 (D17).

Leroy, Stephen S. 1997. Measurement of geopotential heights by GPS radio occultation. *Journal of Geophysical Research: Atmospheres* (1984–2012) 102 (D6): 6971-86.

Leroy, Stephen S., John A. Dykema, and James G. Anderson. 2006a. Climate benchmarking using GNSS occultation. In *Atmosphere and climate.*, 287-301Springer.

Lohmann, Ulrike, and Johann Feichter. 1997. Impact of sulfate aerosols on albedo and lifetime of clouds: A sensitivity study with the ECHAM4 GCM. *Journal of Geophysical Research: Atmospheres* (1984–2012) 102 (D12): 13685-700.

Manabe, Syukuro, and Kirk Bryan. 1969. Climate calculations with a combined ocean-atmosphere model. *Journal of the Atmospheric Sciences* 26 (4): 786-9.

Matsuno, Taroh. 1971. A dynamical model of the stratospheric sudden warming. *Journal of the Atmospheric Sciences* 28 (8): 1479-94.

McCarthy, Mark P., Holly A. Titchner, Peter W. Thorne, SFB Tett, Leopold Haimberger, and DE Parker. 2008. Assessing bias and uncertainty in the HadAT-adjusted radiosonde climate record. *Journal of Climate* 21 (4).

Meehl, Gerald A., Curt Covey, Karl E. Taylor, Thomas Delworth, Ronald J. Stouffer, Mojib Latif, Bryant McAvaney, and John FB Mitchell. 2007. The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bulletin of the American Meteorological Society* 88 (9): 1383-94.

Meehl, Gerald A., Lisa Goddard, James Murphy, Ronald J. Stouffer, George Boer, Gokhan Danabasoglu, Keith Dixon, Marco A. Giorgetta, Arthur M. Greene, and Ed Hawkins. 2009. Decadal prediction can it be skillful? *Bulletin of the American Meteorological Society* 90 (10).

Menne, Matthew J., Imke Durre, Russell S. Vose, Byron E. Gleason, and Tamara G. Houston. 2012. An overview of the global historical climatology network-daily database. *Journal of Atmospheric & Oceanic Technology* 29 (7).

Miller, Gifford H., Scott J. Lehman, Kurt A. Refsnider, John R. Southon, and Yafang Zhong. 2013. Unprecedented recent summer warmth in arctic canada. *Geophysical Research Letters* 40 (21): 5745-51.

Morice, Colin P., John J. Kennedy, Nick A. Rayner, and Phil D. Jones. 2012. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *Journal of Geophysical Research: Atmospheres* (1984–2012) 117 (D8).

Moss, Richard H., Jae A. Edmonds, Kathy A. Hibbard, Martin R. Manning, Steven K. Rose, Detlef P. Van Vuuren, Timothy R. Carter, Seita Emori, Mikiko Kainuma, and Tom Kram. 2010. The next generation of scenarios for climate change research and assessment. *Nature* 463 (7282): 747-56.

Nakicenovic, Nebojsa, and Robert Swart. 2000. Special report on emissions scenarios. *Special Report on Emissions Scenarios*, Edited by Nebojsa Nakicenovic and Robert Swart, Pp.612. ISBN 0521804930. Cambridge, UK: Cambridge University Press, July 2000. 1 .

Pagano, Thomas S., Moustafa T. Chahine, and Eric J. Fetzer. 2010. The atmospheric infrared sounder (AIRS) on the NASA aqua spacecraft: A general remote sensing tool for understanding atmospheric structure, dynamics, and composition. Paper presented at Proc. of SPIE Vol, .

R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>

Randall, DA, RA Wood, S. Bony, R. Colman, and T. Fichefet. 2007. Co-authors. 2007. climate models and their evaluation.

Randel, William J., and Janel B. Cobb. 1994. Coherent variations of monthly mean total ozone and lower stratospheric temperature. *Journal of Geophysical Research: Atmospheres* (1984–2012) 99 (D3): 5433-47.

Randel, William J., Keith P. Shine, John Austin, John Barnett, Chantal Claud, Nathan P. Gillett, Philippe Keckhut, Ulrike Langematz, Roger Lin, and Craig Long. 2009. An update of observed stratospheric temperature trends. *Journal of Geophysical Research: Atmospheres* (1984–2012) 114 (D2).

Rignot, Eric, I. Velicogna, MR Van den Broeke, A. Monaghan, and JTM Lenaerts. 2011. Acceleration of the contribution of the greenland and antarctic ice sheets to sea level rise. *Geophysical Research Letters* 38 (5).

Rocken, C., R. Anthes, M. Exner, D. Hunt, S. Sokolovskiy, R. Ware, M. Gorbunov, W. Schreiner, D. Feng, and B. Herman. 1997. Analysis and validation of GPS/MET data in the neutral atmosphere. *Journal of Geophysical Research: Atmospheres* (1984–2012) 102 (D25): 29849-66.

Schmidt, Gavin A., Drew T. Shindell, Ron L. Miller, Michael E. Mann, and David Rind. 2004. General circulation modelling of holocene climate variability. *Quaternary Science Reviews* 23 (20): 2167-81.

Sherwood, S. C., J. R. Lanzante, and C. L. Meyer. 2005. Radiosonde daytime biases and late-20th century warming. *Science* (New York, N.Y.) 309 (5740) (Sep 2): 1556-9.

Sillmann, J., VV Kharin, X. Zhang, FW Zwiers, and D. Bronaugh. 2013. Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. model evaluation in the present climate. *Journal of Geophysical Research: Atmospheres* 118 (4): 1716-33.

Steiner, AK, G. Kirchengast, M. Borsche, U. Foelsche, and T. Schoengassner. 2007. A multi-year comparison of lower stratospheric temperatures from CHAMP radio occultation data with MSU/AMSU records. *Journal of Geophysical Research: Atmospheres* (1984–2012) 112 (D22).

Steiner, AK, G. Kirchengast, BC Lackner, B. Pirscher, M. Borsche, and U. Foelsche. 2009. Atmospheric temperature change detection with GPS radio occultation 1995 to 2008. *Geophysical Research Letters* 36 (18).

Steiner, AK, BC Lackner, F. Ladstädter, B. Scherllin-Pirscher, U. Foelsche, and G. Kirchengast. 2011. GPS radio occultation for climate monitoring and change detection. *Radio Science* 46 (6).

Stocker, Thomas F., Q. Dahe, and Gian-Kasper Plattner. 2013. Climate change 2013: The physical science basis. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Summary for Policymakers (IPCC, 2013).

Stouffer, Ronald Jay, Syukuro Manabe, and K. Ya Vinnikov. 1994. Model assessment of the role of natural variability in recent global warming.

Syndergaard, S. 1998. Modeling the impact of the earth's oblateness on the retrieval of temperature and pressure profiles from limb sounding. *Journal of Atmospheric and Solar-Terrestrial Physics* 60 (2): 171-80.

Taylor, Karl E., Ronald J. Stouffer, and Gerald A. Meehl. 2012. An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society* 93 (4).

Thorne, Peter W., John R. Lanzante, Thomas C. Peterson, Dian J. Seidel, and Keith P. Shine.

2011. Tropospheric temperature trends: History of an ongoing controversy. Wiley

Interdisciplinary Reviews: Climate Change 2 (1): 66-88.

Thorne, Peter W., David E. Parker, Simon FB Tett, Phil D. Jones, Mark McCarthy, Holly

Coleman, and Philip Brohan. 2005. Revisiting radiosonde upper air temperatures from 1958 to

2002. *Journal of Geophysical Research: Atmospheres* (1984–2012) 110 (D18).

Thorne, PW, DE Parker, BD Santer, MP McCarthy, DMH Sexton, MJ Webb, JM Murphy, M.

Collins, HA Titchner, and GS Jones. 2007. Tropical vertical temperature trends: A real

discrepancy? *Geophysical Research Letters* 34 (16).

Trenberth, KE, PD Jones, P. Ambenje, R. Bojariu, D. Easterling, A. Klein Tank, D. Parker, F.

Rahimzadeh, JA Renwick, and M. Rusticucci. 2007. Observations: Surface and atmospheric

climate change, chap. 3 of *climate change 2007: The physical science basis. contribution of*

working group I to the fourth assessment report of the intergovernmental panel on climate

change [solomon, S., qin, D., manning, M., marquis, M., averyt, KB, tignor, M., miller, HL and

chen, Z.(eds.)], 235–336.

Ware, R., C. Rocken, F. Solheim, M. Exner, W. Schreiner, R. Anthes, D. Feng, B. Herman, M.

Gorbunov, and S. Sokolovskiy. 1996. GPS sounding of the atmosphere from low earth orbit:

Preliminary results. *Bulletin of the American Meteorological Society* 77 (1): 19-40.

Wentz, Frank J., and Matthias Schabel. 1998. Effects of orbital decay on satellite-derived lower-

tropospheric temperature trends. *Nature* 394 (6694): 661-4.

Wickert, Jens, Christoph Reigber, Georg Beyerle, Rolf König, Christian Marquardt, Torsten

Schmidt, Ludwig Grunwaldt, Roman Galas, Thomas K. Meehan, and William G. Melbourne.

2001. Atmosphere sounding by GPS radio occultation: First results from CHAMP. *Geophysical Research Letters* 28 (17): 3263-6.

Wigley, TML, and SCB Raper. 1990. Natural variability of the climate system and detection of the greenhouse effect.

Zhai, Panmao, and Robert E. Eskridge. 1996. Analyses of inhomogeneities in radiosonde temperature and humidity time series. *Journal of Climate* 9 (4): 884-94.

Zou, Cheng-Zhi, Mei Gao, and Mitchell D. Goldberg. 2009. Error structure and atmospheric temperature trends in observations from the microwave sounding unit. *Journal of Climate* 22 (7).

———. 2009. Error structure and atmospheric temperature trends in observations from the microwave sounding unit. *Journal of Climate* 22 (7).

Zou, Cheng-Zhi, Mitchell D. Goldberg, Zhaohui Cheng, Norman C. Grody, Jerry T. Sullivan, Changyong Cao, and Dan Tarpley. 2006. Recalibration of microwave sounding unit for climate studies using simultaneous nadir overpasses. *Journal of Geophysical Research: Atmospheres* (1984–2012) 111 (D19).