Improving an Atmosphere General Circulation Model (AGCM) through parameter optimization

Dissertation

with the aim of achieving a doctoral degree
at the faculty of Mathematics, Informatics and Natural Sciences
Department of Geowissenschaften
of Universität Hamburg

submitted by
Reema Agarwal

Hamburg

2016

Day of oral defense: 17 January 2017

The following evaluators recommend the admission of the dissertation:

Prof. Dr. Detlef Stammer

Dr. Armin Köhl

Contents

Dedication

To my husband,

Neeraj Agarwal

Acknowledgement

I would like to express my sincere gratitude to my scientific adviser Prof. Dr. Detlef Stammer for providing me an opportunity to work in his group.

I also wish to thank my co-adviser Dr. Armin Köhl for his encouragement and many useful discussions. Many thanks to Prof. Dr. Jörn Behrens for chairing my advisory panel and his support in the extension of my funding period.

I thank Dr Ion Matei for helping me familiarize with the PlaSim model. Special thanks to my colleagues Iuliia Polkova and Xueyuan Liu for their support. The support from administration staff from the University of Hamburg was incredible. Special thanks to Frau Kutlu for her support specially when I went back to India.

I would like to thank THOR and NACLIM, a project of the European Union 7th Framework Programme (FP7 2007-2013) under grant agreement n.212643 and n.308299 for financial support. I am also thankful to School of Integrated Climate Sciences (SICSS) for providing courses and giving me the opportunity to expand my knowledge about the Earth System. I am grateful to central IT service (CIS) for timely printing of my thesis.

My parents have always supported my decisions and work. This work is a result of their blessings, love, encouragement and good wishes with which they brought me up. My sincere gratitude to Neeraj, who has been extremely supportive and caring towards me through all these years. I am extremely grateful to my kids Nishita and Naman for their support and love. I feel really blessed because of love and support from my entire family.

Finally I would like to thank God, the almighty, for giving me strength, tolerance and patience that helped me to complete this work.

Abstract

This thesis presents an implementation and evaluation of two different multivariate data assimilation techniques for the optimization of parameters of a global primitive equation Atmospheric General Circulation Model (AGCM), the Planet Simulator (PlaSim). The hypothesis used is that the source of uncertainty in the model is related to parameters from the cloud parameterizations, vertical and horizontal diffusion time scales in the model. The results are evaluated by comparing basic physical state variables of the atmosphere such as surface temperature, precipitation, net heat flux, winds and sea level pressure predicted by the model with observations.

Initially, sensitivity analysis of PlaSim with respect to various parameters used in its different parameterizations is carried out. The variation of the cost function with respect to changes in each control parameter is studied and the most sensitive parameters are identified. The results of the sensitivity analysis serve as a guideline for identifying sensitive model parameters optimization procedures.

Green's function (GF) method of parameter optimization is applied on two different model configurations, with and without moisture related processes (*wet* and *dry* configurations, respectively) in an identical twin model framework. The results are inter-compared with existing results from 4D-variational (4D-var) assimilation scheme in PlaSim. GF procedure successfully estimates model parameters for both shorter (30 days) and longer time (365 days) scales using 3, 5 and 6 control parameters. However, when using real world observations, the GF method is unable to minimize cost function even using a single control parameter.

Another optimization procedure based on stochastic approximation, the simultaneous perturbation stochastic approximation (SPSA) method is implemented in PlaSim and the results are discussed. The advantage of using SPSA method is its ease of implementation and its robustness to noise in cost function. In identical twin experimental framework SPSA method reliably recovers the control parameters. When real observations are used, the errors in optimized state, for example in surface temperature and net heat flux are reduced by 16% and 30% respectively. In addition, the optimized state of PlaSim shows improvement in sea level pressure, zonal winds at 500 hPa and surface precipitation. This study demonstrates the usefulness of a simple data assimilation scheme in a highly non-linear chaotic system and its potential application in tuning of the climate models.

Chapter 1

Introduction

1.1 Climate models

Climate models are a numerical representation of the earth system that are widely used as a tool to understand the observed climate changes. The components of an earth system model are Atmosphere, Ocean, Land and cryosphere. These models are used to diagnose whether the changes are due to natural variability, human activity or a combination of both. Development in climate models is essential so as to predict the accurate state closer to known state of Earth's climate system. Climate models are used for a variety of applications such as understanding the past and present climate system and for predicting future state of the climate of the entire earth system. The projections of future climate is one of the major components of the Inter-governmental Panel for Climate Change (IPCC) reports and its interpretation influences the decisions of national, regional and local importance (Houghton et al., 1997).

Climate projections based on numerical models are biased as compared to the real world and these biases (errors) could be of significant concern (Schwartz et al., 2002). The errors in numerical climate models can be due to several reasons. They could be either due to model's resolution or from parametrization of processes unresolved at grid scale (for example cloud formation), (Murphy et al., 2004). Some of the major sources of errors in climate models are :

- The nonlinear system of the Navier-Stokes equation results in chaotic, unpredictable system (atmosphere in particular) in contrast with linear system. In addition, various assumptions used to formulate and discretized model equations can lead to significant model errors.

- Model configuration, for example the resolution. Sub-grid scale processes are not captured by a coarse resolution model and result in errors that grow with time.
- Uncertainty in the initial conditions which can affect the quality of short term forecast (seasonal to inter-annual). Small uncertainties in today's weather conditions prescribed to climate models as initial conditions may lead to large forecast errors.
- Uncertainties in parameterization schemes used in models can lead to significant errors in forecast and climate projections. In accurate model parameters used in the parameterization of several sub-grid scale processes (for example clouds and turbulence) are the major cause of this uncertainty. These parameters are basically derived from empirical relationships that are developed on a basis of few observations that are poorly sampled in both space and time. Global application of such parameters is one of the biggest sources of model uncertainty.

For meaningful and accurate climate predictions, an important task is to reduce the errors and bring the model closer to observations. By increasing the horizontal and vertical resolution model simulations can be improved (Roeckner et al., 2006), however increasing resolution results in increase in computational costs. Improvement in the atmospheric initial conditions tend to improve short to medium range forecast errors, however their influence on climate scale predictions is negligible. Improved parameterization of sub-grid scale processes (like convection, clouds) play an important role to improve climate simulations and affect climate sensitivity (Iacobellis et al., 2003).

1.2 Data Assimilation

Climate models are corrected to bring them closer to the real world observations by using data assimilation techniques (Daley, 1993; Kalnay, 2003; Malanotte-Rizzoli, 1996). Data assimilation tend to reduce model errors, regardless of the cause of these errors such as numerical and physicals schemes or model parameters (as discussed in the previous section). This procedure involves a combination of the observational data and the physical processes that govern the system to be optimized, (Robinson and Lermusiaux, 2000). Data assimilation procedures are mainly used to:

- provide better initial conditions in order to get a better forecast.

- produce a model state which is consistent with observations.
- provide estimates of model parameters used in parameterization.

Immense progress has been made in the field of data assimilation in numerical models since last three decades. Most of the efforts in data assimilation are taking place in the Numerical weather prediction (NWP) community, where observations are assimilated as and when available to improve the operational weather forecast. This type of continuous data assimilation systems are meant to correct the initial conditions of a short to medium range forecasting system. While improving initial conditions through data assimilation, help in short to medium range forecasting, longer time scale simulations are barely affected by the quality of initial conditions (mainly the atmosphere). Model parameters used in different parameterizations play a vital role in controlling model drifts from observed states during the course of long term simulations. Hence it is necessary to accurately prescribe these model parameters. This procedure of correcting model parameters and hence improving model's sub grid scale parameterizations by means of data assimilation is known as parameter optimization. Parameter optimization plays a major role in building a realistic numerical climate model which can be used to make reliable climate predictions. Using observed climate to tune model parameters in a climate model is a complex process.

There have been several efforts to include model parameters into data assimilation as control variables using four-dimensional (4-D Var) or ensemble Kalman filter (EnKF) methods (Wunsch, 1996; Anderson, 2001; Annan et al., 2005; Sugiura et al., 2008; Evensen, 2009). These techniques have been developed and implemented both in AGCMs and Ocean General Circulation Models (OGCMs). The successful implementation of optimization procedures on these models is one of the main concerns as the the manual tuning of through trial and error requires intensive work and computational cost is proportional to number of control parameters. Therefore an automatic optimization procedure is required to find the best estimate of a set of parameters.

Parameter estimation on climate scale simulations has been a challenge, particularly in atmospheric models. The problem becomes even more complex when multiple parameters are used as control variables for optimization. While many studies have shown success in estimating parameters in synthetic observation framework, there are difficulties in applying these techniques to the real world problems. Annan et al. (2005) implemented and tested EnKF (Ensemble Kalman Filter) in an AGCM and found that few control parameters were successfully retrieved in identical

twin experiments. Application using reanalysis data resulted in an improvement in temperature fields whereas precipitation remained unaffected. Schirber et al. (2013) found that the estimated parameter values lead to an overall error reduction on short timescales, the error of the model state increases on climatological timescales. Center for Earth System Research and Sustainability (CEN) at the University of Hamburg has developed a 4-D var based coupled data assimilation system for their earth system model (CESAM). CESAM consists of an atmospheric model (component of the Planet Simulator, PlaSim) coupled to the Massachusetts institute of technology general circulation model (MITgcm). The assimilation procedure (Blessing et al., 2014) is an iterative minimization of cost function through variation of control parameters (θ) by a Quasi Newton algorithm (Fletcher and Powell, 1963) which determines the search direction through gradient of cost function with respect to control parameters. The gradient is calculated by automatic differentiation (Griewank et al., 1989) of the source code through transformation of algorithm in fortran (TAF) (Giering and Kaminski, 1998). The performance of assimilation scheme used is limited by the time window in case of the atmosphere model. However switching off the non-linear processes helped in getting good results.

Some optimization algorithms are based on gradient approximation from measurements of cost function. The main advantage of using such algorithms is that they do not require detailed knowledge of functional relationship between the parameters being optimized and the cost function being minimized that is required in gradient based algorithms. This category includes simulated annealing (Jackson et al., 2004; Gonzalez et al., 2007; Lakshmanan and Derin, 1989), Monte Carlo methods (Lewis and Bridle, 2002), genetic algorithms (Yao and Sethares, 1994) and statistical estimation algorithms (Altaf et al., 2011; Menemenlis et al., 2005). These algorithms have been successfully used for optimization in chaotic systems. As pointed in above discussion that parameter optimization in non-linear, chaotic systems is a difficult problem. Variational assimilation techniques applied on climate models sometimes fail to yield impressive results because of its inherent non-linear characteristics. In these circumstances, stochastic optimization and statistical estimation algorithms are quite useful because of their ability to cope with inherent system noise and high nonlinearities. This thesis aims to build an alternate assimilation system to optimize parameters of PlaSim, which is the atmospheric component of CESAM. The study is funded by European Union research project NACLIM (https://www.cen.uni-hamburg.de/en/research/coordinating-projects/naclim-seite.html).

1.3 Goal of the thesis and approach

The goal of this thesis is to test an alternate assimilation system for optimizing model parameters used in different physical parameterizations of PlaSim. In this study two optimization algorithms; Green's function (GF) based on statistical estimation and Simultaneous perturbation stochastic approximation (SPSA) based on stochastic optimization will be implemented and tested with PlaSim at longer integration time scales. The basic motive is to evaluate the potential of these two approaches in optimizing model parameters to get a physical state of the atmosphere which is more closer to the observations as compared to the state obtained using the control configuration of PlaSim. Another objective of this study is to evaluate and discuss the advantages/disadvantages of parameter estimation relative to the existing adjoint based approach of data assimilation in PlaSim. This study is an attempt to explore the possibility of applying less complex data assimilation procedure to a highly non-linear chaotic system in an effective manner to reduce model errors.

1.3.1 Approach

In order to achieve the above mentioned goals, following approach has been adopted in this thesis:

- Study PlaSim's sensitivity to model parameters used in different parameterization schemes so that the parameters which must be used as control variables in pilot optimization approaches can be identified.

- Experiments with GF based parameter estimation approach using both *dry* and *wet* configuration of PlaSim in an identical twin framework and comparison of its performance relative to the existing 4D-var assimilation method in PlaSim.

- Experiments with SPSA based optimization procedure applied on PlaSim.

The feasibility of parameter estimation using these techniques will be first tested in an identical twin framework where the basic assumption will be that the main source of model uncertainties are the biased parameters. Once this is done, then these optimization procedures will be tested using real data.

1.4 Structure of the thesis

The thesis has been divided into seven chapters (including introduction which is chapter 1).

Chapter 2 gives a detailed description of the atmosphere model (PlaSim) used in the study. The chapter further discusses model's control run and data sets used in the study. The results of comparison of PlaSim's control run with European Center for medium-range weather forecasts reanalysis (ERA-interim) are also discussed.

Chapter 3 briefly describes three optimization techniques that are used in the thesis: Green's function, Adjoint and SPSA .

Sensitivity experiments performed with PlaSim are presented in chapter 4. The sensitivity of the model is based upon variations in the cost function with respect to the perturbations applied to model parameters. Impact of model parameters on basic state variables like temperature, precipitation, winds and sea level pressure (SLP) is also analyzed. The results from sensitivity analysis are useful to provide information on model parameters that could be chosen for optimization.

Chapter 5 discusses the application of GF on PlaSim. An inter comparison of the Adjoint and the GF for the *dry* model configuration is made. Further, the results of optimization for the *wet* model configuration using the GF are shown in this chapter.

Chapter 6 is based on the application of SPSA on PlaSim. The optimization algorithm assimilates pseudo as well as ERA-Interim reanalysis data for integration time of 1 year. An improvement in optimized state is compared with original state of model in this chapter.

Finally the conclusions from the thesis are compiled in Chapter 7.

Chapter 2

Model description

In this study, an earth system model of intermediate complexity, PlaSim (Fraedrich et al., 2005), is used. PlaSim, mainly an atmospheric model, has all the components of an earth system model such as sea-ice, ocean, land surface and biosphere, however these component models are of reduced complexity[1]. It is a σ-coordinate (where σ is the pressure normalized by the surface pressure), primitive equation model based on the basic laws of conservation of momentum, mass and energy. The equations are solved using the spectral transform method (Orszag, 1970) and integrated in time using a leap-frog semi-implicit time stepping scheme (Hoskins and Simmons, 1975; Simmons et al., 1978) with Robert/Asselin time filter (Robert, 1981; Asselin, 1972). Major unresolved processes of boundary layer fluxes, diffusion, radiation, moisture and clouds are included by simplified parametrization.

This model has been widely used in research involving timescales ranging from seasonal to inter-annual and up to climate scales. Several climate related studies have been performed using PlaSim, for example, the impact of Atlantic and Pacific ocean sea surface temperature anomalies on the north Atlantic multidecadal variability (Grosfeld et al., 2008), thermodynamics of climate change (Lucarini et al., 2010), the effect of global warming and global cooling on the distribution of the latest Peiman climate zones (Roscher et al., 2011), effect of mountains and ice sheets on global ocean circulation (Schmittner et al., 2011). Although, a complete description of the atmosphere model is available in the model document[2], a brief overview of the model's equation and the configuration that has been used in this study is provided

[1]The complete model including sources and documentation is available at (www.mi.uni-hamburg.de/plasim)

[2]http://www.mi.uni-hamburg.de/fileadmin/files/forschung/theomet/planet_simulator/downloads/PS_ReferenceManual.pdf

in the following sections.

2.1 Model equations

The horizontal coordinates of the model are φ (representing latitude) and λ (representing longitude). The temperature is linearized about a reference profile T_0, hence, prognostic equation for temperature deviations $T' = T - T_0$ are used where $T_0 = 250K$ is constant for all σ levels.

Model equations are moist primitive equations which are dimensionless set of differential equations obtained by scaling of

- vorticity ζ and divergence D with respect to angular velocity of the Earth Ω.
- pressure p by mean sea level pressure $p_s = 101325Pa$
- temperature T and T' by $\frac{a^2\Omega^2}{R}$ and
- Orography and geopotential (Ψ) by $\frac{a^2\Omega^2}{g}$ (where g is the acceleration due to gravity and R is the gas constant for dry air).

The dimensionless equations in (λ,μ,σ)- coordinates (Hoskins and Simmons, 1975) are as shown below

Conservation of momentum leading to equations for for vorticity (ζ) and divergence (D)

$$\frac{\partial \zeta + f}{\partial t} = \frac{1}{(1-\mu^2)}\frac{\partial F_v}{\partial \lambda} - \frac{\partial F_u}{\partial \mu} + P_\zeta \tag{2.1}$$

$$\frac{\partial D}{\partial t} = \frac{1}{(1-\mu^2)}\frac{\partial F_u}{\partial \lambda} + \frac{\partial F_v}{\partial \mu} - \nabla^2 E - \nabla^2(\phi + T_0 l p_s) + P_D \tag{2.2}$$

Hydrostatic approximation

$$0 = \frac{\partial \phi}{\partial ln\sigma} + T \tag{2.3}$$

Conservation of mass (continuity equation) leading to pressure equation

$$\frac{\partial lnp_s}{\partial t} = -\int_0^1 A d\sigma \tag{2.4}$$

and the Thermodynamic equation

$$\frac{\partial T\prime}{\partial t} = F_T - \dot{\sigma}\frac{\partial T}{\partial \sigma} + \kappa W T + \frac{J}{c_p} + P_T \tag{2.5}$$

The notations used above are

$$F_u = (\zeta + f)V - \dot{\sigma}\frac{\partial U}{\partial \sigma} - T\prime\frac{\partial ln p_s}{\partial \lambda}$$
$$F_v = -(\zeta + f)U - \dot{\sigma}\frac{\partial V}{\partial \sigma} - (1-\mu^2)T\prime\frac{\partial ln p_s}{\partial \mu}$$
$$F_T = -\frac{1}{(1-\mu^2)}\frac{\partial (UT\prime)}{\partial \lambda} - \frac{\partial (VT\prime)}{\partial \mu} + DT\prime$$
$$E = \frac{U^2 + V^2}{2(1-\mu^2)}$$
$$\dot{\sigma} = \sigma\int_0^1 A d\sigma - \int_0^\sigma A d\sigma$$
$$W = \frac{\omega}{p} = V.\triangledown ln p_s - \frac{1}{\sigma}\int_0^\sigma A d\sigma$$
$$A = D + V.\triangledown ln p_s = \frac{1}{p_s}\triangledown .p_s V$$

Here $\dot{\sigma}$ is the vertical velocity in the σ system. Diabatic heating per unit mass and kinetic energy per unit mass are given by J and E respectively. For velocity, the stream function (ψ) is the nondivergent part and the velocity potential (χ) is the irrotational part.

The model includes parametrization of river runoff, soil hydrology and soil and land temperatures. Although a simple zero layered thermodynamic sea ice model is included in the model, it is not used in this particular study. Maximum albedo for sea ice is set to 0.7 while the minimum and maximum albedo for glaciers are 0.6 and 0.8, respectively. and that for snow are 0.4 to 0.8 respectively. The albedo of water is set to 0.069, (Romanova et al., 2006). The model setup used in the thesis is explained in next section.

2.2 Model configuration and control run

2.2.1 Model configuration

Two different configurations of PlaSim, namely the *wet* and the *dry* have been used in this study. In the *wet* configuration, PlaSim includes all major processes and

components except the land-biosphere model, while in the *dry* configuration moisture related processes are excluded, i.e. there is no precipitation, evaporation and runoff. In addition, the soil moisture is set to climatology there is no cloud radiative feedback.

For all experiments in this study, the horizontal resolution of PlaSim is set to spectral T21 (approximately 5.6° grid spacing) with 64 points in longitude and 32 points in latitude. There are 10 vertical sigma levels and an integration time step of 45 minutes. This configuration is same as used by Blessing et al. (2014) in their experiments. There were a couple of reasons for choosing this setup:

- Being a coarse resolution model the experiments, specially the sensitivity runs (which were around 3200 model years) could be performed in a reasonable time.

- The results of 4D-var experiments using the same configuration were available at hand that would allow an inter-comparison with the results of this study.

2.2.2 Control run

Beginning from a cold start, a forward run of PlaSim was carried out for 25 years using monthly averaged climatological sea ice concentration (sea ice model is switched off), sea surface temperature (SST), glacier coverage, albedo and soil temperature from ERA40 reanalysis (Uppala et al., 2005) as boundary conditions. A restart file was recorded at the end of 25 years and this restart which corresponds to 1st January is used as initial condition in all the experiments. A forward simulation is then made for a period of one year using the initial and boundary conditions mentioned above and the default values of various parameters used in model's parameterization scheme (Table 4.1). This simulation is referred to as the *control* run and is used as a reference in all the experiments. The control run is referred to as model's reference state, i.e. the state before parameter optimization. Although we mentioned two different configurations of PlaSim in Section 2.2, in most of the cases, the *wet* configuration is used (except when exclusively mentioned otherwise).

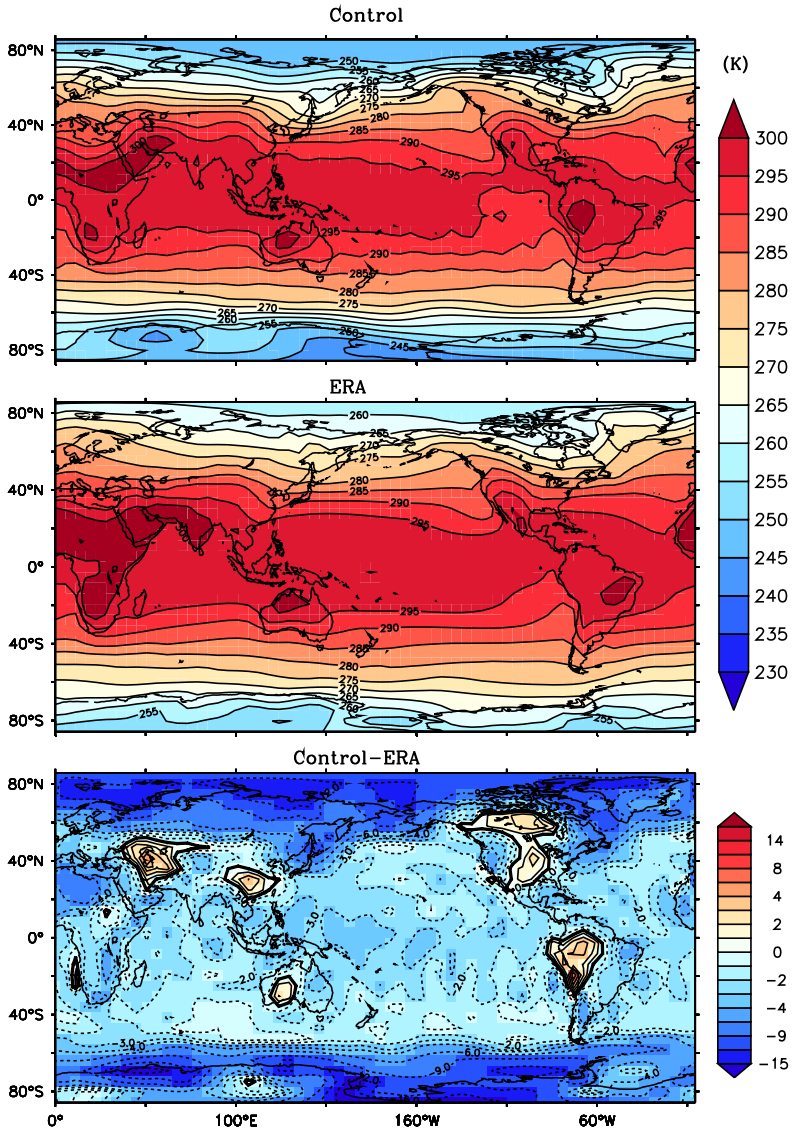

Figure 2.1— Plot of Annual mean temperature (K) at the lowest model level from (top) control run, (middle) observations and (bottom) difference (control-observations). Control run represents the annual mean state resulting from one year run discussed in section 2.2. Observations represents annual mean of 21 years (1989-2009) from ERA-Interim reanalysis.

2.3 Comparison of the control run with ERA-Interim reanalysis

It is of interest to compare the model's reference state with observations in order to get the information about the model biases. For this purpose, a comparison

of the annual mean temperature at model's lowest level, sea level pressure, zonal winds at 500 hPa and surface precipitation simulated from the control run (1 year period as discussed in section 2.2) is made with annual mean ERA-Interim reanalysis (referred to as observations hereafter). ERA-Interim reanalysis (Dee et al., 2011) is available from the European center for medium range weather forecasting (ECMWF; http://data-portal.ecmwf.int/) and it covers period from 1979 onwards overlapping the earlier ECMWF 40 year reanalysis. The data is available on a native 0.7°x 0.7°grid and is an output of an atmospheric model with an advanced 4D-Var data assimilation technique. For this study, annual mean observations are constructed using 21 years of data (from 1989 to 2009) and the values are averaged on to the model grid to compute model data difference.

Figure 2.1 shows the comparison of annual mean temperature at the lowest model level (which is near the surface). Overall patterns and range of temperature distribution in the control run (top panel) agree well with the observations (middle panel). However from the difference control - observations plot (bottom panel), it can be clearly seen that PlaSim in its original state in general has colder temperature with respect to observations. Exceptions are the regions over middle east, eastern China, south-west Australia, most parts of central North America and over north west region over South America. In tropics the difference is around -2 to -4 K, however in polar regions the model is far too cold ($\sim$ -10K to -15K), specially over the Arctic Ocean.

Surface precipitation patterns shown in Figure 2.2 reveal that PlaSim (top panel) has a poorly resolved Inter-tropical Convergence Zone (ITCZ) that appears to be too much diffused. This might be attributed to the coarse resolution used in our setup. However, the east-west gradient in precipitation patterns in the Pacific are qualitatively similar to observations (Figure 2.2, middle panel). The differences in precipitation are between ± 6 mmday^{-1} (Figure 2.2, bottom panel). PlaSim overestimates precipitation (> 4 mmday^{-1}) mainly over east Africa, Arabian Sea, Peninsular Indian region and East China. Tropical East Indian Ocean and west Pacific oceans receive less rains in the control run (< 4 mmday^{-1}) as compared to the observations. Precipitation in the regions of prominent ITCZ between central and east Pacific, just above the equator is also underestimated in the model.

Model simulated SLP is shown in Figure 2.3 (top panel). In the Northern Hemisphere the patterns and magnitude of low and high pressure systems (in the subtropical, sub-polar Pacific and Atlantic) are consistent with the observations (Figure 2.3, middle panel). However, in the southern hemisphere, south of 40°S, the simulated

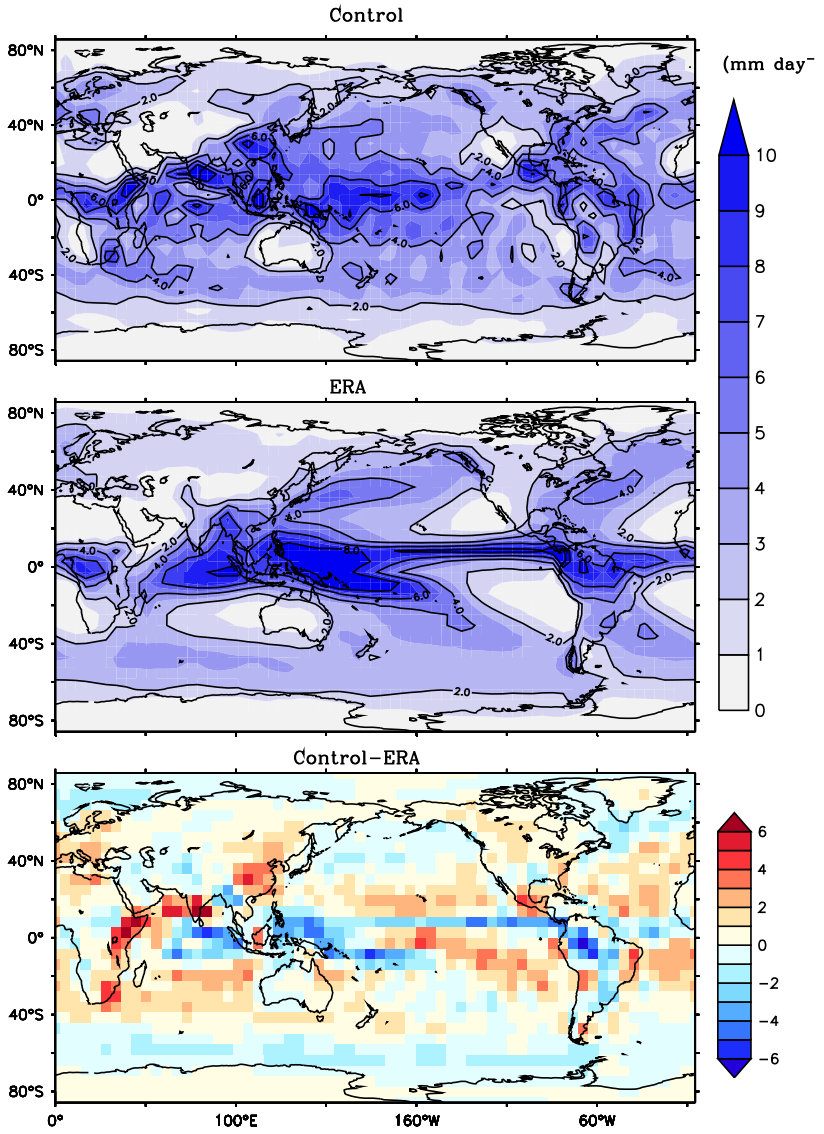

Figure 2.2— same as in Figure 2.1 but for total precipitation ($mmday^{-1}$)

pressure fields are highly underestimated (Figure 2.3, bottom panel). These are the regions of very strong zonal winds, and the difference in SLP in these regions is quite high ($>$ 16 hPa).

500 hPa zonal winds from the model (Figure 2.4, top panel) show a well developed westerly flow in both the hemispheres and the locations of maximum magnitude are consistent with those in observations (Figure 2.4, middle panel). The difference plot between model and observations (Figure 2.4, bottom panel) reveal weaker ($<$

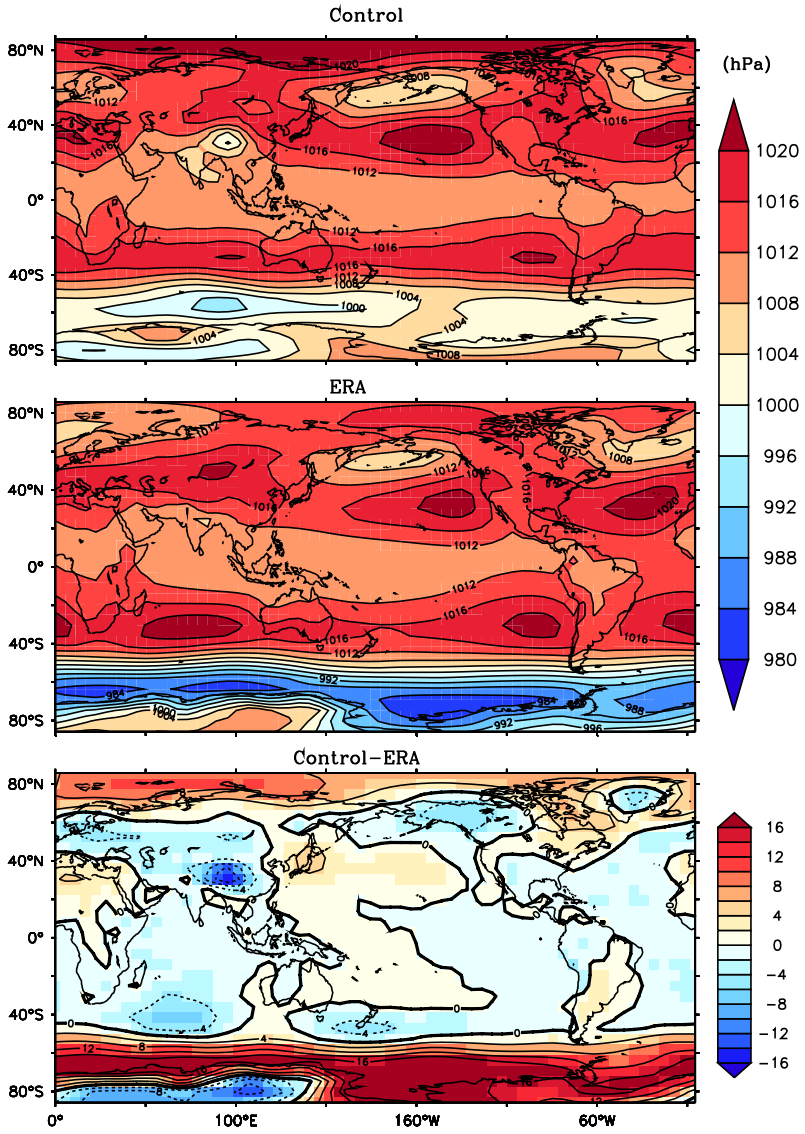

Figure 2.3— same as in Figure 2.1 but for Sea Level Pressure (hPa)

5-10 ms^{-1}) zonal winds between 60°S and 45°S mainly in the Indo Pacific region. Correspondingly between 40°S and 20°S , winds are stronger ($>$ 3-16 ms^{-1}) in the control run. Apart, from these and barring a few regions in the northern hemisphere between 30°N and 60°N, the differences in 500 hPa zonal winds between model and observations are less than 4 ms^{-1}.

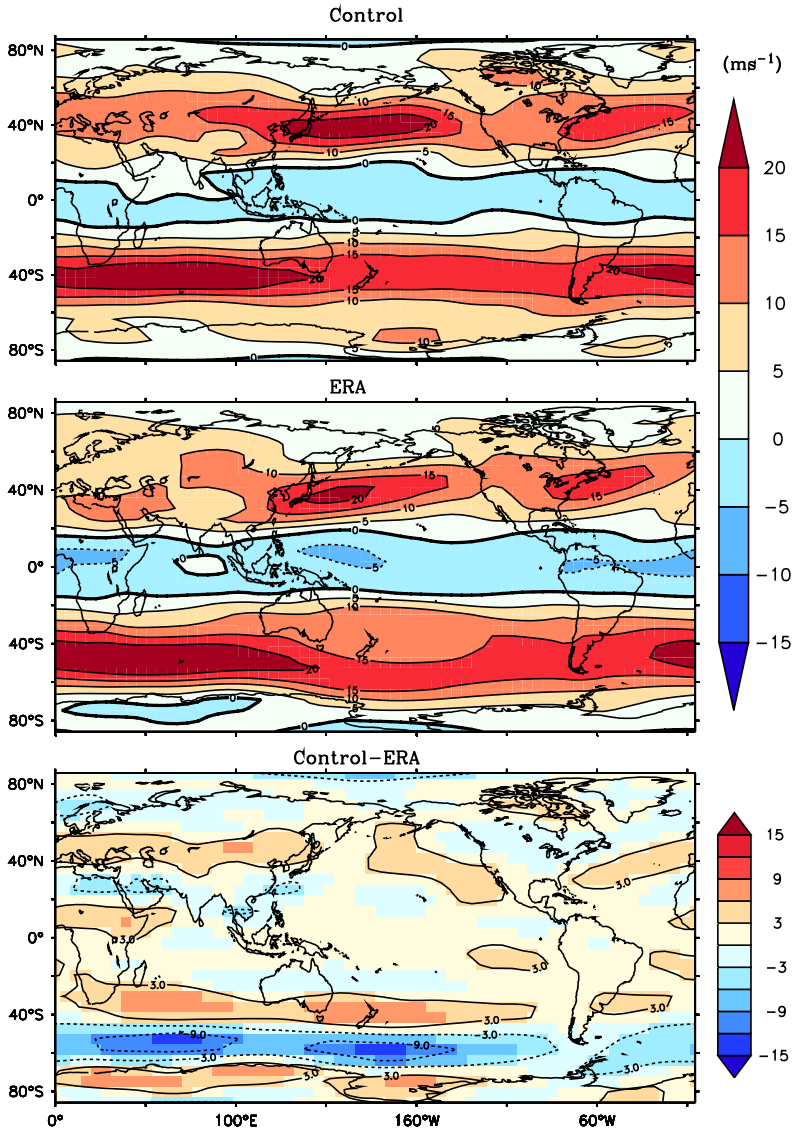

Figure 2.4— same as in Figure 2.1 but for zonal winds(ms^{-1})at 500 hPa

2.4 Summary

This chapter provided details of the numerical model (PlaSim) used in this study. The setup, initial and boundary conditions and configurations used in the control run were discussed. A qualitative evaluation of basic variables simulated from the control run was performed with respect to the ERA-Interim reanalysis and differences were pointed out. Overall, PlaSim simulates colder surface temperatures, weaker

zonal winds in the Southern Ocean and a poorly resolved ITCZ. It is seen from the above comparison that there are some prominent regions of disagreement between the control run and the observations, specially in the southern hemisphere and in the Arctic. At this moment, it is quite difficult to identify the real cause of these differences. Not all the differences are caused due to errors in model parametrization as previous studies (Roeckner et al., 2006) have shown the dependence of model performance on grid resolution. Nevertheless it will be interesting to find out how much impact does the model parameters have in these regions. In the next chapter, the effect of model parameters on PlaSim simulations will be examined through sensitivity experiments performed at different timescales using dry and wet configurations. This will give insight into the magnitude and patterns of model differences due to errors in model parameters.

Chapter 3

Optimization Techniques

Data assimilation system comprises of three components: a numerical model , reference data (observations) and a assimilation technique. Several data assimilation techniques are available for parameter optimization in numerical models. The difference between the various methods is the way in which the model skill is calculated to measure distance between model and observations and also corresponding weights given to observations and prior estimates. Another difference is the choice of parameters that are allowed to be adjusted in order to produce the final estimate. A brief overview of various assimilation techniques was provided in Section 1.2 This chapter discusses the optimization methods that will be used in this study.

3.1 Green's function approach

The Green's function (GF) approach follows the discrete inverse theory (Menke, 1989) which is used to solve inverse problems. The model's Green functions are calculated by perturbing model parameters one at a time and are used to linearize the GCM. Later, the discrete inverse theory is used to estimate uncertain model parameters of a GCM. The aim is to estimate model parameters from a combination of model and observations. The GF method was successfully applied on an Ocean General Circulation Model (OGCM; Menemenlis et al. (2005); Menemenlis and Wunsch (1997); Stammer and Wunsch (1996)) and to optimize sea ice and ocean model parameters of a coupled ocean-sea ice model (Nguyen et al., 2011). The GF has also been used for solving geophysical inverse problems (Wunsch, 1996; Gloor et al., 2001; Fan et al., 1999; Challis and Sheard, 2003). The conventional methods of adjusting model parameters involve trial and error sensitivity experiments where the perturbed parameters are used repeatedly until optimum state of the model is

obtained. This is a time consuming approach and requires lot of experiments and patience. Menemenlis et al. (2005) successfully demonstrated GF approach for optimizing a small set of parameters in an OGCM. The GF technique used in this study briefly discussed (following Menemenlis et al. (2005)) as below:

Let an AGCM be written as discrete non linear equations:

$$\boldsymbol{x^f}(t_{i+1}) = M_i[\boldsymbol{x^f}(t_i)] \tag{3.1}$$

The state vector $\boldsymbol{x^f}(t_i)$ includes temperature, humidity, vorticity and wind velocity on a predefined grid at discrete time t_i for i = 0,...,N-1. M is a non linear operator representing the advancement in the model state from time t_i to time t_{i+1}. It includes known parameters describing the system, initial and boundary conditions that serve as an input to the model and drive the system.

Let the true state of the ocean $\boldsymbol{x^t}$ on discrete grid be represented by equation 3.1 plus a random noise in uncertain model parameters, $\boldsymbol{\theta}$ having zero mean and covariance $\boldsymbol{Q}$:

$$\boldsymbol{x^t}(t_{i+1}) = M_i(\boldsymbol{x^t}(t_i), \boldsymbol{\theta}) \tag{3.2}$$

The observations are related to the model state by the equation:

$$\boldsymbol{y^{obs}}(t) = P(t, \boldsymbol{x(t)}) + \boldsymbol{\epsilon} \tag{3.3}$$

where;
t = 0,.....,N
$\boldsymbol{y^{obs}}(t)$ is a vector of all available observations at time t
$P(t, \boldsymbol{x}(t))$ is non linear operator that includes transformations and grid interpolations.
$\boldsymbol{\epsilon}$ is a vector of noise process having zero mean and covariance R.
To find the best estimates of the control parameters $\boldsymbol{\theta}$ for the model to fit the observations $\boldsymbol{y^{obs}}$, an objective function (Y) is defined as :

$$Y = \boldsymbol{\theta^T Q^{-1} \theta} + \boldsymbol{\epsilon^T R^{-1} \epsilon} \tag{3.4}$$

Model Green's function is a vector comprising of model simulations using perturbed parameters (denoted by $\boldsymbol{\theta}$) relative to the baseline integration (control run). The Green's function $\boldsymbol{G}$ is then used to project back the simulated response on to the model state vector (Menemenlis and Wunsch, 1997). Green's function $\boldsymbol{G}$ relates observations to the model parameters by using Equations (3.2) and (3.3):

$$\boldsymbol{y^{obs}}(t) = \boldsymbol{G[\theta]} + \boldsymbol{\epsilon} \tag{3.5}$$

Several optimization algorithms (Adjoint, EnKF) are based on linearization of model dynamics and minimize Y with respect to the model. Green's function method makes use of linear assumption and this leads to the simplification of equation (3.5) to:

$$\boldsymbol{y^d} = \boldsymbol{y^{obs}} - \boldsymbol{G[0]} = \boldsymbol{G\theta} + \boldsymbol{\epsilon} \tag{3.6}$$

where 0 is null vector, i.e. the set of parameters without perturbation whose corresponding state is the control run, $\boldsymbol{y^d}$ is model data difference The j^{th} column of matrix $\boldsymbol{G}$ is

$$\boldsymbol{g}_j = \frac{G(\boldsymbol{e}_j) - G(\mathbf{0})}{e_j} \tag{3.7}$$

where $\boldsymbol{e_j}$ represents the vector of perturbation applied to each model parameter. G(0) is baseline integration. In GF method, the AGCM's sensitivity experiments are performed by perturbing the parameters represented by vector $\boldsymbol{\theta}$ and the matrix $\boldsymbol{g}$ is constructed which is the difference between the perturbed and the baseline run relative to the size of perturbation. Optimum parameters are then estimated using the linear inverse theory by minimizing the model data differences. The minimization of equation (3.4) and (3.6) is a discrete linear inverse problem with solution:

$$\boldsymbol{\theta}^a = \boldsymbol{H}\boldsymbol{G}^{\boldsymbol{T}}\boldsymbol{R}^{-\mathbf{1}}\boldsymbol{y^d} \tag{3.8}$$

where $\boldsymbol{H}$ is the uncertainty covariance matrix given by,

$$\boldsymbol{H} = (\boldsymbol{Q}^{-\mathbf{1}} + \boldsymbol{G}^{\boldsymbol{T}}\boldsymbol{R}^{-\mathbf{1}}\boldsymbol{G})^{-\mathbf{1}} \tag{3.9}$$

Generally θ^a gives estimates of the model parameters. These estimates are further used to calculate new parameter values at each iteration in a nonlinear system. In the GF approach, each model parameter is perturbed one at a time, and corresponding measurements of cost function (Y) are obtained. Each component of gradient estimate matrix is formed by taking the difference of the corresponding cost function (Y) values and then dividing by a difference interval which is the corresponding change in one of the parameter. The gradient is a vector of the p partial derivatives which is same as number of control parameters. The aim is to reduce the cost function by searching an optimal set of parameters in the search direction.

At each iteration, cost function, gradients are evaluated until Euclidean norm of the control parameters and gradients are minimized.

3.2 4D-var assimilation method (Adjoint)

The 4D-VAR assimilation method was implemented with the PlaSim model by Blessing et al. (2014). In this implementation, adjoint of the PlaSim required to calculate the gradients of the cost function was constructed with the help of the source transformation tool Transformation of Algorithms (Giering and Kaminski 1998, TAF). The assimilation consists of an iterative minimization of a cost function through variation of control parameters (θ) by a Quasi Newton algorithm (Fletcher and Powell, 1963). This procedure determines the search direction through the gradient of a cost function with respect to the control parameters. In the case of PlaSim, the gradient of the cost function with respect to control parameter is calculated by automatic differentiation (Griewank et al., 1989) of the source code through TAF (Giering and Kaminski, 1998). The adjoint model is a very efficient tool to compute the gradient of the cost function. For linear models, it can be shown that it is equivalent to the optimal Kalman Filter solution under the hypothesis of a perfect model (Daley, 1991; Lorenc, 1986). In the nonlinear case and if the model is assumed to be perfect, the variational data assimilation is successful if the tangent-linear hypothesis is valid for the duration of the assimilation period.

Lea et al. (2000, 2002) demonstrated that variational parameter estimation with an adjoint model does not work well for tuning of chaotic models due to their sensitive dependence on initial conditions. In a study by Annan and Hargreaves (2007) it is shown in the limit of infinite time scales, gradients of the cost function become undefined therefore on those time-scales gradient based estimation of parameters become impractical.

3.3 Parameter estimation using SPSA technique

SPSA is essentially a randomized Kiefer Wolfowitz method where the gradient is estimated using only two evaluations per iteration regardless of the dimension of the optimization problem (Spall, 1998). The evaluations are calculated based on two simultaneous perturbations. The SPSA algorithm has gathered a great deal of interest over last decade and has been used in variety of applications (Altaf et al., 2011; Hutchison and Hill, 2001; Spall, 2000; Gerencsér and Vágó, 2001; Kong et al., 2011).

As a result of the stochastic perturbation the calculated gradient is also stochastic, however the expectation value of the stochastic gradient is the true gradient (Li and Reynolds, 2011).

The gradient based algorithms converge faster than any gradient approximation (SPSA algorithm) when speed is measured in terms of the number of iterations. However, the total cost to achieve effective convergence depends not only on the number of iterations required but also on the cost needed to perform these iterations, which is large in gradient based algorithms. This cost may include computational cost and additional human effort required for coding the gradients.

The details of SPSA are given by Spall (1998); however the technique is briefly discussed here. In SPSA the cost function $\mathrm{Y}(\hat{\theta})$ is minimized using the iterative procedure

$$\hat{\theta}_{k+1} = \hat{\theta}_k - a_k \hat{g}_k(\hat{\theta}_k) \tag{3.10}$$

where $\hat{g}_k(\hat{\theta}_k)$ is a stochastic approximation of $\nabla Y(\hat{\theta}_k)$, which denotes the gradient of the cost function with respect to $\hat{\theta}$ evaluated at previous iterate $\hat{\theta}_k$. The stochastic gradient $\hat{g}_k(\hat{\theta}_k)$ in SPSA is calculated by

$$\hat{g}_k(\hat{\theta}_k) = \frac{Y(\hat{\theta}_k + c_k\Delta_k) - Y(\hat{\theta}_k - c_k\Delta_k)}{2c_k\Delta_k} \tag{3.11}$$

Δ_k is a p-dimensional perturbation vector generated by the Monte Carlo method where each of the p components of Δ_k are independently generated from a zero mean probability distribution. Here Δ_k is sampled from a Bernoulli $\pm$ distribution with probability of $\frac{1}{2}$ for each ± 1 outcome.

The gain sequences (a_k and c_k) are defined as

$$a_k = \frac{a}{(A+k+1)^\alpha}, c_k = \frac{c}{(k+1)^\gamma} \tag{3.12}$$

The choice of gain sequences (a_k and c_k) is critical for the performance of SPSA. It is observed that in a high noise setting it is necessary to pick smaller a and larger c than in a low noise setting. To attain optimal efficiency of the algorithm, the choice of step size a_k is crucial. The step size is flexible which can be used to make the algorithm work efficiently for highly non linear models. The asymptotically optimal values of α and γ are 1.0 and 1/6 respectively (Fabian, 1971; Chin, 1997). As a rule of thumb (with the Bernoulli ± 1 distribution for the elements of Δ_k), it is effective

to set c at a level approximately equal to the standard deviation of the noise in the model cost function Y (equation 4.2). The SPSA method begins with an initial set of parameter values . Two model simulations are performed by perturbing the parameters. The cost functions are calculated and the gradients are approximated according to equation (3.11). Further, parameter values are corrected so that the algorithm progressively moves towards the parameter values where cost function is minimized.

SPSA technique is very easy to implement because it needs only two cost function evaluations independent of number of model parameters.

3.4 Summary

Three data assimilation techniques are briefly discussed in this chapter. These techniques are iterative algorithms with aim of reducing model-data difference at each iteration and obtain estimates of parameters. The choice of measure of distance between model and observations and also corresponding weights given to observations and prior estimates in these algorithms is very crucial. 4D-var data assimilation technique (Adjoint) is gradient based algorithm. In Adjoint method gradients gives information to determine the search direction at every iteration of the algorithm. In stochastic optimization algorithms: GF and SPSA, gradients are approximated and there is a random choice in the search direction at each iteration towards the solution. The optimization results of these techniques are discussed in Chapter 5 and 6.

Chapter 4

Sensitivity Analysis of the model

4.1 Sensitivity Analysis

A sensitivity analysis (SA) is usually performed as a series of tests in which values of parameters used in various parametrization schemes are changed and their effect on the model is studied. By studying the model's response to changes in parameter values, one can get useful information for the model development as well as model evaluation. SA helps to understand the effect of uncertainties that are often associated with parameters in the models. In this respect, SA allow us to determine which model parameters are responsible for creating maximum change of the simulated model state. It gives an idea of what range of perturbation in parameters is corresponding to errors in the model and this information can be useful in model development. Experimenting with a wide range of values of model parameters can offer insights into model's behavior.

In order to estimate parameters of the model, prior knowledge about impact of parameter changes on model is important, therefore, performing SA is inevitable. The change in model parameters also give an idea as to how much effect these parameters have in model simulations, particularly on climate scales. In this chapter, results of these sensitivity experiments are presented. Sensitivity of PlaSim is studied from the behavior of model's cost function with respect to perturbation in model's control parameters. Cost function is an estimate of the model data misfit and therefore its magnitude in the sensitivity experiments can be used to know the positive or negative impact of model parameters on model's performance.

4.2 Control parameters

In an AGCM, the errors in parameters in different physical parametrization schemes contribute to the model errors. The choice of model's control parameters in optimization plays a crucial role. The model is not equally sensitive to all parameters. Some parameters have large impact on the climate of the model while others have a moderate or less effect. The idea of performing sensitivity experiments is to find out those parameters whose variations have huge impact on the model simulations.

According to Intergovernmental panel on Climate Change (IPCC) first assessment report most of the climate uncertainties are due to representation of cloud radiative feedbacks. Cloud and radiation parameters are commonly chosen in optimization procedures (Annan et al., 2005; Severijns and Hazeleger, 2005; Mauritsen et al., 2012) as they are mainly responsible for accurate temperature and precipitation simulations. Recently, Blessing et al. (2014) used diffusion time scale related parameters in their coupled model optimization experiments using the 4D-var method. In this study, 15 control parameters are selected for SA. These selected parameters are listed in Table 4.1 along with their definitions and their default values which were coded in the model. The control variables selected are parameters used in the parameterization of long wave, short wave radiation and clouds, time scale for Rayleigh friction in the uppermost two atmospheric layers, the diffusion time scales for divergence, vorticity and temperature. The choice of these parameters is similar to those used by Blessing et al. (2014), although there are many more parameters in the model that can be considered for optimization . Since one of the objectives of this study is to inter-compare the results of optimization with those of the 4-D var method applied on PlaSim, so it was thought to use same control parameters as were used in experiments with 4-D var method.

The parameters tfrc1 and tfrc2 are rayleigh friction timescales in level 10 and level 9 respectively, while, tdissz, tdissd and tdisst represents time scale of damping in the atmosphere and are used in horizontal diffusion parametrization. For calculation of mixing lengths, drag and transfer coefficients, the parameters vdiff_lamm, vdiff_b, vdiff_c and vdiff_d are used. The parameters tpofmt, th2oc, tswr1, tswr2, tswr3 and acllwr are used in cloud parametrization. A high value of th2oc increases the absorption of radiation by the moisture in the atmosphere leading to an increase in temperature, which results in more evaporation and hence increases moisture in the air. A small value of parameter tswr1 increases cloud albedo and leads to more precipitation. For a balance of precipitation and evaporation these two parameters

Nr	*Acronym*	*Control parameter*	*units*	*Default Value*
1	tfrc1	rayleigh friction timescale	days	20
2	tfrc2	rayleigh friction timescale	days	100
3	tdissz	diffusion time scale for vorticity	days	1.1
4	tdissd	diffusion time scale for divergence	days	0.2
5	tdisst	diffusion time scale for temperature	days	5.6
6	vdiff_lamm	tuning parameter for vertical diffusivity	m	160
7	vdiff_b	tuning parameter for vertical diffusivity	m	5
8	vdiff_c	tuning parameter for vertical diffusivity	m	5
9	vdiff_d	tuning parameter for vertical diffusivity	m	5
10	tpofmt	tuning of point of mean transimittivity in layer	_	1
11	th2oc	absorption coefficient for h2o continuum	_	0.04
12	tswr1	tuning of cloud albedo range	_	0.04
13	tswr2	tuning of cloud back scattering range2	_	0.048
14	tswr3	tuning of cloud scattering albedo range2	_	0.004
15	acllwr	mass absorption coefficient for clouds	_	0.1

Table 4.1— List of 15 control parameters used in optimization procedure

should be consistent. The parameters tswr2, tswr3 and acllwr are responsible for cloud back scattering.

4.3 Methodology for sensitivity experiments

4.3.1 Cost function (model data misfit)

In this study, it is assumed that the difference between data and simulation results is only due to measurement errors and incorrectly prescribed model parameters. The measure of model skill, is based on a cost function which is defined as the weighted measure of the mean squared difference between the model and the observations. Mathematically, cost function for N different fields d (for example, temperature, precipitation etc.) and model predictions M(θ) at N points and normalized by the inverse of a N X N model-data covariance matrix C_d^{-1}. C_d represents uncertainties in the observations and the model. C_0 contains information about error statistics of the background field. The problem of estimation is then solved by minimizing the cost function (Y) with respect to the parameters $\hat{\theta}$

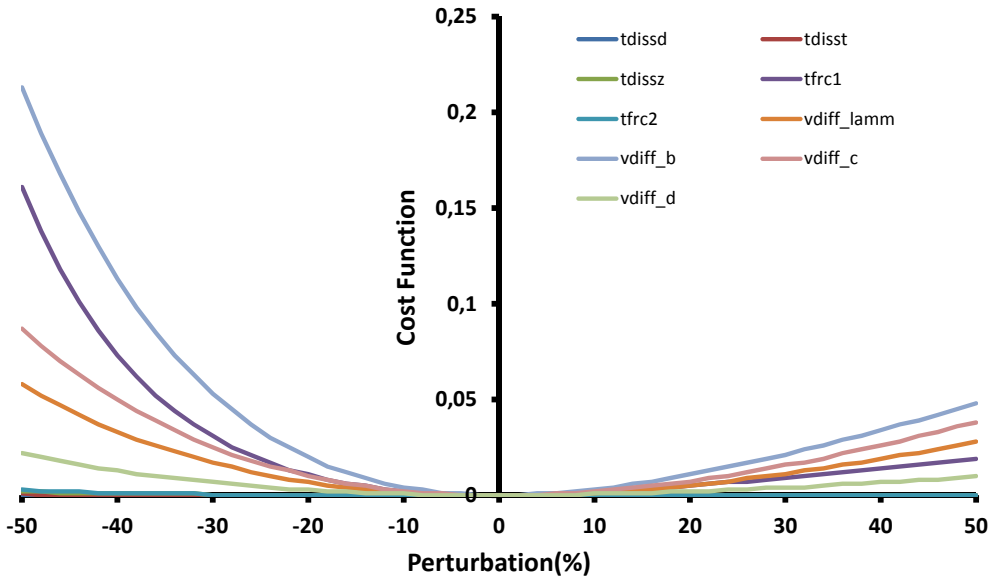

Figure 4.1— Plot of model cost function (Y) vs. perturbation applied to model control parameters (θ, 9) in the *dry* configuration of PlaSim. The cost function (Y) is computed from pseudo data set and an averaged cost function is plotted for a period of 7 days.

$$Y(\hat{\theta}) = 0.5[(M(\theta) - d)^T C_d^{-1}(M(\theta) - d) + (\theta - \theta_0)^T {C_0}^{-1}(\theta - \theta_0)] \tag{4.1}$$

where θ_0 is model's background state vector. θ is model's original state. C_0 represents uncertainties in the model respectively. In equation 4.2, θ is a vector of model parameter values and 'T' indicates the matrix transpose.

In this entire study the cost function computation includes the contribution of 16 model variables which are listed in Table 4.2. These are temperature at all model levels, surface fluxes, radiation fields and precipitation. The Table 4.2 also shows the corresponding standard deviations of each of these variables computed from 34 years of ERA-Interim reanalysis data set spanning period 1979-2012.

For SA experiments are performed in the identical twin framework, i.e. the cost function is computed with respect to the output that is generated by the model itself (called pseudo observations). Pseudo observations are a result of the control run which is discussed in section 2.2. In this run the control parameters used in various parameterization (in our case 15 parameters) are set to their default values (listed in table 4.1). In equation 4.2, the first term shows the model data misfit and

Datafield	*STD*	*Units*
Temperature (level 10(surface))	0.657	K
Temperature (level 9)	0.640	K
Temperature (level 8)	0.616	K
Temperature (level 7)	0.578	K
Temperature (level 6)	0.564	K
Temperature (level 5)	0.568	K
Temperature (level 4)	0.724	K
Temperature (level 3)	0.984	K
Temperature (level 2)	1.212	K
Temperature (level 1)	1.302	K
Large scale precipitation (LSP)	$3.390\text{x}10^{-9}$	ms^{-1}
Convective precipitation (CP)	$5.007\text{x}10^{-9}$	ms^{-1}
Surface Sensible Heat Flux (SSHF)	3.155	Wm^{-2}
Surface Latent Heat Flux (SLHF)	6.502	Wm^{-2}
Surface Solar Radiation (SSR)	5.054	Wm^{-2}
Surface Thermal Radiation (STR)	2.490	Wm^{-2}

Table 4.2— Variables contributing cost function computation. The second column list standard deviation (STD) that are globally averaged values computed from annual averaged data sets spanning (1979-2012). The data set used to compute standard deviation is ERA-Interim data.

second term represents the prior information (background term). As mentioned in section 2.2, the total number of horizontal grid points in the model are 64*32 = 2048, therefore in our experiments of one year we will have the number of pseudo observations = 2048*16*365 = 11960320. Since the number of observations are much larger than the total number of control parameters, the background term in equation 4.2 can be safely dropped. Therefore in equation 4.2 the background term is set to zero i.e. prior information on the control variables is not considered in this study. This approach is similar to the one used in for computing the cost function in Menemenlis et al. (2005) and Blessing et al. (2014). Revised equation 4.2 is now expressed as

$$Y(\hat{\theta}) = 0.5[(M(\theta) - d)^T C_d^{-1}(M(\theta) - d)] \tag{4.2}$$

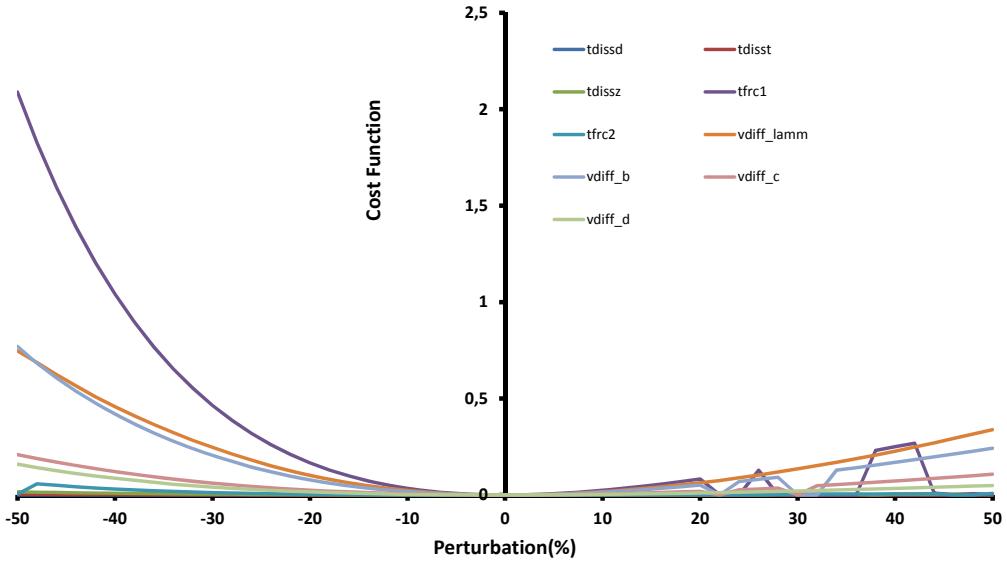

Figure 4.2— Same as Figure 4.1 except for 30 day integration period.

The pseudo observations used to compute the cost function for SA comprises the contribution from 16 variables involving temperature (all 10 levels), precipitation and heat flux . Two sets of pseudo observations are generated from the control run. One set corresponds to the *dry* configuration of the model while the other to the *wet* configuration. Both configurations were described in section 2.2. The model data differences are time averaged over the assimilation window and the cost function is evaluated at the end of the assimilation window.

4.4 Sensitivity experiments and results

4.4.1 Sensitivity analysis with the *dry* configuration of PlaSim

In the *dry* configuration of PlaSim, there is no hydrological cycle and all moisture related processes are excluded. Hence the parameters related to cloud parametrization are not used. These parameters are tpofmt, th2oc, tswr1, tswr2, tswr3 and acllwr. Therefore in this case, SA of PlaSim is performed with the remaining 9 control parameters. The main purpose of these experiments with dry configuration is to compare the assimilation results with those obtained from the 4D var adjoint method. The results of the 4D var were available only for the dry configuration of the model.

In the first experiment, model integrations are performed by perturbing the above mentioned 9 control parameters one by one. The range of perturbation which is applied to each parameter varies from -50% to 50% (with an interval of 2%) of the default control parameter value. The model is integrated forward in time for 7 seven days and the time averaged cost function corresponding to each perturbation is evaluated with respect to pseudo-observations. This resulted in a total of 51x9 = 455 simulations. Figure 4.1 shows the plot of cost function with respect to perturbation applied to the control parameter. Different line colors indicate the control parameters that are used in the experiments. The variation in the cost function due to changes applied to the control parameters is very smooth. This is mainly due to two reasons, one the period of integration is very small (7 days) and second the non-linearity doesn't set in during this small integration period as the moisture related processes are not active (dry configuration). Larger perturbation applied results in larger deviations of the model from pseudo-observations which can be seen in relatively higher values of the cost function. Reduction in parameter values (negative perturbations) causes more deviations in the model. The maximum effect of perturbations comes from the tuning parameter for vertical diffusivity (vdiff_b) and the parameter for rayleigh friction timescale (tfrc1). It is concluded from Figure 4.1 that the cost function has a no local minima.

In the second experiment, with the same set of perturbed parameters as used in the first experiment, model integrations are performed over 30 days. Here too 455 model runs were carried out corresponding to different control parameters range. The time averaged cost function is then plotted against the parameter perturbations (Figure 4.2). It can be seen from a couple of cases (when tfrc1 and vdiff_b are perturbed) that the cost function has multiple minima.

The range of the cost function is higher for 30 days as compared to 7 days. This is consistent with the fact that as the integration period increases, the model deviates from pseudo observations and hence errors become larger. It can also be seen from Figures 4.1 and 4.2 that for 7 days the cost function is most sensitive to parameters vdiff_b and vdiff_c while for 30 days tfrc1 is the most influential parameter when its value is reduced relative to its default value. From these two experiments we can say that for shorter time scales, the model is quite linear in its *dry* configuration and there is a unique minimum of the cost function and this is one important requirement for the success of many assimilation algorithms. This result may change when all parameters are perturbed at a time.

4.4.2 Sensitivity analysis with the *wet* configuration of PlaSim

In the *wet* configuration of PlaSim, the hydrological cycle is included and all moisture related processes are active in the model. Therefore the parameters used in cloud parameterization are included as control parameters in the SA. In this case, SA is performed using a total of 15 control parameters. Once again, as in section 4.4, two set of experiments are performed, for 7 days and for 30 days. Each set consisted of 51*15 = 765 model integrations. All data are time averaged over the assimilation window and the cost function is evaluated against pseudo observations.

Figure 4.3 shows the cost function vs. parameter plot for the 7 day model run. The fluctuations in cost function with respect to change in parameter values are quite large and therefore the function is not very smooth. The range of the cost function goes up to 60. Unlike dry model configuration in which the cost function plot was very smooth , the wet configuration of PlaSim due to several non-linear processes involved results in large variations in the model cost function. On 7 day timescale the cost function is most sensitive to water vapor absorption coefficient (th2oc) and parameter related to tuning of albedo due to cloud scattering (tswr3).

For 30 day integration the cost function's fluctuations are even larger (Fig 4.4)

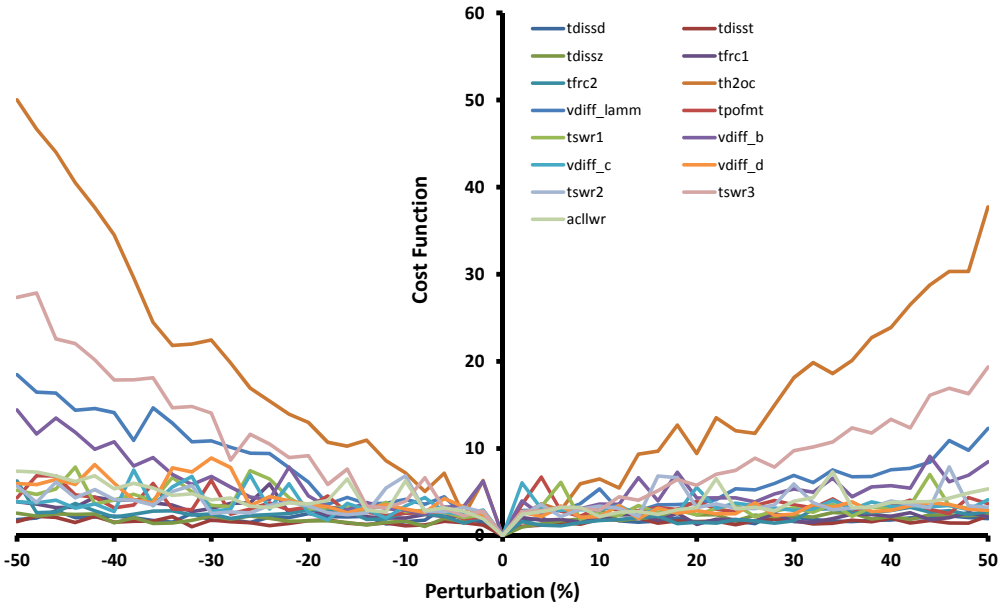

Figure 4.3— Same as Figure 4.1 except for the *wet* configuration of PlaSim and for 15 control parameters.

with the range of the cost function reaching up-to 140 corresponding to large perturbations in water vapor absorption coefficients (th2oc). Other parameters that have large influence on model on 30 day integration period are tuning parameters for vertical diffusivity (vdiff_lamm and vdiff_b).

From Figures 4.3 and 4.4, it can be seen that the cost function has a parabolic shape with several secondary minima.

4.5 Model response to parameter perturbation on longer timescales

SA for model runs for shorter time-scales in previous two sections was mainly performed keeping in mind comparison of the optimizations techniques with 4-D var adjoint assimilation scheme, results for which were available for shorter timescales and for dry configuration of PlaSim. Since one of the objectives of the thesis is to test and to implement the Green's function and the SPSA approach into PlaSim for longer time scales, therefore, SA with the *wet* configuration for 365 days is also carried out. In this case, the control parameters are perturbed from -90% to 500% (with an interval of 10%) of the original parameter value and the model is integrated for a period of 1 year. This entire process involved 61*15 = 915 model runs each

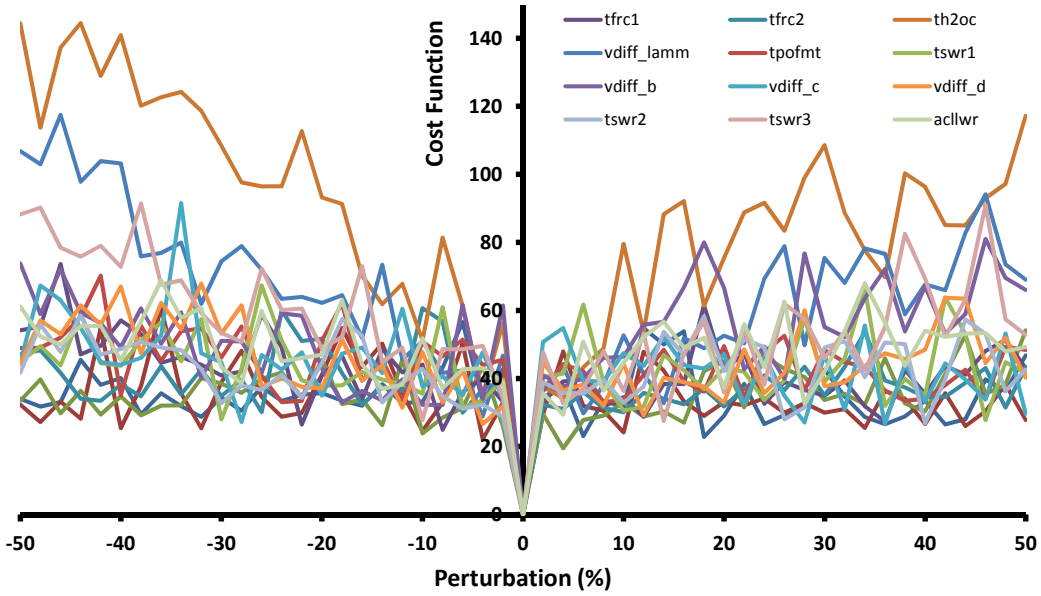

Figure 4.4— Same as Figure 4.2 except for the *wet* configuration and for 15 control parameters.

of 1 year length. The cost function is computed for the averaged values of variables for the entire year and is plotted with respect to perturbation range (Figure 4.5) of parameters. Consistent with the previous results, here too the range of cost function increases further with integration time.

The cost function is most sensitive to parameters tswr3 and th2oc (see Figures 4.3, 4.4 and 4.5). From Figure 4.5, it is visible that the cost function is not smooth and several secondary minima exist. This behavior is quite similar to what is seen in Figure 4.4. This is an indicator of nonlinear behavior of the model cost function with respect to perturbation in each parameter. Due to the non-linearity of an atmosphere system, for longer integration time scales, several secondary minima exist. These secondary minima are no longer resolved and appear as stochastic deviations and therefore gradient descent methods fail to find the absolute (true) minima. Instead of the true minimum, the optimization gets caught in a nearby minima or a range of minimum cost function (Köhl and Willebrand, 2002).

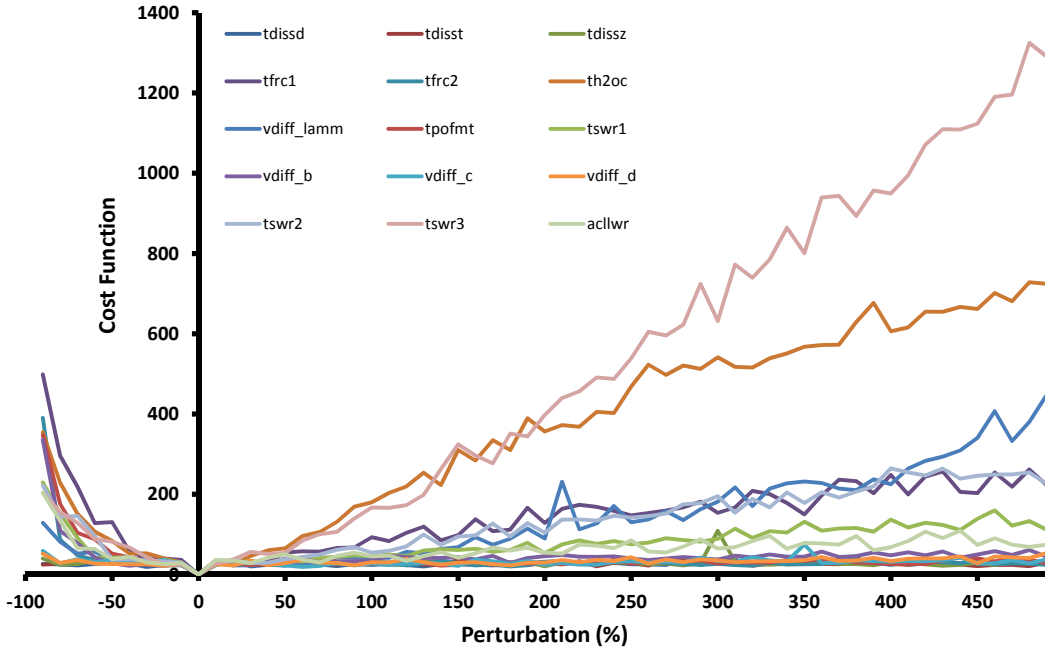

Figure 4.5— Plot of model cost function (Y) vs. perturbation applied to model control parameters (θ, 15) in the *wet* configuration of PlaSim. The cost function is computed from pseudo data and is based on annual mean values of 1 year model integration.

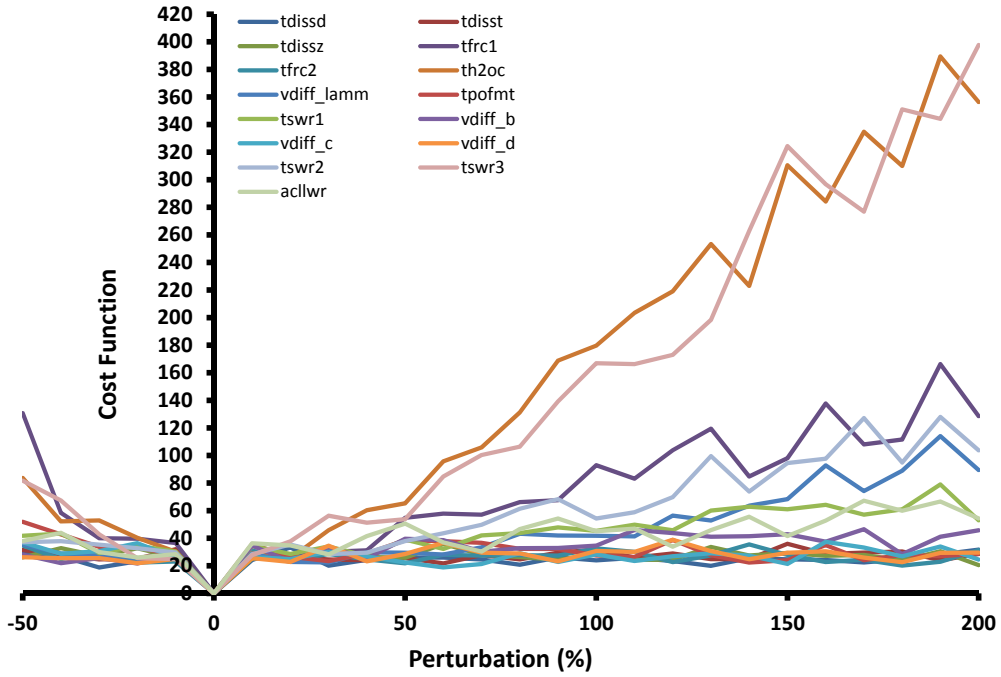

Figure 4.6— Same as Figure 4.5 except the range of perturbation from -50% to 200%. This plot clearly shows the range of minimum cost function that can achieved if parameter perturbations are not precisely zero.

These sensitivity experiments gives an idea of the range of perturbation that each parameter should be perturbed so as to get their reasonable estimate by cost function minimization. The range of true minimum of cost function in case of identical twin experiment can be identified from the sensitivity plot. In our case the range of cost function values corresponding to perturbation in parameters vary between 0 and 400 shown in Figure 4.6 which is the magnified plot of Figure 4.5. The range of perturbation in Figure 4.6 is -50% to 200%. From the Figure it is clear that the range of minimum cost function that can be achieved if parameter perturbations are not precisely zero is between 20 and 30. Therefore, it can be said that the best estimates of parameters in case of identical twin experiments would result in this range of cost function.

4.5.1 Impact of model parameters on state variables

Up to now, in this chapter, the focus was on the sensitivity of the cost function to model parameters. The cost function is a measure of the cumulative errors in the model variables and therefore the above SA does not provide any information on the behavior of individual model variables. Hence, it will be of interest to know the consequence of parameter perturbation on individual model variables in different regions. In this subsection, the regional impact of perturbation to model parameters is investigated on basic atmospheric state variables. This particular analysis provides an information about the parameters that might be crucial in the regions where PlaSim has errors (as discussed in the previous chapter). The aim is not to discuss much about the physical processes related to each parameter but to identify the regions which have the maximum impact due to perturbation in each of these parameters.

For this purpose, the sensitivity runs discussed in section 4.5 are analyzed. There are 60 realizations corresponding to each of the 15 parameters. The basic state variables of temperature at 1000 hPa, surface precipitation, SLP, zonal winds at 500 hPa, surface humidity and net heat flux are considered. Each of these variables are averaged for the entire period of 365 days, therefore providing 60 values for each model parameter. Root Mean Square error (RMSE) for each variable is computed at every model grid point with respect to the control run over those 60 simulations and the results are discussed. The RMSE is computed as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (X_i - X_c)^2}{N - 1}} \tag{4.3}$$

where N denotes the total number of model runs performed for each parameter. In this case N = 60. X_i denotes the state variable from the i^{th} run and X_c denotes the state variable from the control run.

Figure 4.7 shows the RMSE for temperature at 1000 hPa. The 15 panels shown correspond to each model parameter that is perturbed. Most prominently for most of the parameters the major influence of perturbation is seen over land particularly over Northern Russia, North America, Sahara region and Brazil and the errors are in the range of 0.5 to 3.5 K in these regions. Over the oceans the major impact of perturbation is seen in the Arctic, north of 40°N in Pacific and Atlantic and south of 50°S in the Southern Ocean. In tropical oceans, there is almost no impact seen from most of the diffusion and viscosity related parameters except from vdiff_lamm, which has some influence in the tropical west Pacific Ocean. The parameter vdiff_lamm

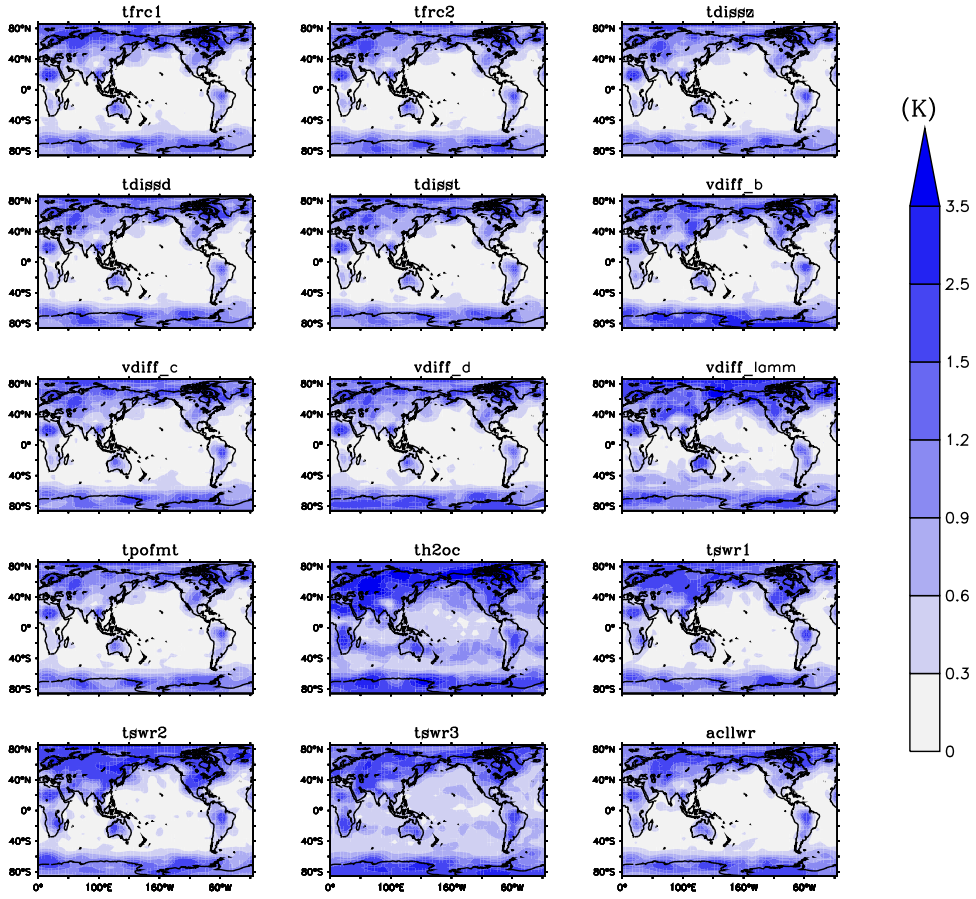

Figure 4.7— Plot of root mean square error (RMSE) in annual mean temperature (K) at 1000 hPa due to perturbations applied to 15 model parameters with respect to the control run.

also affects the regions south of 40°S. Cloud and moisture related parameters (mainly th2oc and tswr3) have influence over the oceanic regions, although weaker as compared to the influence on land. Overall, in temperature at 1000 hPa, the maximum sensitivity seems to come from th2oc, tswr3 and vdiff_lamm.

The effect on surface precipitation (Figure 4.8) from all parameters is mainly seen in the Northern Indian Ocean, north east China, the Indian subcontinent and to some extent in the Western Pacific. Cloud and moisture related parameters (th2oc, tswr1, 2 and 3) have an impact on the entire ITCZ region. In addition to this, perturbations in vdiff_lamm influences north west and central Pacific between 30°N and 50°N and also in the narrow band along 40°S mainly in the Atlantic and the Indian Ocean. Overall, the range of errors in surface precipitation due to the applied

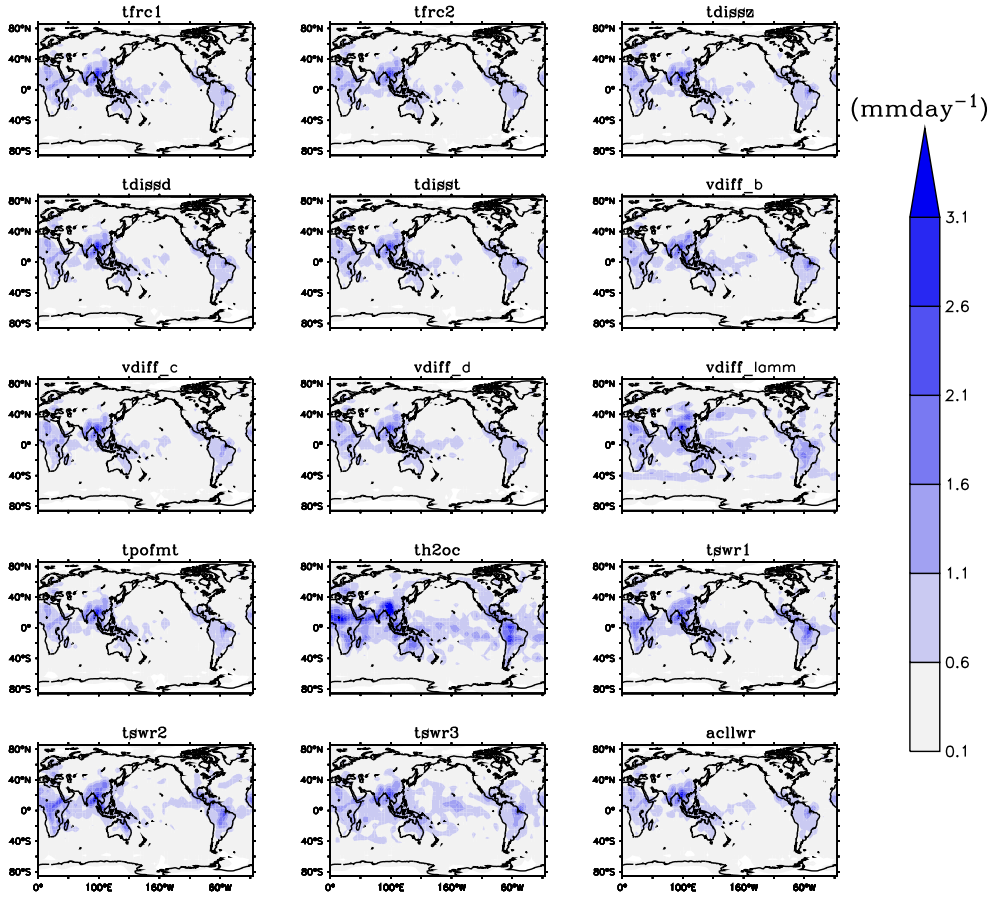

Figure 4.8— same as in figure 4.7 but for surface precipitation (mmday^{-1}).

perturbations is between 0.1-3 mmday^{-1}. Interestingly, these regions coincide with the regions where the surface precipitation had errors with respect to observations (seen in chapter 2).

For SLP (Figure 4.9) the sensitive regions are mainly outside the tropical regions and coincide with the regions having major SLP changes, i.e. North Pacific, Arctic, North Atlantic and the entire Southern ocean. Parameter vdiff_lamm has the strongest influence (errors $>$ 5 hPa) mainly in the north Pacific between 40°N and 60°N and in Southern Ocean south of 40°S in strong winds regime. Sensitivity is also observed in the Indian Ocean and in the central Pacific between 10°N-30°N. There are some regions in the central tropical Pacific between 140°W and 120°W which are sensitive to th2oc. Remaining parameters have an almost similar domain of influence, albeit with varying magnitudes of changes (between 1-5 hPa).

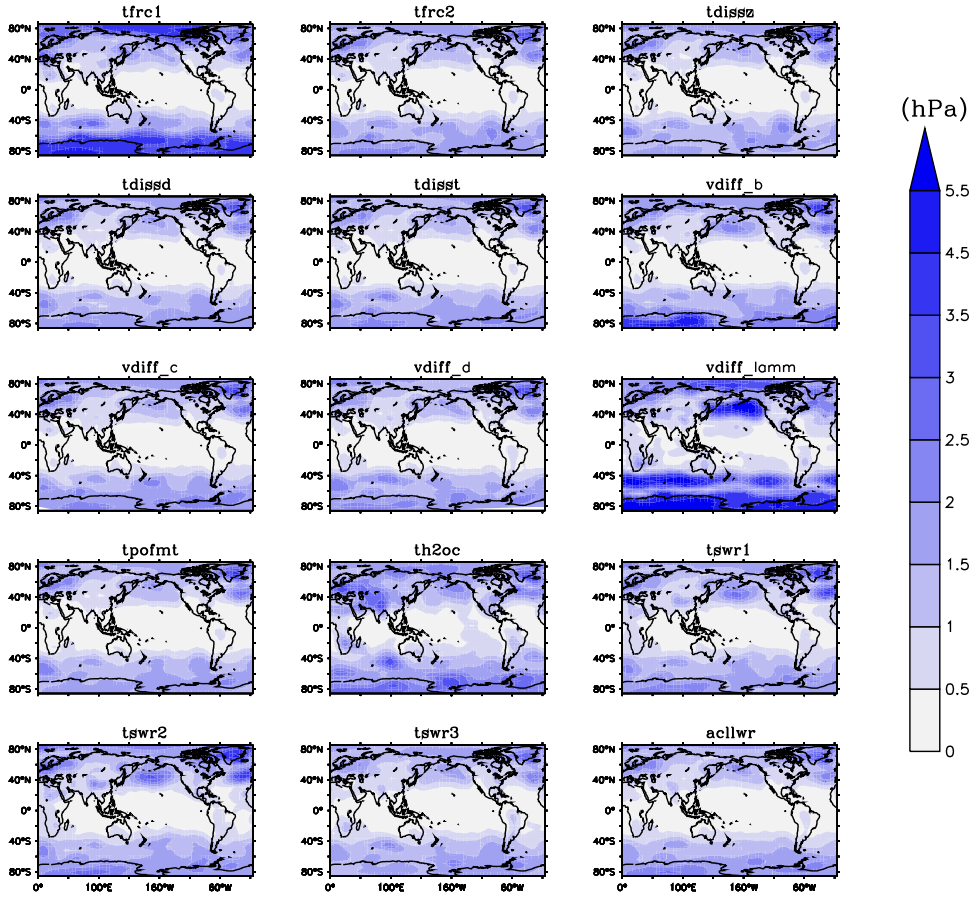

Figure 4.9— same as in figure 4.7 but for SLP (hPa).

Zonal winds at 500 hPa are affected in almost all the regions (Figure 4.10). The exception is the western equatorial Pacific region where some of the parameters have no effect on the upper level winds. Among all parameters, zonal winds at 500 hPa have the maximum sensitivity to vdiff_lamm especially in the southern ocean below 35S and central north Pacific ocean (above 35N). One can recollect from previous chapter that these regions had major errors in the model simulations of zonal winds at 500 hPa. Few other parameters like tfrc1, th2oc and tswr2 have somewhat stronger impact as compared to rest of the parameters.

Apart from the above four basic variables, the effect of sensitivity experiments on surface net heat flux and humidity is also examined. For net heat flux (Figure 4.11) the major contribution to the response comes from all cloud and radiation based parameters (th2oc, tswr1,2 and 3) and almost all of the oceanic regions are sensitive.

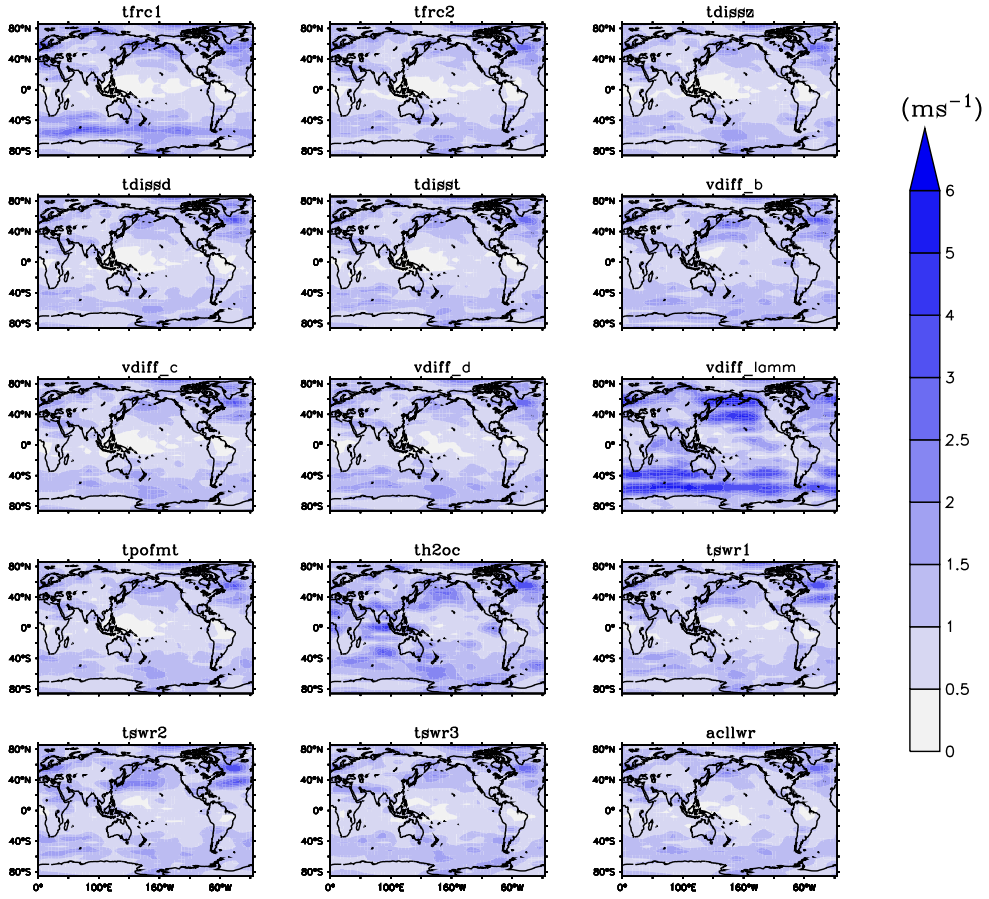

Figure 4.10— same as in Figure 4.7 but for zonal winds (ms^{-1}) at 500 hPa.

However, there are some regions like the North Pacific (above 30N), the Northwest Atlantic and North Indian Ocean, which have somewhat pronounced impact due to perturbations in these parameters. For all other remaining parameters, the errors are within 10-20 Wm^{-2}. Humidity (Figure 4.12)at 1000 hPa is sensitive to th2oc and tswr3 mainly in the tropical oceans and over continents. Remaining parameters influence over the oceans but identically affect the regions over continents.

Next, we look at the sensitivity in the upper levels of the model. For this, the RMSE of the global mean values at each level for temperature, zonal winds and humidity were computed and the results are shown in Figure 4.13. It can be clearly seen that for temperature (left panel, 4.13), at all levels most sensitive parameters are mainly belonging to parametrization of clouds and moisture (th2oc, tswr2, tswr3). While th2oc has maximum influence on the surface, tswr3 contributes to maximum

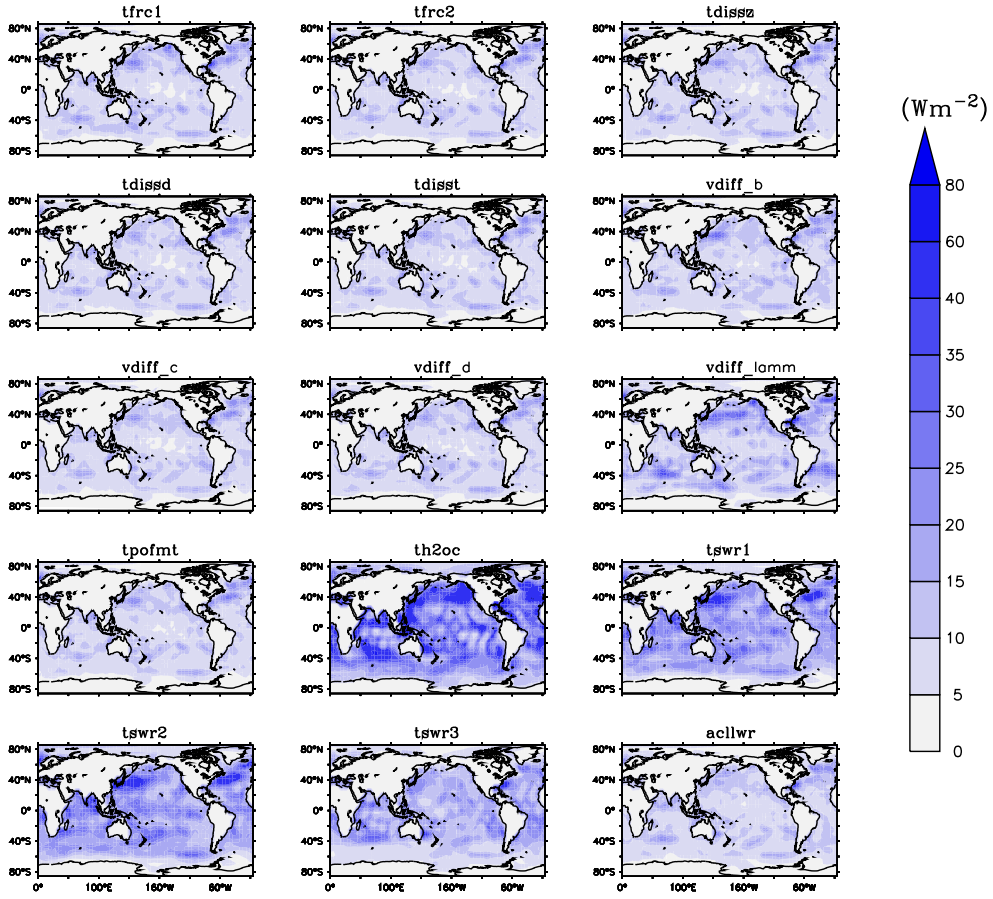

Figure 4.11— same as in Figure 4.7 but for surface net heat flux (Wm^{-2}).

errors at 300 hPa. There is a contribution from vdiff_lamm at 200 hPa and at levels close to surface. From this plot, six main parameters appear to have the maximum impact on temperature and they are th2oc, tswr1,2,3, acllwr and vdiff_lamm. For winds (middle panel, 4.13) the outstanding response is seen above 600 hPa due to changes in th2oc with maximum at 200 and 300 hPa. All other parameters (except for vdiff_lamm) seems to have very little influence on the global mean zonal winds. The peak in response in humidity (right panel, 4.13)is mainly seen at 800 hPa with an exception of tswr3 whose maximum impact is seen at the surface. Along with th2oc, humidity at 800 hPa is sensitive to diffusion related parameters (vdiff_b and vdiff_d).

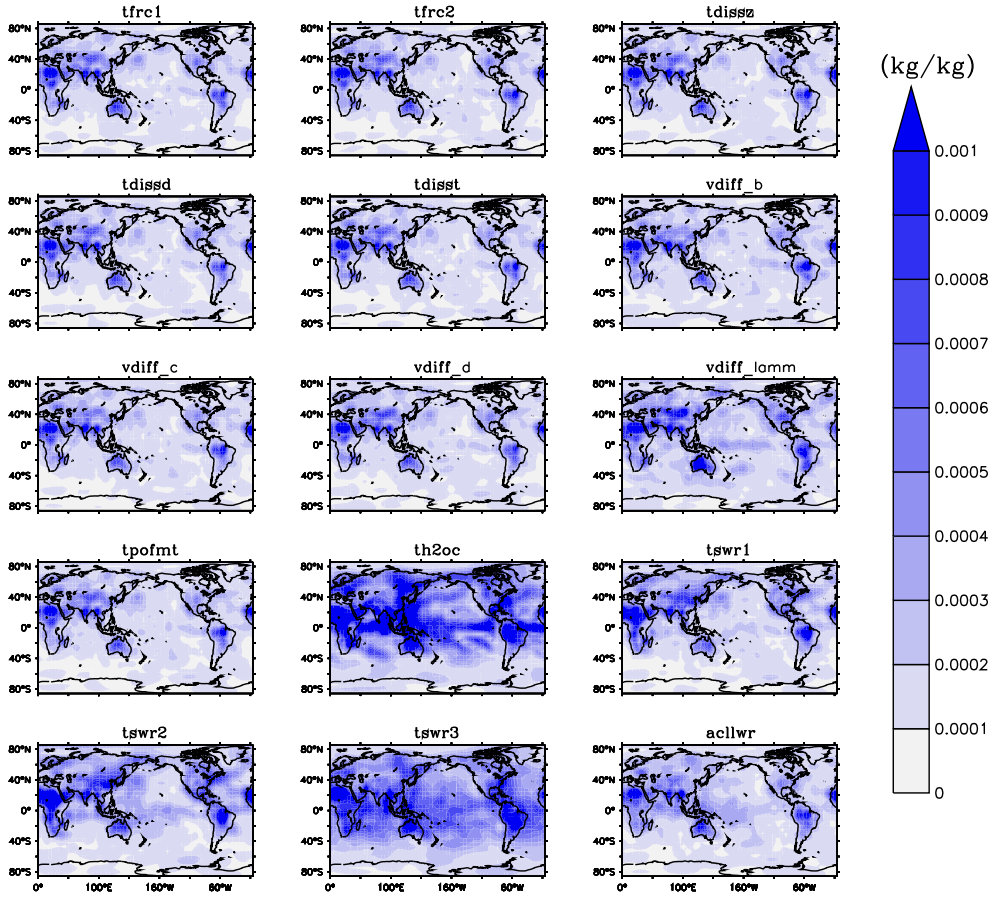

Figure 4.12— same as in Figure 4.7 but for humidity at 1000 hPa (kg/kg).

4.6 Summary

In this chapter, the sensitivity experiments are analyzed in terms of the behavior of the model's cost function with respect to perturbation in the control parameters. It has been found that the cost function is quite sensitive to cloud related parameters and also to the length of time of model integration. Longer integration time results in growing nonlinearities in the cost function and it becomes more difficult to identify true minima. The SA gives an idea about the model response due to perturbation in individual parameters. The regional effect of model parameters on the basic state variables is also examined. SA results will serve as a guideline for choosing suitable parameters including a useful range of applied perturbations for tests in identical twin experiments. Our next aim is to perform parameter optimization in PlaSim

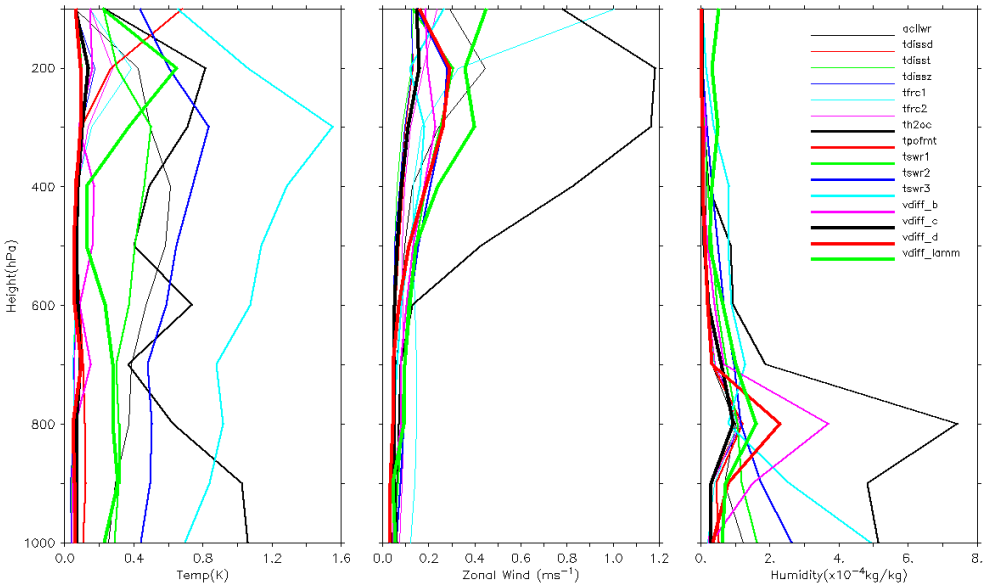

Figure 4.13— Plots of height vs. RMSE between model simulations due to perturbations applied to 15 model parameters and the control run for (left panel) Temperature(K), (middle panel) zonal winds (ms^{-1}) and (right panel) humidity (kg/kg).

using the GF approach which is presented in next chapter.

Chapter 5

Parameter optimization using Green's function

In this chapter, the GF based approach is applied and tested with PlaSim. The technique is briefly discussed in section 3.1. Since this is for the first time the GF is applied to PlaSim, it is of interest to inter-compare two techniques of parameter optimization and this is done in the following sections. Firstly the comparison of two optimization algorithms (Adjoint and GF approach) for the *dry* configuration of PlaSim for shorter time scales based on identical twin experiments is discussed.

5.1 Identical twin experiments

Identical twin experiments were first carried out in planning the Global Atmospheric research program and were also called observing system simulation experiments (OSSE) (Bengtsson et al., 1981). Identical twin experiments provide the initial test of an assimilation system. They assimilate a data set (pseudo data) produced by model itself in a run with the known values of all control parameters that take the role of the optimal values. Identical twin experiments provide a good test for the applicability of the method because the errors are controlled. Therefore methods that perform well in identical twin experiments are also considered for performing data assimilation of real observations.

In the GF approach, the iterative assimilation procedure starts from a perturbed control vector. The experiment is successful if we can approximately recover the default values of control parameters through assimilation of data with a reduction in cost function. In identical twin experiments the optimal values of control parameters

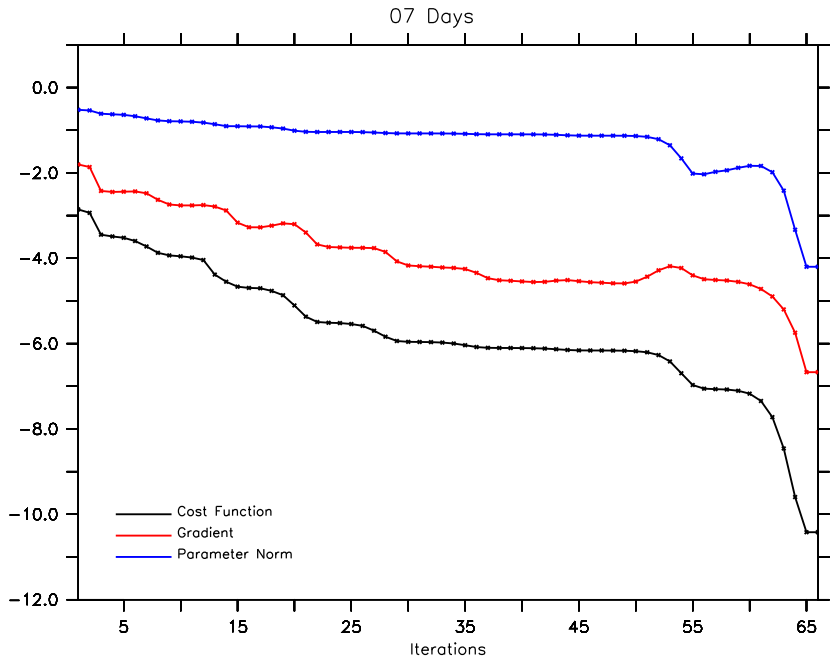

Figure 5.1— Convergence of the Adjoint method using the *dry* configuration for, 9 control parameters, the assimilation of pseudo observations and an integration time 7 days. Cost function (black), norm of its gradient (red) and norm of control vector to default value (blue) over the iteration number. Y axis is on log_{10} scale.

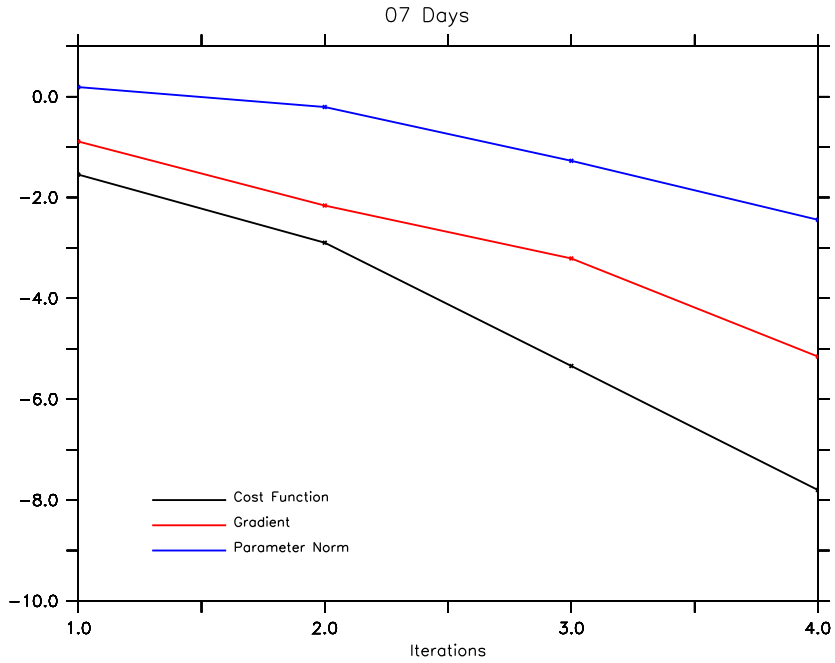

Figure 5.2— Convergence of the GF approach using the *dry* configuration for, 9 control parameters, assimilation of pseudo observations and integration time 7 days. Cost function (black), norm of its gradient (red) and norm of control vector to default value (blue) over the iteration number. Y axis is on log_{10} scale

are known so it is possible to calculate the parameter error norm. The parameter error norm is defined as :

$$D = (\sum_{i=1}^{N}(\theta_i - \theta_0)^2)^{0.5} \tag{5.1}$$

where N is number of control parameters Here θ_0 are the default values of control parameters. θ_i is parameter value at each iteration(i).

The condition that all control parameters converge to their original value, which means the parameter error norm decreases to zero indicates the success of optimization. Since, at the moment optimization results using PlaSim's adjoint in the *dry* configuration are available, in the next section we will compare them with the results using the GF approach implemented with PlaSim.

5.1.1 Inter comparison of Adjoint and GF approach in PlaSim

For comparison, the optimization runs for identical twin experiments were available from the adjoint of PlaSim in its *dry* configuration (Dr. Ion Matei, personal communication). These results were available for integration times of 7, 21 and 56 days. The pseudo observations for identical twin experiments come from the control run (discussed in section 2.2) using default values of control parameters (as listed in Table 4.1) and the optimization was performed using the following 9 control parameters:

- rayleigh friction(tfrc1,2)
- diffusion time scale for vorticity (tdissz)
- diffusion time scale for divergence (tdissd)
- diffusion time scale for temperature (tdisst)
- tuning parameter for vertical diffusivity (vdiff_b, vdiff_c, vdiff_d)
- tuning of point of mean transimittivity in layer (tpofmt)

Using same setup as in the adjoint method, the optimization is performed using the GF approach for the integration time periods of 7, 21 and 56 days. The GF matrix given by equation (3.4) is computed by applying perturbation to the control parameters. The perturbation to the parameters applied is $\sim 10^{-7}$. This choice of perturbation was based on several trial experiments where the smoothness of the

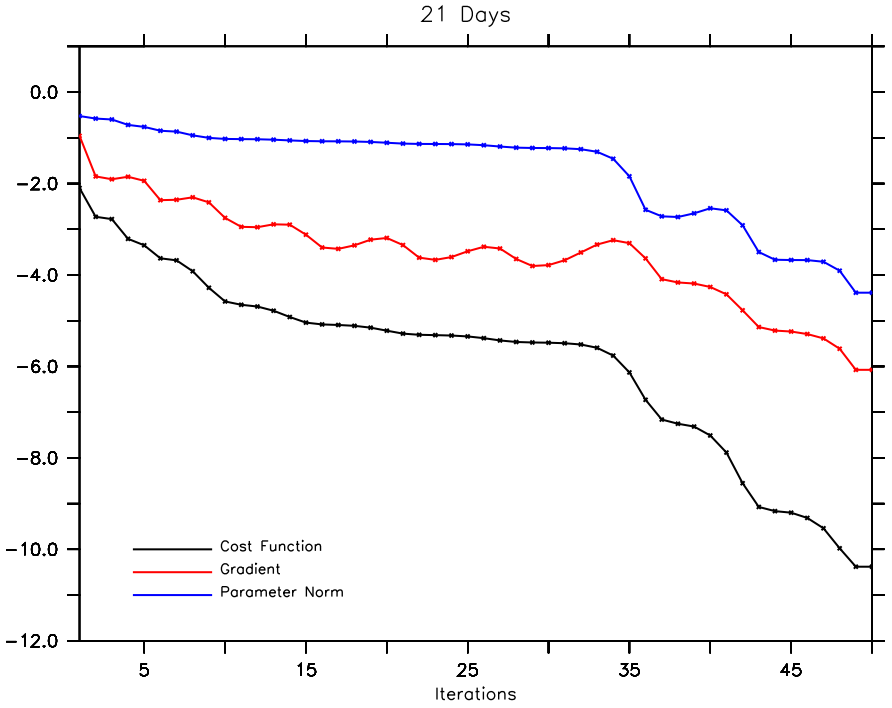

Figure 5.3— Same as Figure 5.1 except for an integration time of 21 days

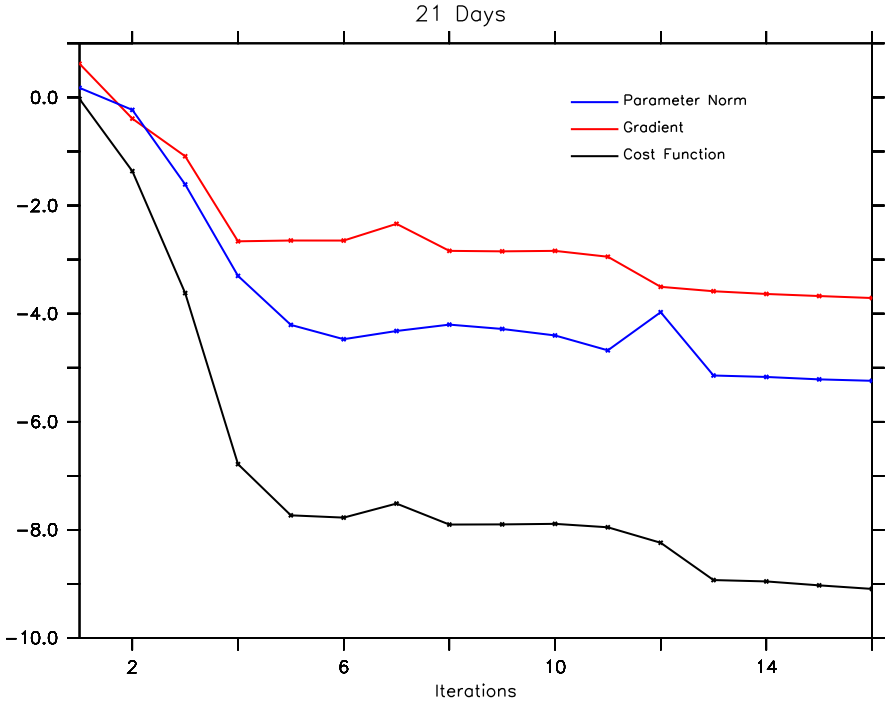

Figure 5.4— Same as Figure 5.2 except for an integration time of 21 days

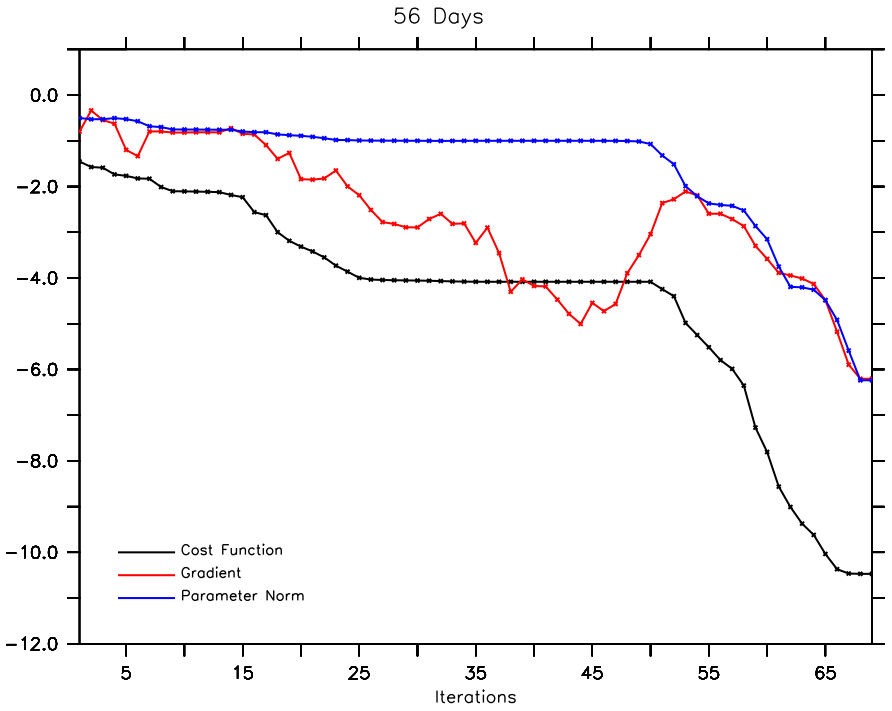

Figure 5.5— Same as Figure 5.1 except for an integration time of 56 days

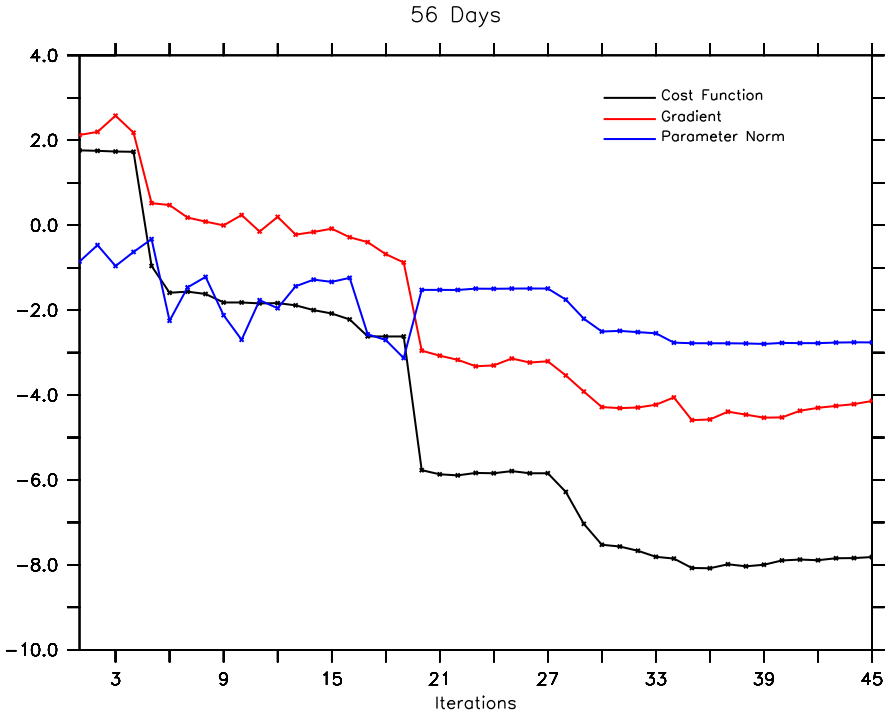

Figure 5.6— Same as Figure 5.2 except for an integration time of 56 days

cost function played a very important role. The estimated parameters are obtained as according to equation (3.5) and these values of new parameters are again used to compute the GF and a new estimate of the parameters is again made. The iterative procedure is stopped if the gradient of the cost function show little changes between successive iterations as compared to the start of the iterative procedure when the gradients were quite large. At this point, it is assumed that no further reduction in cost function can be obtained. Figures 5.1 and 5.2 show the results of the adjoint Method and the GF, respectively, for an integration time of 7 days. The plots show a reduction in the cost function, the parameter error norm and the gradients with respect to the number of iterations. It can be seen from the figures 5.1 and 5.2 that the number of iterations required to reach the minimum cost function is large in case of the adjoint method as compared to the GF approach. For the adjoint method, a minimum cost function is achieved in around 65 iterations whereas with the GF approach only 4 iterations achieve convergence. The parameter error norm (blue line) and the gradients (red line) also show a consistent decrease.

Similar experiments are performed for an integration time of 21 and 56 days. Figures 5.3 and 5.4 show the results of adjoint and the GF approach, respectively for integration time of 21 days. It is to be noted that while for GF the number of iterations that it took to achieve convergence increased in 21 days experiment as compared to 7 days experiment, while it is not for the adjoint method. In the 21 days experiment using the adjoint method, the convergence is achieved in 50 iterations (as compared to 65 iterations in 7 days experiment) while it took only 14 iterations in case of the GF. Similarly, it can be seen that for 56 days the minimum cost function is attained using the adjoint method in 70 iterations (Figure 5.5) and using the GF in only 40 iterations (Figure 5.6).

It is clear from the above mentioned results that with the GF approach, minimum cost function can be achieved in less number of iterations as compared to adjoint method. From these three different experiments it can be inferred that with the GF method the number of iterations required to minimize the cost function and achieve convergence increases when the assimilation window length increases, however it is not so in the adjoint method.

A complete adjoint-method optimization may require approximately 4 times as much time to complete as a forward-model integration. The adjoint method requires a tangent-linear model and in addition, its computational expensiveness, implementation is technically demanding and it is unable to handle chaotic systems. The

chaotic dynamics of the AGCM restricts the applicability of Adjoint method to very short time scales.

In the next section, we demonstrate the application of the GF on the *wet* configuration of PlaSim for both shorter (30 days) and longer time scales (1 year)

5.1.2 Application of Green's function approach in the *wet* configuration of PlaSim

In this section, the *wet* configuration of PlaSim is used. Since the *wet* configuration includes the hydrological cycle in PlaSim therefore cloud related parameters (as shown in the Table 4.1) can also be included for the optimization. The selection of control parameters is based on the sensitivity of model's cost function. The sensitivity experiments were discussed in chapter 4. Only the parameters which generated large deviations in the cost function due to change in their values were selected for further experiments using GF method. The following parameters were selected:

- tuning parameter for vertical diffusivity (vdiff_lamm)
- absorption coefficient h2o continum(lwr) (th2oc)
- tuning of cloud albedo range1 (tswr1)
- tuning of cloud back scattering range2 (tswr2)
- tuning of cloud scattering albedo range2 (tswr3)
- mass absorption coefficient for clouds (lwr) (acllwr)

The results from the sensitivity experiments (section 4.4.2) with the *wet* configuration reveal that the model cost function is not smooth and contains several local minima. Therefore for using GF in the wet configuration of PlaSim the value of perturbation applied to estimate gradients will be different than what was used in GF experiments with *dry* configuration. Since for the *wet* configuration cost function is not smooth, smaller perturbation for estimating gradients will result in wrong search direction. Therefore, the perturbation applied to estimate gradients should be large enough to overcome the noise level of the cost function. The sensitivity experiments performed in chapter 4 gives an idea for the necessary range of perturbations applied to estimate gradients.

Experiments	Control parameters						Integration time (in days)	Cost function reduction(%)
	vdiff_lamm	th2oc	tswr1	tswr2	tswr3	acllwr		
EXP-6P-30	160	0.04	0.04	0.048	0.004	0.1	30	97
EXP-5P-30	-	0.04	0.04	0.048	0.004	0.1	30	92
EXP-3P-30	160	0.04	-	-	0.004	-	30	96
EXP-6P-365	160	0.04	0.04	0.048	0.004	0.1	365	95
EXP-5P-365	-	0.04	0.04	0.048	0.004	0.1	365	94
EXP-3P-365	160	0.04	-	-	0.004	-	365	97

Table 5.1— Three experiments performed with different sets of control parameters used in the optimization procedure for different integration time scales in identical twin experiments.

Different set of experiments are performed using pseudo data sets for an integration time of 30 days and 365 days. In the first experiment (called EXP-6P-30), all 6 selected parameters were used as control, while in EXP-5P-30, 5 control parameters that are related to cloud processes were used. In EXP-3P-30, only the three most sensitive parameters are considered. The control parameters used and their values are listed in Table 5.1. The purpose of choosing different control parameters was to check the dependence of the performance of the GF approach on the choice of parameters. The perturbation applied to each parameter for calculating gradients is 100% of the parameter value. The cost function in EXP-6P-30 reduces from a value of 100 to a minimum value of 3 (Figure 5.7) in around 9 iterations. In EXP-5P-30, the cost function reduces to a minimum value of 7 (from an initial value of 92) in around 7 iterations along with a decrease in parameter error norm and gradients. In EXP-3P-30, the cost function is reduced from a value of 114 to 5 in 4 iterations. Table 5.1 shows the percentage reduction in cost function for each of the experiments. The middle panel of Figure 5.7 shows the parameter error norm as calculated using equation 5.1. The parameter error norm represents the deviation in the parameter values from the control values and it is seen from Fig 5.7 (middle panel) that with successive iterations the parameter error norm reduces and become close to zero at the point of attaining minimum cost function. The parameter error norm getting closer to zero suggests that the optimization algorithm is successfully able to retrieve the original parameter values. Simultaneously, the gradients are also reduced until about 10 iterations. In the begining, gradients shows maximum decrease, however beyond 10 iterations there is no clear trend and the fluctuations in the gradients become quite large and no further convergence can be achieved (Figure 5.7, lower panel). In this case of identical twin experiments, the true minimum should ideally

be zero. However, due to the stochastic behavior of the cost function the minimum of the perfect minima cannot be found. This minimum can only be reached using the precise set of parameter values which were used in the control run of the experiments. Any small deviation in any or all of the parameter values result in a sizable increase in the value of the cost function. From sensitivity experiments we saw that unless we use exact parameter values, the minimum value of the cost function that could be attained is between 10 and 20 for a 365 day integration period. Further reduction in the value of cost function is not attainable. In fact continuing further with iterations result in large deviations in parameter error norm and the cost function begins to rise again.

Beyond 10 iterations there is an increase in the value of the cost function, the gradients show random fluctuations. It is also seen that for all three experiments, the parameter error norm increases after attaining a minimum. This is due to the large fluctuations in the gradients that the deviations in the parameters again begin to increase and therefore the cost function also show a steady rise. From the above experiments it can be seen that GF is able to obtain a unique minimum cost function and parameter error norm for wet configuration of PlaSim in identical twin experiments performed at shorter time scales. Blessing et al. (2014) noted that the Adjoint approach was not successful for even less days of PlaSim simulations in *wet* configuration.

We now attempt to demonstrate parameter optimization using GF for longer model integration periods. For this one year model integration time is chosen. The model integration period of one year will be just enough to demonstrate the robustness of the optimization capability of GF for climate applications. The above experiments were repeated for cost function calculations for an integration time of one year. The convergence of the cost function and the parameter error norm is quite similar to the 30 day experiments except for the different values of initial cost function which in this case is slightly higher because of longer integration times (Figure 5.8). The minimum cost function as shown in Figure 5.8, top panel is achieved in 18, 10 and 6 iterations for EXP-6P-365, EXP-5P-365 and EXP-3P-365 respectively. The minimum value of cost function is within the expected minimum range (as shown in Figure 4.6). At the point of attaining minimum cost function the parameter error norm (Figure 5.8 middle panel) also becomes minimum and is close to zero. The minimization of norm indicates that the optimization procedure has reached convergence and the parameter values retrieved are close to the actual values. Iterations carried beyond this point leads to an increase in the norm indicat-

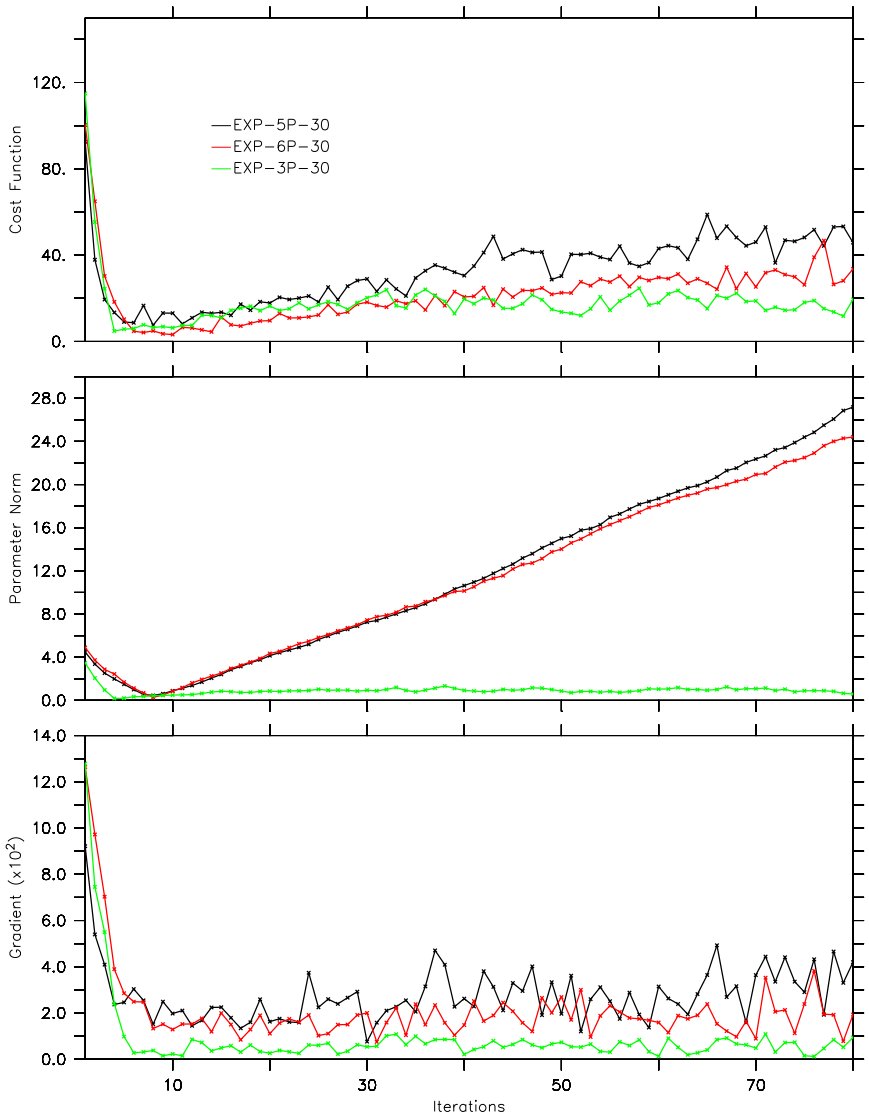

Figure 5.7— Convergence of the GF approach in the *wet* configuration using pseudo observations for EXP-6P-30, EXP-5P-30 and EXP-3P-30. Cost function (top panel), parameter error norm (middle panel) and gradients (bottom panel) versus iteration numbers. The integration time is 30 days.

ing that the parameter values are starting to diverge away from the actual values. The larger fluctuations in the gradient (Figure 5.8 bottom panel) after the point

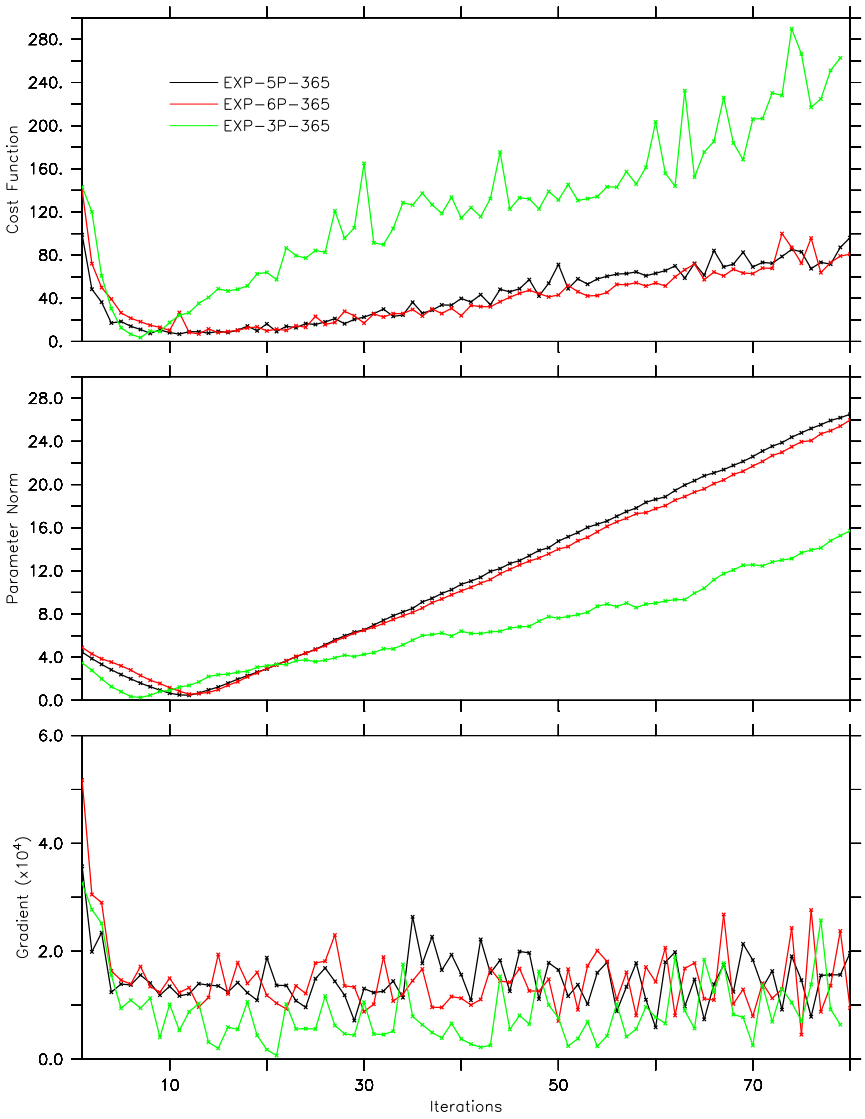

Figure 5.8— Convergence of the GF approach in the *wet* configuration using pseudo observations for EXP-6P-365, EXP-5P-365 and EXP-3P-365. Cost function (top panel), parameter error norm (middle panel) and gradients (bottom panel) versus iteration numbers. The integration time is 365 days.

of minima suggest that continuation of iterations beyond this point leads to large uncertainties. From the above results, it appears that this optimization procedure

for identical twin experiments is quite successful for longer integration periods of the atmosphere model.

5.2 Optimization with ERA-Interim reanalysis data

After successfully assimilating pseudo data using the GF approach, the next step is to perform the experiments with the real observations. ERA-Interim reanalysis data as described in section 2.3 has most of the available observational data assimilated and therefore it can be considered as one of the most accurate representation of the climate system. As a next step towards optimizing the atmosphere model we make use of this data set. For computing the cost function, instead of pseudo data, annual means of state variables from ERA-interim data set are used. The period of data from which the annual means are constructed spans year 1979 to 2012. The cost function computation comprise of same state variables which were used in the identical twin experiments as described in section 5.1.

Optimization experiments are performed similar to those in section 5.1.2 using 3, 5 and 6 control parameters described in table 5.1 for shorter and longer (30 and 365 days respectively) integration time periods and with different combination of the above used parameters. The default values of the control parameters as mentioned in table 5.1 is used as starting values in the iterative procedure. The perturbation applied in successive iterations is large and comparable to the parameter value itself. The cost function using the control values of model parameters with respect to ERA-interim observations is $\sim 10^5$ which indicates that model is far away from the observations. There is no significant reduction in the cost function with successive iterative steps of GF approach and fluctuations in the gradients are observed. Subsequent efforts for optimizing the model were carried out by reducing the number of control parameters, however the behavior was similar and the cost function curve remained noisy for entire iterations and no meaningful optimization could be attained. Finally, the most sensitive parameters vdiff_lamm and th2oc are used for the procedure. Two experiments are performed using ERA-Interim data with the GF approach. The Figures 5.9 and 5.10 shows the results of the iterations for 30 and 365 days and with the above mentioned control parameters. The cost function is of the order of 10^5. It is observed from these two figures that there is no reduction in the cost function. Also, gradients are not reducing with respect to iterations and they are not helpful (Figure not shown) in finding minimum cost function. In experiments with real data optimal values of control parameters are not known and

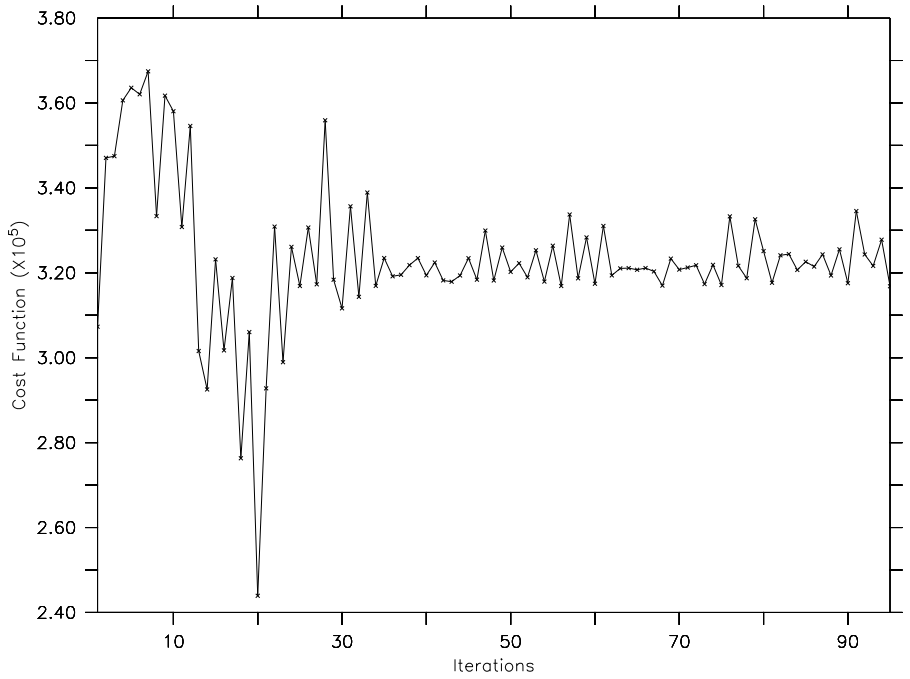

Figure 5.9— Reduction in cost function wrt Iterations when optimization is done with ERA-Interim data using 2 control parameters and integration time is 30days.

therefore it is not possible to get a measure of the parameter error norm.

5.3 Summary

In this chapter, the implementation and application of GF on PlaSim is discussed. It is found that GF and Adjoint method can successfully optimize parameters in the *dry* configuration of PlaSim and their results are comparable. The advantage of using GF is the ease implementation of code as compared to the adjoint method. The performance of the GF with the *wet* configuration and using pseudo data is quite satisfactory. GF approach used on *wet* configuration of PlaSim is able to obtain optimal model parameters within 10 iterations for 30 and 365 days integration time in identical twin experiments framework. However, GF fails to optimize PlaSim when real observations are used to compute the cost function.

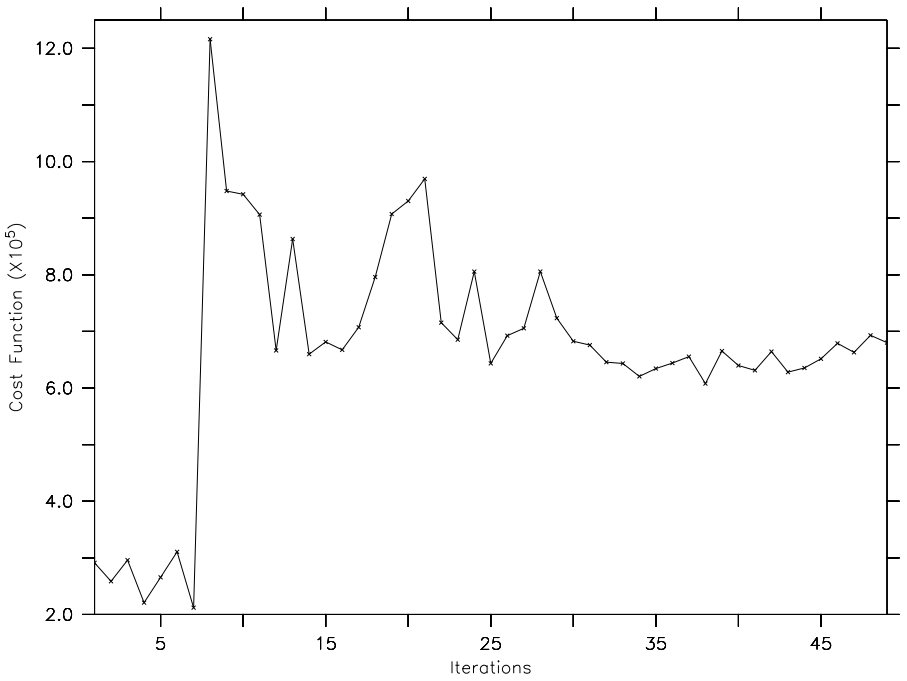

Figure 5.10— Same as in Figure 5.9 but integration time is 365 days.

One of the drawbacks of the GF approach is that the computational cost increases linearly with the number of control parameters that are used in the procedure. Another major limitation that is found in GF approach is that it is unable to optimize parameters with ERA-interim reanalysis data .One of the reasons for this could be large biases in the PlaSim with respect to ERA-interim observations which grow with longer integration time. In the next chapter, stochastic optimization method SPSA is implemented and discussed.

Chapter 6

Parameter optimization using SPSA

In the last chapter, it was seen that the GF approach is successfully able to retrieve optimal parameters for both *dry* and *wet* configurations of PlaSim in identical twin experiment framework. It was also found that Green's function performance is computationally better than the adjoint method in the *dry* configuration of PlaSim. The major drawback of GF approach from the experiments performed is that it fails when applied to real atmospheric conditions, i.e. when the data used for cost function computation are not pseudo observations but real observations. Another major drawback of GF approach is that the computational cost increases linearly with the number of control parameters. To overcome above drawbacks of GF a stochastic approach, the Simultaneous Perturbation stochastic Approximation (SPSA) is considered. The concept of SPSA is developed by (Kiefer and Wolfowitz, 1952) and these stochastic algorithms perform random search in parameter space. SPSA is based on highly efficient and easily implemented simultaneous perturbation approximation to the gradient. This gradient approximation for the central difference method uses only two cost function evaluations independent of the number of parameters being optimized. As a result of the stochastic perturbation the calculated gradient is also stochastic, however the expectation of the stochastic gradient is true gradient (Li and Reynolds, 2011).

The advantage of stochastic algorithms is its ability to deal with random noise in the measurement of cost function. Also these algorithms rely on an approximation to the gradient from the noisy measurements of cost function. These algorithms can be easily implemented with any numerical model hence no huge programming

effort is needed. The SPSA algorithm has gathered a great deal of interest over last decade and has been used in variety of applications (Hutchison and Hill, 2001; Spall, 2000; Gerencsér and Vágó, 2001; Kong et al., 2011). The gradient based algorithms (Adjoint,EnKF) are faster to converge than any gradient approximated algorithms (GF, SPSA algorithm) when speed is measured in terms of the number of iterations. The total cost to achieve effective convergence depends not only on the number of iterations required but also on the cost needed to perform these iterations which is more in gradient based algorithms. This cost may include computational cost and additional human effort required for coding the gradients. In the following sections, implementation and experiments of SPSA are discussed.

6.1 Identical twin results

SPSA scheme based upon formulation given in section 3.3 is implemented in PlaSim. Following this, a set of identical twin experiments are performed. Different sets of experiments with different control parameters (as in chapter 5) with pseudo data and with different integration time scales are carried out. The cost function (model data difference) comprises of 16 variables which are listed in Table 4.2 together with their corresponding globally averaged STD (C_0). A default setup (as explained in section 2.2) is used to generate pseudo data for one year period. The experiment is successful if we can approximately recover the default values through assimilation of pseudo data via reduction in the cost function. All experiments using SPSA are performed on the *wet* configuration of PlaSim.

First experiment is carried out using the same 5 control parameters as used in EXP-2 of chapter 5 for integration time 30 days. The perturbation applied to each parameters is 200% of their default value. The value of a and c in the gain sequence given by equation 3.12 were chosen to be 0.05 and 1.0 respectively. According to Spall (1998) the value of a should be very small as compared to c. It is effective to set c as standard deviation of the noise in the cost function. The values of a and c that are used in this study are a result of several trial experiments where the criteria of cost function convergence was used. Figure 6.1 shows the result of optimization with SPSA. Top panel shows the reduction in the cost function with respect to each iteration. The cost function is reduced from a value of 35 to $\sim$5 in 50 iterations. Simultaneously the parameter error norm defined in equation 5.1 has reduced approximately to zero. Because of the nonlinear behavior of cost function

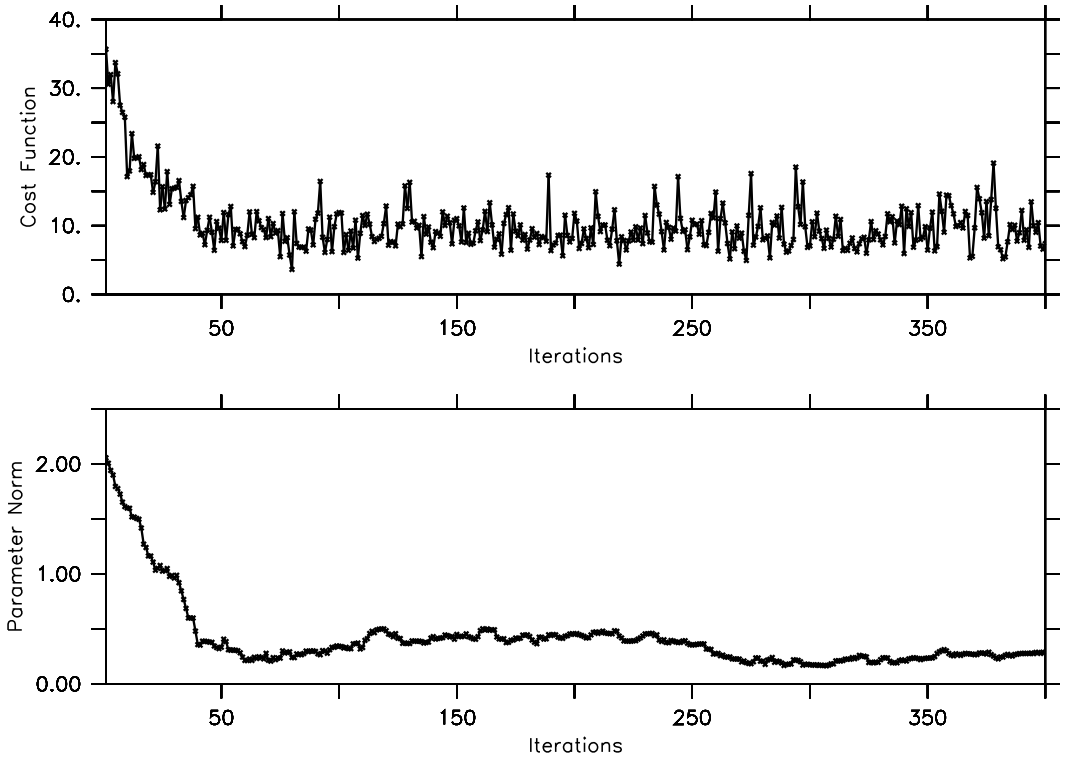

Figure 6.1— Cost function vs. iteration number for selected 5 parameters (Table 4.1) in the identical twin experiment. Cost function is computed for a period of 30 days model integration

exact minimum cannot be achieved but a nearby minimum is achieved and this minimum is in the same range as explained in sensitivity experiments. The number of iterations required to achieve convergence are larger than those in GF approach. However, it must be noted that in GF for 5 parameter case, each iteration required 5 model runs (one corresponding to each parameter perturbation) while in SPSA only two model runs per iteration are needed for any number of control parameters.

Next experiment is performed for integration time of 365 days and with same setup as for 30 days. Figure 6.2 shows the corresponding result. Top panel shows the reduction in cost function with respect to iterations and it is reduced to minimum range of 10-20. This is the same range of minimum as shown in sensitivity experiments Figure 4.6. The minimum cost function is found in 150 iterations and parameter error norm is also reduced in 150 iterations.

In the GF approach we reduced the number of control parameters that were

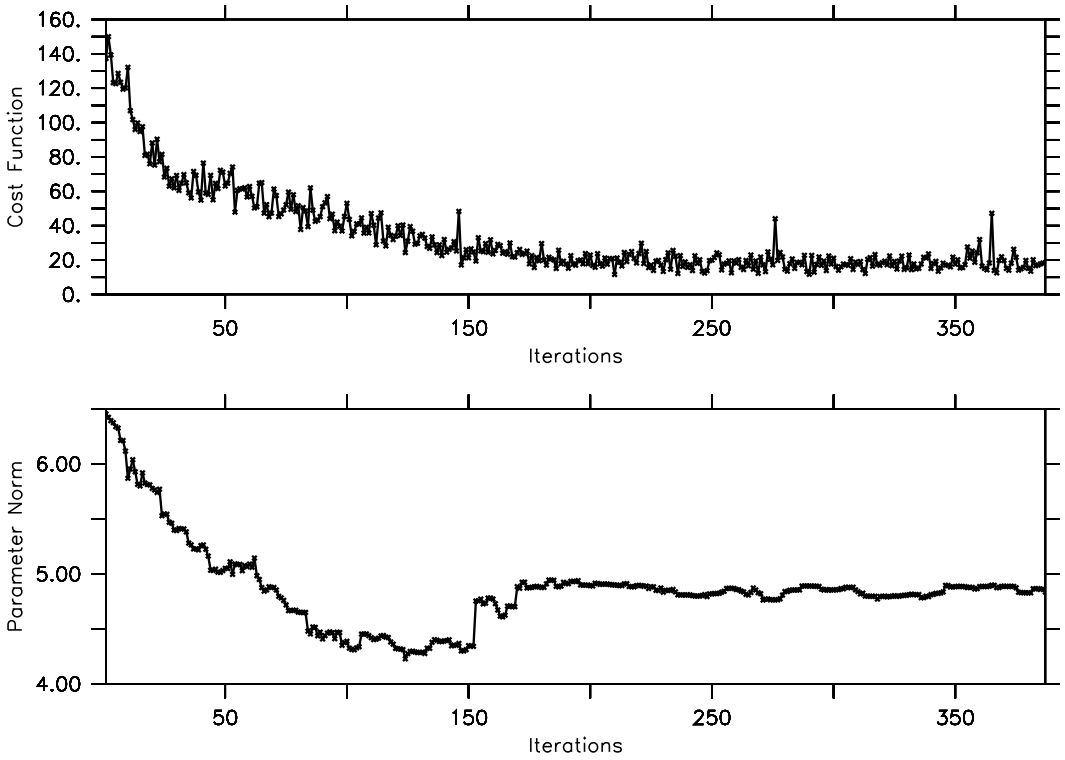

Figure 6.2— Cost function vs. iteration number for selected 5 parameters (Table 4.1) in the identical twin experiment. Cost function is computed for a period of one year model integration

actually used in the procedure. The main reason for this was the computational time associated with each parameter. The total time taken for one iteration in GF is dependent on the number of control parameters used. Hence we restricted the control parameters to 6 based on the sensitivity experiments of PlaSim However, SPSA is not limited by this constraint as the perturbation applied to all the parameters is simultaneous and iteration time is independent of the number of control parameters. Therefore an attempt is made to apply SPSA on PlaSim using all the 15 control parameters that are mentioned in Table 4.1. The iterative procedure is started using a perturbed vector of selected 15 parameters as mentioned in Table 4.1, known as the control vector. The size of perturbation is based upon the sensitivity of cost function to each parameter. As seen in Figure 4.5, the parameters tfrc1, tfrc2, tdissz, tdissd, tdisst, vdiff_b, vdiff_c, vdiff_d and tpofmt are contributing almost same as the noise in the cost function and therefore the size perturbation applied to these is larger ($\sim$ 200% of the original value) as compared to the remaining parameters ($\sim$ 20% of the original value). The assimilation is carried out over one year integration time

scale. The experiment is successful if we can approximately recover the default values through the assimilation of the data via reduction of the cost function.

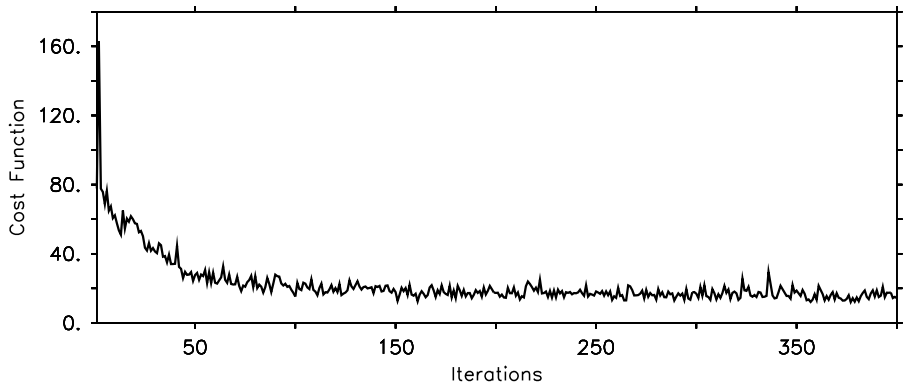

Figure 6.3— Cost function vs. iteration number for selected 15 parameters in the identical twin experiment. Cost function is computed for a period of 1 year model integration

The control run as explained in section 2.2 is used as reference to compute the cost function using equation 4.2 of section 4.3.1. The value of a and c in the gain sequence used are 0.01 and 0.2 respectively. a is used for gain sequence a_k which relates to the inverse of Hessian in the Quasi-Newton algorithm. The value of c is decided by the variability of the cost function over certain range of perturbations in parameters (Spall, 1998). All 15 parameters (normalized to 1) are perturbed simultaneously and using the SPSA method described in section 3.3, gradients are computed and corrections to the parameter are applied.

Figure 6.3 shows the cost function vs. iterations plot. It can be seen that the cost function achieves the minimum value of $\sim$15 in around 120 iterations and then continues to fluctuate about this value further on. As discussed in section 4.5 this value of cost function lies in the range of an acceptable minimum and it cannot reduce further. At this stage about 2/3 of the cost corresponds to temperature and 1/3 to the surface fluxes.

Ideally, in identical twin experiments, the reduction in cost function and in Euclidean distance of control vector to the default values is achieved simultaneously. The absolute values of the differences of control parameters are plotted with successive iterations is shown in Figure 6.4. Here, the absolute difference ($\Delta\theta$) is defined as the

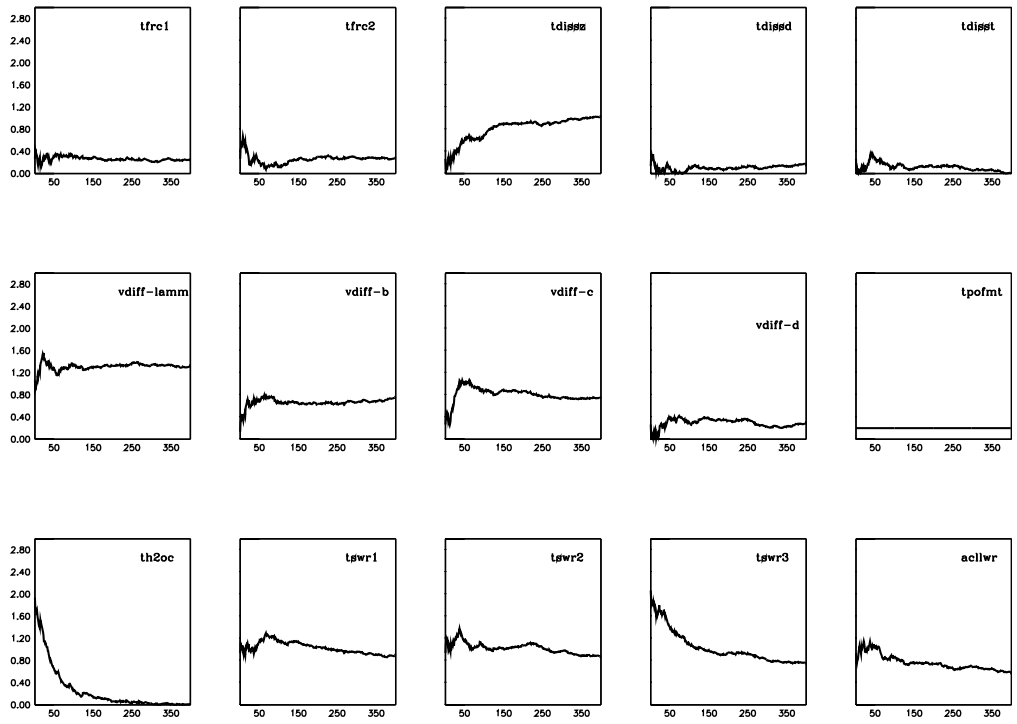

Figure 6.4— Absolute differences ($\Delta\theta$) vs Iteration plotted individually for all 15 parameters in identical twin experiments. Integration time is 1 year. The absolute differences is computed as shown in equation (6.1)

distance between the default parameter ($\hat{\theta}_i$) value and value of parameter at each iteration (θ_{ki})

$$\Delta\theta_i = |\hat{\theta}_i - \theta_{ki}| \tag{6.1}$$

where $\hat{\theta}_i$ represents control parameter values.

In identical twin experiments, the parameters must reach to their default value, however due to the effectively stochastic behavior of the model this may not happen in case of multiple parameters. The parameters are expected to converge to their original values, however in our case only values for 8 out of 15 parameters were retrieved, while remaining 7 parameters were not converging to their default values. As mentioned above, for 8 out of 15 parameters the absolute differences $\Delta\theta_i$ has decreased meaning that the parameter is getting closer to its true value. From the sensitivity experiments curve in Figure 4.5 it can be seen that the cost function is less sensitive to the parameters which were not correctly retrieved and that some parameters were not contributing in the cost function more than the noise level. So

it is difficult to retrieve those parameters. However, the cost function minimum in our case was achieved which indicate that the technique has a potential application to our model. Annan et al. (2005) got similar results where they also found that not all parameters converge to their default values. Their argument was that the data used to constraint the parameters were inadequate and so this type of result was expected.

In the next section ERA-interim observations are used in the optimization procedure.

6.2 Application using reanalysis data

The observational data used are from the ERA-Interim reanalysis. These datasets are interpolated on model levels and then averaged to the model horizontal grid. We have used annual means of the temperature (all levels), large scale precipitation (LSP), convective precipitation (CP), surface sensible heat flux (SSHF), surface latent heat flux (SLHF), surface solar radiation (SSR) and surface thermal radiation (STR).

As expected, with ERA-Interim data the cost function has a larger value (order 3 larger) as compared to cost function against pseudo data and therefore the value of a in the algorithm used to be low ($\sim 1e^{-4}$) for stable results in successive iterations. The model was integrated for one year using default values of parameters, the gradients with respect to the ERA-Interim data were used to calculate corrections to parameters iteratively. At each iteration the cost function reduces until it reaches a value after which there is no substantial decrease and the cost function fluctuates around that minimum value. This happens after about 500 iterations and by then the reduction in cost function is around 25%. The main contribution for reduction in cost function is from tfrc1 (14%), tpofmt (41%), th2oc(11%), tswr1 (16%) and tswr3 (14%) , while the contribution of the remaining parameters is significantly smaller.

The robustness of the experiment results is assessed by an ensemble with three members, each of which start from a different point in control space. Three ensembles (Ens-1, Ens-2 and Ens-3) using different initial values of control parameters are carried out with 15 control parameters. In Ens-1, Ens-2 and Ens-3 the value of each control parameter is increased by 0% (default parameter values), 60% and 100% respectively. Figure 6.5 shows the reduction in cost function with successive model iterations for three ensembles.

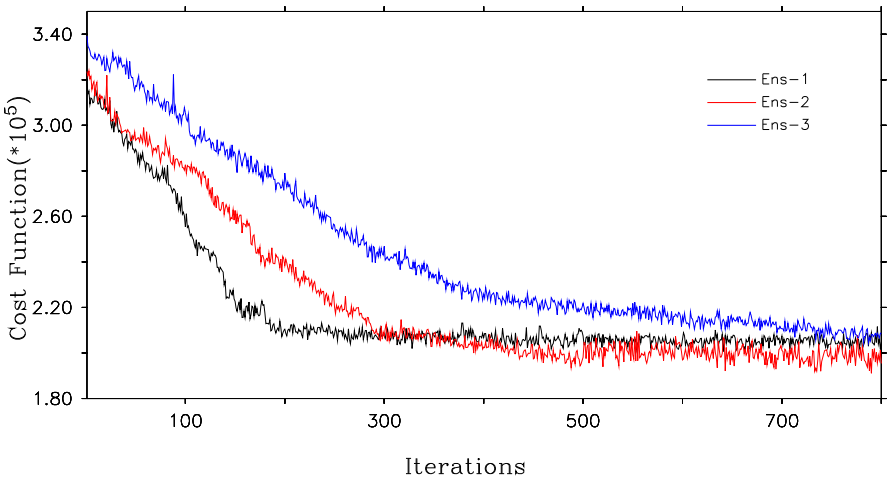

Figure 6.5— Total cost function(using ERA-Interim reanalysis data) vs iteration number. Reduction in cost function with respect to iterations for three ensemble members.

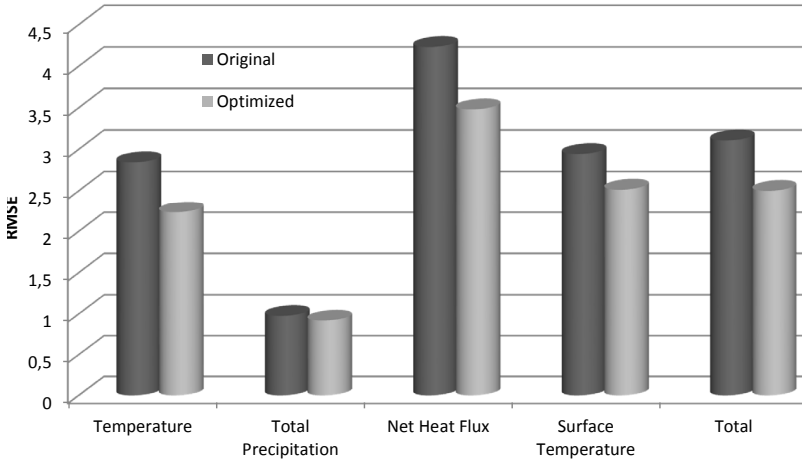

Figure 6.6— RMSE of all (total) and individual(temperature, precipitation, net heat flux and total cost) contributions in original and optimized model states. The reduction is $\sim$ 20%, 5%, 11%, 18% and 21% for temperature, precipitation, net heat flux, surface temperature and total cost function.

The cost function in all the cases attains *near* identical values around 700 iterations. All three cases converging to same value of the cost function gives the impression that the technique is robust and identifies unique minimum instead of multiple minima. The performance of optimization was judged by means of Root Mean Square Error (RMSE) which is the square root of normalized cost function given by

$$RMSE = \sqrt{\frac{Y(\hat{\theta})}{N}} \tag{6.2}$$

where N represents the total number of observations. The model state resulting after using optimized parameters shows improvement on the original state when compared to ERA-Interim reanalysis.

The total RMSE which was $\sim$ 3.1 in the Original run is reduced by $\sim$16% to 2.6 while for the individual variables the reduction is 5%, 11% and 18% for precipitation, net heat flux and surface temperature respectively (Figure 6.6). The optimized control parameters are used for a forward run of PlaSim for 1 year and an optimized state is made. The RMSE of the global mean temperature profile from the original and an optimized state is shown in Figure 6.7. The errors have reduced at almost all levels barring level 2. The levels of maximum errors experienced significant

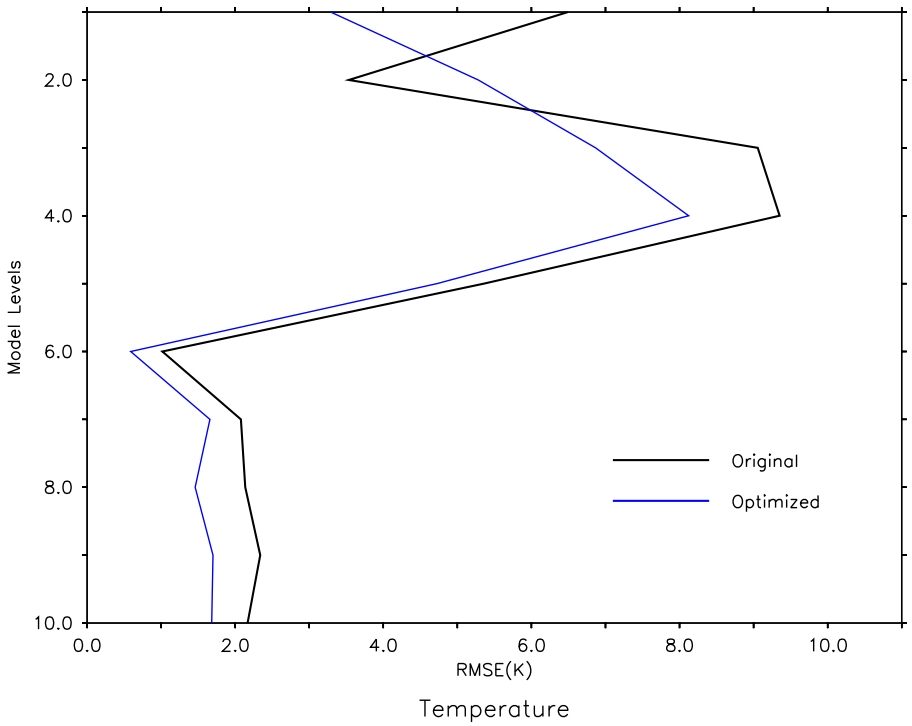

Figure 6.7— RMSE of global mean Temperature between ERA-Interim and Original and optimized model temperature. Y axis denotes height of the atmosphere represented as sigma levels with model level 10 corresponding to the bottom and model level 1 corresponding to the top of the atmosphere

reduction. Overall there is ∼10-20% reduction of error in temperature at various levels in the atmosphere. Spatial distribution of errors of surface temperature and surface net heat flux from original and optimized state is shown in Figures (6.8 and 6.9). Most of the error reduction in surface temperature takes place in the equatorial West Pacific, North Atlantic and Southern Indian Ocean regions. Net heat flux shows improvement almost in every region with large errors of $> 50\ Watts/m^2$ coming down within 10 $Watts/m^2$ in the North Pacific, south-east Africa and also in South Atlantic ocean. There are some places where optimization has resulted in larger errors, most prominent of these are the regions around Greenland in the North Atlantic and in the eastern equatorial Indian ocean. Not much improvement is seen in total precipitation, however there are some regions in the equatorial West Pacific and in the Indian Ocean region where the RMSE in total precipitation has decreased (Figure not shown).

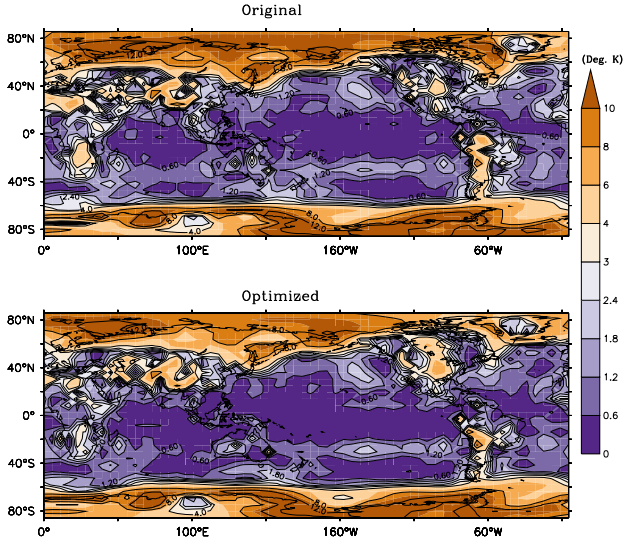

Figure 6.8— Absolute mean difference of near surface temperature between ERA-Interim and (top panel) Original, (bottom panel) Optimized

In chapter 2, annual mean differences of PlaSim with respect to the ERA-interim reanalysis for near surface temperature, precipitation, sea level pressure and 500hPa winds (2.1 - 2.4) were shown. Figures 6.10 - 6.13 show the differences in the optimized state resulting from SPSA procedure for the same variables as discussed in section 2.3. For temperature at 1000 hPa, the differences have reduced mainly in the northern Polar Regions (Arctic and Greenland) where the temperatures were underestimated. Marginal improvements are also seen over equatorial oceans. How-

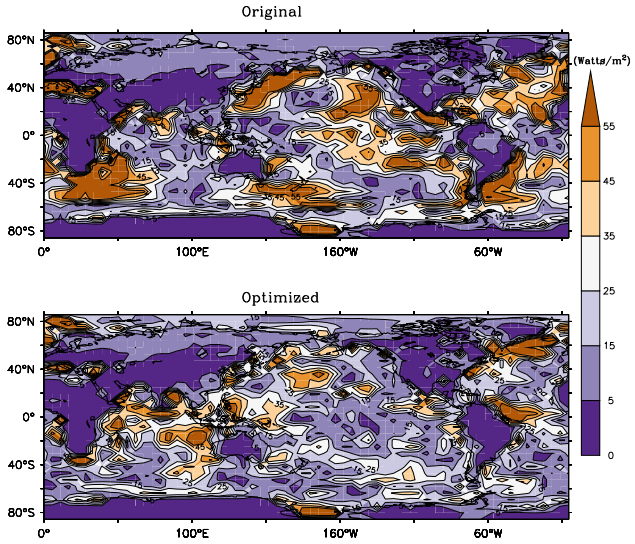

Figure 6.9— same as Figure 6.8 but for Net surface heat flux. Contour intervals for net heat flux is 5 watts m^{-2}

ever, over North America and over West Asia, the differences have increased after optimization. Precipitation shows slight less differences over Indian Ocean and over South America (as discussed in previous section too), however there are pockets of regions where the differences have increases such as in eastern Africa and over eastern China. Overall, there is slight improvement in the precipitation ($< 10\%$) in the optimized simulations. Differences have reduced in the Sea level pressure (Figure 6.12) simulations, mainly over Antarctic and Arctic regions. High positive and negative differences have now reduced after optimization in these regions. The only region where the differences have slightly increased are around 40N in the Pacific Ocean. Sea level pressure is also an indicator of surface winds, and therefore these results may suggest also the differences in surface winds in the optimized runs. Winds at 500hPa show less differences than control in the region between 40-60S , over central Africa and West Asia. In other regions, however, there is either no improvement or degradation seen. Overall, there is only slight improvement in 500hPa winds. The above results suggests that, although, overall the optimization procedures have resulted in reduced cost function, there are regions over the globe where the procedure has yielded negative results.

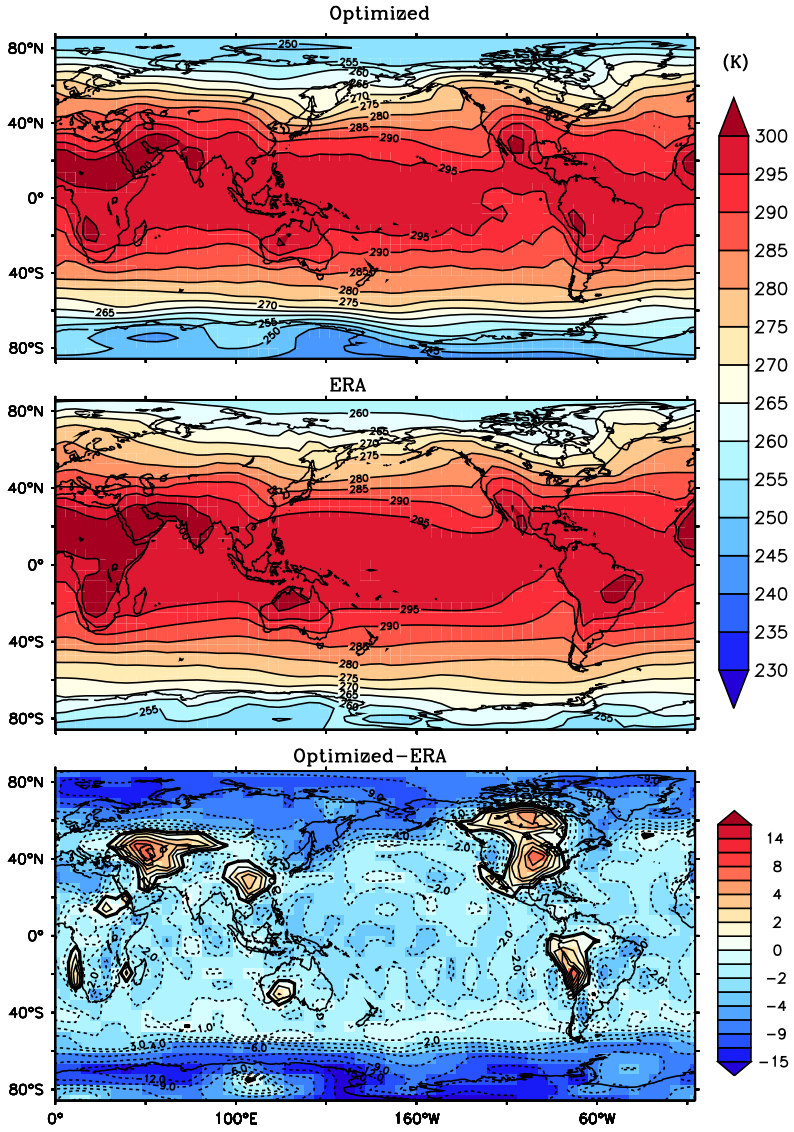

Figure 6.10— Annual mean Temperature (K) at 1000 hPa from (Top) PlaSim Optimized, (middle) ERA-Interim reanalysis data and (bottom) absolute difference (PlaSim Optimized - ERA-interim). PlaSim Optimized represents the annual mean state resulting from one year optimization run using SPSA as discussed in section 6.2. ERA-interim data represents annual mean of 21 years (1989-2009)

6.3 Conclusion

In this chapter, SPSA technique is applied to optimize the AGCM PlaSim. The experiments are performed with 15 control parameters simultaneously and for time

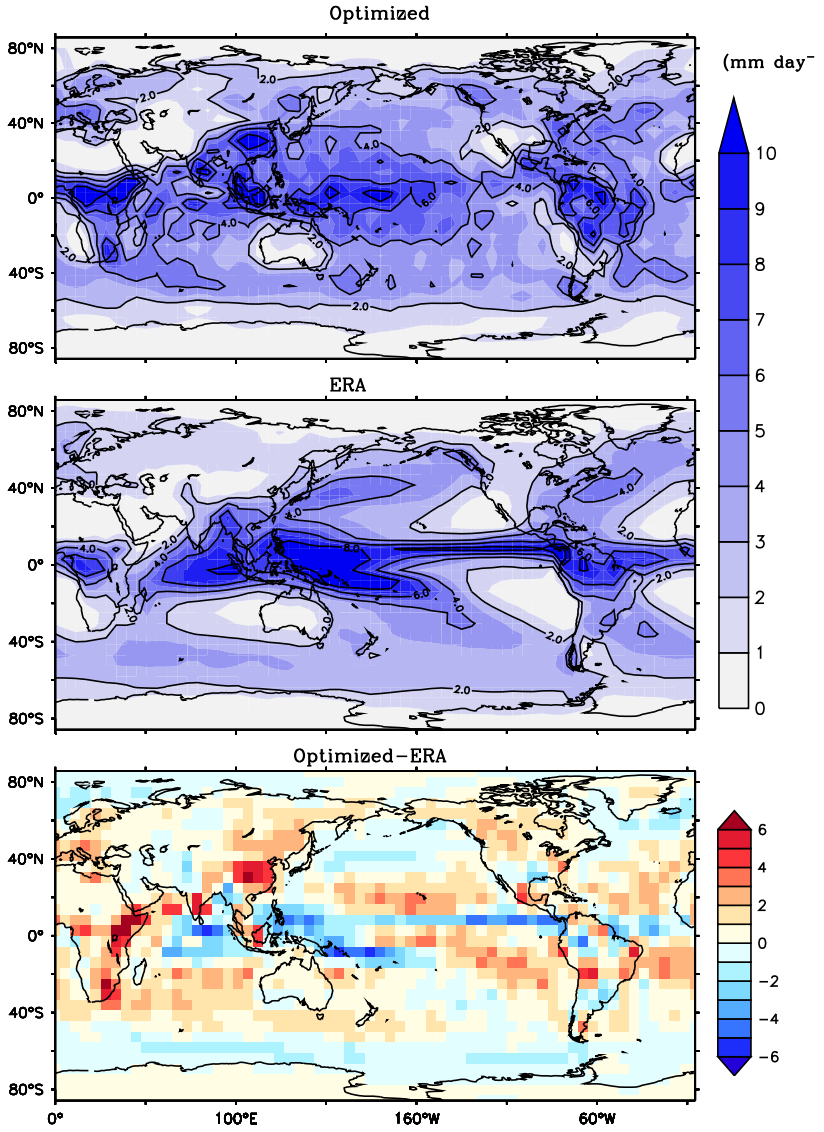

Figure 6.11— same as in 6.10 but for total precipitation ($mmday^{-1}$)

scale of 1 year using the *wet* configuration of PlaSim. The experiments are performed with synthetic as well as real data. SPSA technique is used to minimize the cost function comprising of contributions of mainly temperature, precipitation and heat fluxes.

SPSA technique can efficiently optimize parameters of the atmospheric model by finding the minimum cost function. The technique is robust and works well both with identical twin and reanalysis data. In identical twin experiments, the default

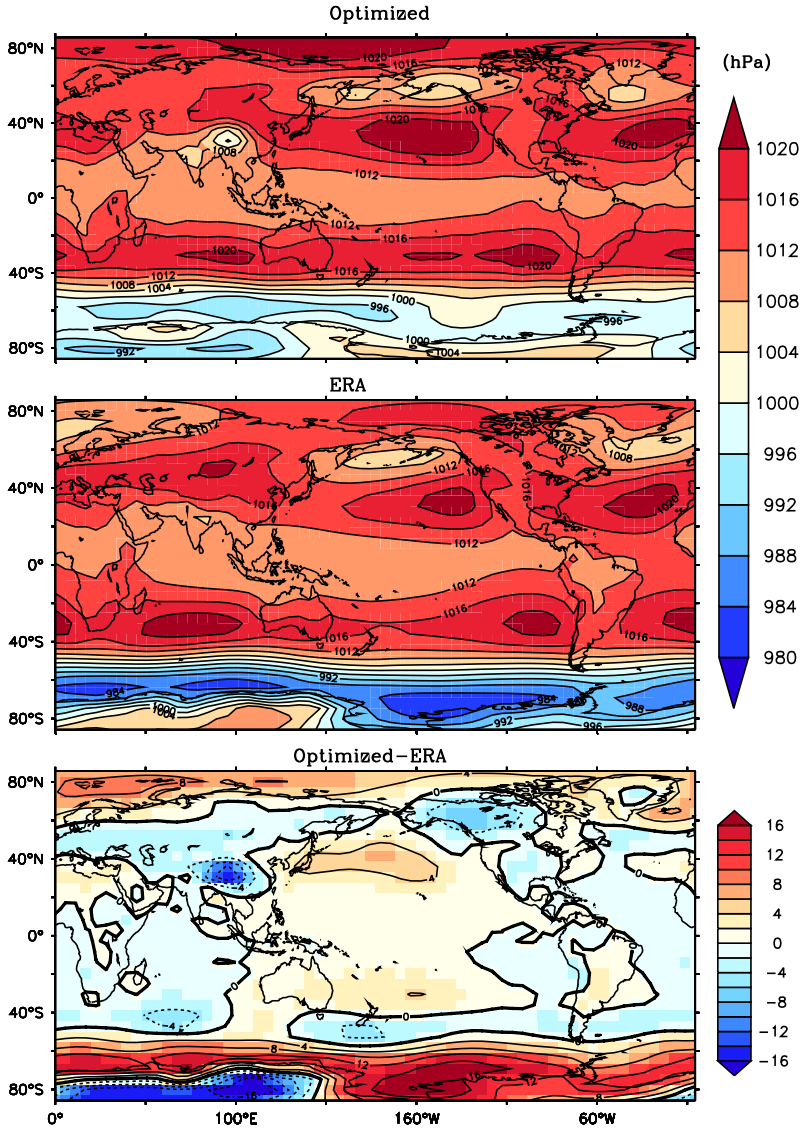

Figure 6.12— same as in 6.10 but for Sea Level Pressure (hPa)

values for 8 parameters are recovered and cost function is also reduced to acceptable range of minimum cost function. The results with ERA-Interim data shows an overall reduction in RMSE of $\sim 21\%$ while temperature shows an improvement up to 20%. Precipitation shows slight improvement especially in the Indian Ocean. The primary virtues of SPSA are the ease of implementation and the lack of need for a cost function gradient, theoretical and experimental support for relative efficiency, robustness to noise in the cost function measurements and ability to find a global

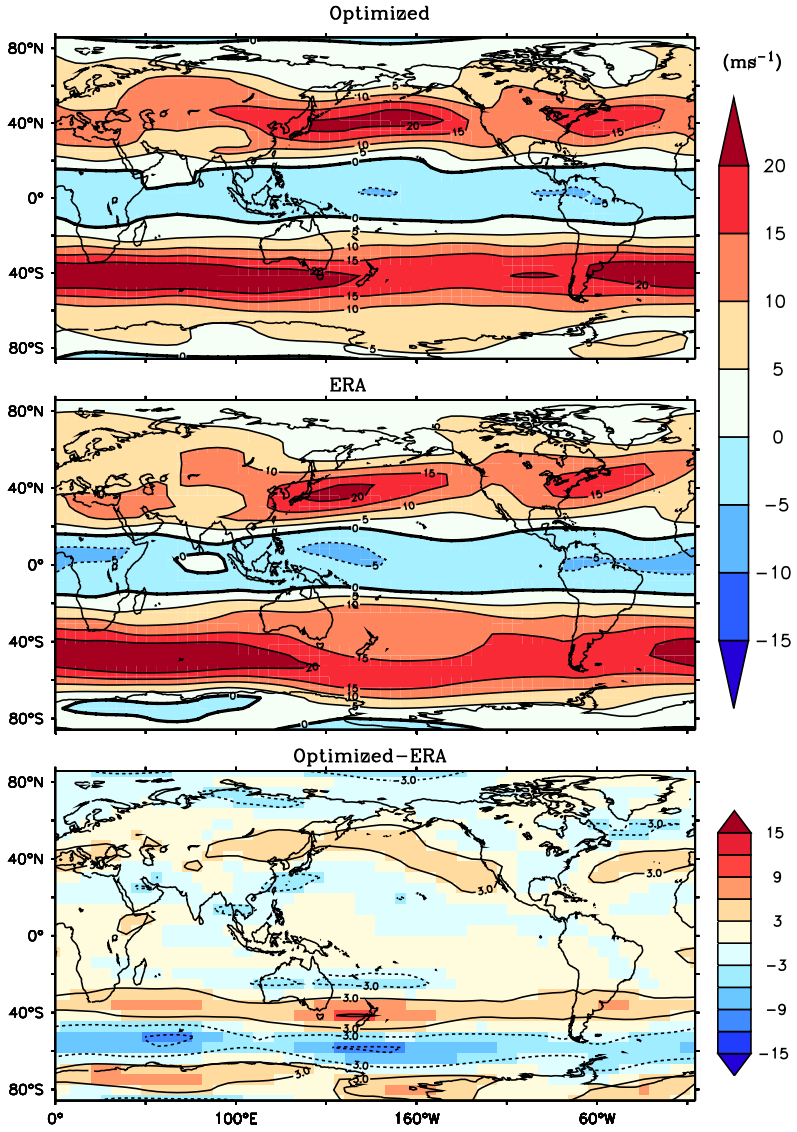

Figure 6.13— same as in 6.10 but for zonal winds at 500 hPa

minimum when multiple minima exist.

Chapter 7

Conclusions

Climate models are essential tools for climate study and future climate projections. Long term climate simulations results in the drifts of the simulated state with respect to the observations and hence the efforts are on,globally, for minimizing these differences between model and observations by means of data assimilation. This thesis is an effort in that direction, in which the atmospheric component, PlaSim, of the university of Hamburg earth system model, CESAM, is used to test the performance of two different approaches of tuning the model parameters through data assimilation. The basis of the work is an assumption according to which the errors in the model simulation are caused by the errors in the values of the parameters used in different parameterization schemes of the model. The parameter optimization is aimed for longer time scales so that the optimized model can be used effectively for climate simulations.

Before applying any optimization procedure, a sensitivity analysis of PlaSim was performed (explained in chapter 4 of the model) by perturbing the vector of control parameters over a wide range and then analyzing the changes in the model data difference (i.e. the cost function). This was done to understand the effect of uncertainties that are often associated with control parameters in models. In this study the sensitivity analysis of PlaSim was performed using the *wet*, the *dry* configurations with 15 parameters and for different integration time scales. It is found that discontinuities in the variation of cost function grew by including complex processes and by increasing model integration time. For longer model integration (up to 365 days) PlaSim's cost function which is computed using contributions from temperature, precipitation and heat flux components is highly sensitive to tuning parameter for

vertical diffusivity, point of mean transimittivity in layer and cloud related processes. Remaining parameters have smaller contribution to the cost function as compared to above mentioned parameters. An important information that is obtained from the sensitivity experiments is that range of the minimum cost function that could be achieved when the parameters are not precisely the same as control parameters. This information was very useful in optimization procedure in identical twin experiments.

The first technique used for optimizing PlaSim is a least square estimate based Green's function (GF) approach. In Greens function method (Menemenlis et al., 2005), each individual control parameter is perturbed one at a time and partial derivatives of cost function with respect to each parameter are obtained. Green's function approach was found to be effective in estimation of parameters under identical twin framework. One of the goals of the thesis was to inter-compare the performance of adjoint based and Green's function based optimization techniques. Following conclusions can be made in this regard:

- Implementation of Green function was quite simple whereas adjoint method is known to be quite complicated to implement and needs immense time for development.
- While adjoint based approach applied on PlaSim is limited by the length of the assimilation window even in the model where the non-linear processes are minimal (Blessing et al., 2014), Green's function approach was able to retrieve the parameters successfully even for longer time scales.
- For shorter integration time (7 days) Green's function method effectively takes less time to converge as compared to the adjoint method, however, for longer timescales (56 days) Green's function was more time consuming although it required less number of iterations to converge.

GF approach is advantageous over other parameter optimization methods (in comparison with Adjoint and EnKF) due to its ease of implementation. Not much efforts are required to calculate the gradients of the cost function. Therefore, this method can be applied to any numerical model with very few changes in the model.

GF approach, however, has some shortcomings. When applied to real observations (ERA-interim reanalysis, used in this study), this method fails to get an estimate of the parameters even for one day assimilation window. The reasons for this were not explored in this thesis, however according to Sumata et al. (2013) finite

difference based optimization techniques generally fail to work when the model's response is non-linear with respect to the control parameters. In our case, with real observations, it was found that the model data misfit is quite large and due to several local minima in the cost function, attaining optimization becomes quite difficult. This technique is based on linearization of model dynamics and therefore in highly non linear systems, this approach is not able to optimize parameters. In GF approach there is a possibility of finding a local minimum of cost function instead of unique minimum. To overcome this disadvantage, the perturbation applied for calculating gradients should be larger than the noise level in the cost function. Another major drawback in using this approach is that the computational time for each iteration is dependent on the size of the control vector. The cost for calculating gradients increases linearly with number of parameters to be estimated. So moderate and highly sensitive model parameters should be selected to enhance the performance of the GF approach.

It was seen that although the Green's function approach took less number of iterations for optimization, the actual time required for the entire procedure was larger than that for the adjoint based optimization approach. This is simply because the model has to be integrated separately for each perturbed parameter. With this constraint GF approach is unfit for any practical application in atmospheric models for parameter optimization especially where we want to consider optimization of a large number of parameters.

Stochastic techniques are known to be an efficient tool for the solution of optimization problems in case of chaotic models. The second technique used in this study is the simultaneous perturbation stochastic approximation (SPSA). In SPSA the gradient is estimated using only two measurements per iteration regardless of the dimension of the optimization problem. Unlike GF, in SPSA all considered parameters are perturbed simultaneously. It is found that SPSA takes more number of iterations to find minimum cost function range as compare to the GF in identical twin experiments. However, in SPSA the computational time for each iteration is independent of the size of the control vector.

SPSA was applied to ERA-Interim reanalysis data and unlike the GF approach, cost function reduction was achieved. The optimized set of parameters was then used to generate the optimized state which was then compared with the observations. It was found that the errors in temperature have reduced (as compared to the control state) at almost all levels barring second level just below the stratosphere. Overall there was $\sim$10-20% of reduction of error in temperature at various levels in the

atmosphere. Most of error reduction in surface temperature occurred in the equatorial West Pacific, North Atlantic and Southern Indian Ocean regions. Net heat flux improved almost in every region with large errors of $> 50\ Watts/m^2$ coming down within 10 $Watts/m^2$ in the North Pacific, south-east Africa and also in South Atlantic ocean. However there was not much improvement in total precipitation, however there are some regions in equatorial west Pacific and in the Indian Ocean region where the RMSE in total precipitation has decreased. Overall SPSA method successfully optimized PlaSim both in identical twin as well as in real experiments.

A major limitation of all stochastic methods is the large number of iterations that are required for optimization. A recent research (Schirber et al., 2013) demonstrates that single parameter estimation (using Ensemble Kalman Filter, EnKF) for shorter time scales (30 days) is quite successful. The gradient based algorithms (Adjoint, EnKF) are faster to converge than any cost function based gradient approximation (SPSA algorithm).

7.0.1 Closing Remarks

An attempt to use two different optimization methods, one based on gradient descent approach and another based on gradient approximation approach was made in this thesis. The aim was to explore an alternate assimilation approach along with already existing 4-D variational assimilation system in PlaSim. With the experiments performed in this experiments, it is found that the stochastic approach of optimization works quite effectively in bringing down the model data differences. The major advantage of using SPSA has been its simple formulation and ease of implementation. In this aspect SPSA scores over many other existing techniques such as the Kalman filter, variational methods etc. There are however certain limitations which are quite serious and might require further efforts to make such techniques practical for routine use. The choice of control parameters is very crucial in determining the success of these optimization techniques. Many more experiments with different initial parameter values must be performed in future to get even better results. Another significant point of concern is that in this study a very coarse resolution model is used for the experiments. For any practical application the resolution of the model has to be fine enough so that it can capture meso to sub-mesoscale variability of the climate system. Applying SPSA in this framework where the model resolution is fine

is a challenging task as the number of iterations required for successful optimization is quite large.

In this study, SPSA technique has been successfully applied for the first time to an atmospheric model, however, because of the above mentioned limitations, it might be possible in near future with the development of computer infrastructure and parallel computations to use it practically for our real world applications.

List of Figures

List of Tables

Bibliography

Altaf, M. U., Heemink, A. W., Verlaan, M., and Hoteit, I.: Simultaneous perturbation stochastic approximation for tidal models, Ocean Dynamics, 61, 1093–1105, 2011.

Anderson, J. L.: An ensemble adjustment Kalman filter for data assimilation, Monthly weather review, 129, 2884–2903, 2001.

Annan, J. and Hargreaves, J.: Efficient estimation and ensemble generation in climate modelling, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365, 2077–2088, 2007.

Annan, J., Hargreaves, J., Edwards, N., and Marsh, R.: Parameter estimation in an intermediate complexity Earth system model using an ensemble Kalman filter, Ocean modelling, 8, 135–154, 2005.

Asselin, R.: Frequency filter for time integrations, Monthly Weather Review, 100, 487–490, 1972.

Bengtsson, L., Ghil, M., and Källén, E.: Dynamic meteorology: data assimilation methods, vol. 36, Springer New York, NY, 1981.

Blessing, S., Kaminski, T., Lunkeit, F., Matei, I., Giering, R., Köhl, A., Scholze, M., Herrmann, P., Fraedrich, K., and Stammer, D.: Testing variational estimation of process parameters and initial conditions of an earth system model, Tellus A, 66, 2014.

Challis, L. and Sheard, F.: The green of green functions, Physics Today, 56, 41–46, 2003.

Chin, D. C.: Comparative study of stochastic algorithms for system optimization based on gradient approximations, Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 27, 244–249, 1997.

Daley, R.: Atmospheric data analysis, cambridge atmospheric and space science series, Cambridge U niversity Press, 4, 57, 1991.

Daley, R.: Atmospheric data analysis, 2, Cambridge university press, 1993.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., et al.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, 2011.

Evensen, G.: Data assimilation: the ensemble Kalman filter, Springer, 2009.

Fabian, V.: Stochastic approximation, Department of Statistics and Probability, Michigan State University, 1971.

Fan, S.-M., Sarmiento, J. L., Gloor, M., and Pacala, S. W.: On the use of regularization techniques in the inverse modeling of atmospheric carbon dioxide, Journal of Geophysical Research: Atmospheres (1984–2012), 104, 21 503–21 512, 1999.

Fletcher, R. and Powell, M. J.: A rapidly convergent descent method for minimization, The Computer Journal, 6, 163–168, 1963.

Fraedrich, K., Jansen, H., Kirk, E., Luksch, U., and Lunkeit, F.: The Planet Simulator: Towards a user friendly model, Meteorologische Zeitschrift, 14, 299–304, 2005.

Gerencsér, L. and Vágó, Z.: The mathematics of noise-free SPSA, in: Decision and Control, 2001. Proceedings of the 40th IEEE Conference on, vol. 5, pp. 4400–4405, IEEE, 2001.

Giering, R. and Kaminski, T.: Recipes for adjoint code construction, ACM Transactions on Mathematical Software (TOMS), 24, 437–474, 1998.

Gloor, M., Gruber, N., Hughes, T., and Sarmiento, J. L.: Estimating net air-sea fluxes from ocean bulk data: Methodology and application to the heat cycle, Global biogeochemical cycles, 15, 767–782, 2001.

Gonzalez, O. R., Küper, C., Jung, K., Naval, P. C., and Mendoza, E.: Parameter estimation using Simulated Annealing for S-system models of biochemical networks, Bioinformatics, 23, 480–486, 2007.

Griewank, A. et al.: On automatic differentiation, Mathematical Programming: recent developments and applications, 6, 83–107, 1989.

Grosfeld, K., Lohmann, G., and Rimbu, N.: The impact of Atlantic and Pacific Ocean sea surface temperature anomalies on the North Atlantic multidecadal variability, Tellus A, 60, 728–741, 2008.

Hoskins, B. and Simmons, A.: A multi-layer spectral model and the semi-implicit method, Quarterly Journal of the Royal Meteorological Society, 101, 637–655, 1975.

Houghton, J. T., Meira Filho, L. G., Griggs, D. J., and Maskell, K.: An introduction to simple climate models used in the IPCC second assessment report, Intergovernmental Panel on Climate Change Geneva, 1997.

Hutchison, D. W. and Hill, S. D.: Simulation optimization of airline delay with constraints, in: Simulation Conference, 2001. Proceedings of the Winter, vol. 2, pp. 1017–1022, IEEE, 2001.

Iacobellis, S. F., McFarquhar, G. M., Mitchell, D. L., and Somerville, R. C.: The sensitivity of radiative fluxes to parameterized cloud microphysics, Journal of climate, 16, 2979–2996, 2003.

Jackson, C., Sen, M. K., and Stoffa, P. L.: An efficient stochastic Bayesian approach to optimal parameter and uncertainty estimation for climate model predictions, Journal of Climate, 17, 2828–2841, 2004.

Kalnay, E.: Atmospheric modeling, data assimilation and predictability, Cambridge university press, 2003.

Kiefer, J. and Wolfowitz, J.: Stochastic estimation of the maximum of a regression function, The Annals of Mathematical Statistics, 23, 462–466, 1952.

Köhl, A. and Willebrand, J.: An adjoint method for the assimilation of statistical characteristics into eddy-resolving ocean models, Tellus A, 54, 406–425, 2002.

Kong, X., Yang, Y., Chen, X., Shao, Z., and Gao, F.: Quality Control via Model-Free Optimization for a Type of Batch Process with a Short Cycle Time and Low Operational Cost, Industrial & Engineering Chemistry Research, 50, 2994–3003, 2011.

Lakshmanan, S. and Derin, H.: Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 11, 799–813, 1989.

Lea, D. J., Allen, M. R., and Haine, T. W.: Sensitivity analysis of the climate of a chaotic system, Tellus A, 52, 523–532, 2000.

Lea, D. J., Haine, T. W., Allen, M. R., and Hansen, J. A.: Sensitivity analysis of the climate of a chaotic ocean circulation model, Quarterly Journal of the Royal Meteorological Society, 128, 2587–2605, 2002.

Lewis, A. and Bridle, S.: Cosmological parameters from CMB and other data: A Monte Carlo approach, Physical Review D, 66, 103 511, 2002.

Li, G. and Reynolds, A. C.: Uncertainty quantification of reservoir performance predictions using a stochastic optimization algorithm, Computational Geosciences, 15, 451–462, 2011.

Lorenc, A. C.: Analysis methods for numerical weather prediction, Quarterly Journal of the Royal Meteorological Society, 112, 1177–1194, 1986.

Lucarini, V., Fraedrich, K., and Lunkeit, F.: Thermodynamic analysis of snowball earth hysteresis experiment: efficiency, entropy production and irreversibility, Quarterly Journal of the Royal Meteorological Society, 136, 2–11, 2010.

Malanotte-Rizzoli, P.: Modern approaches to data assimilation in ocean modeling, vol. 61, Elsevier, 1996.

Mauritsen, T., Stevens, B., Roeckner, E., Crueger, T., Esch, M., Giorgetta, M., Haak, H., Jungclaus, J., Klocke, D., Matei, D., et al.: Tuning the climate of a global model, Journal of Advances in Modeling Earth Systems, 4, 2012.

Menemenlis, D. and Wunsch, C.: Linearization of an oceanic general circulation model for data assimilation and climate studies, Journal of Atmospheric and Oceanic Technology, 14, 1420–1443, 1997.

Menemenlis, D., Fukumori, I., and Lee, T.: Using Green's functions to calibrate an ocean general circulation model, Monthly weather review, 133, 1224–1240, 2005.

Menke, W.: Geophysical data analysis: discrete inverse theory, vol. 45, Access Online via Elsevier, 1989.

Murphy, J. M., Sexton, D. M., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A.: Quantification of modelling uncertainties in a large ensemble of climate change simulations, Nature, 430, 768–772, 2004.

Nguyen, A. T., Menemenlis, D., and Kwok, R.: Arctic ice-ocean simulation with optimized model parameters: Approach and assessment, Journal of Geophysical Research: Oceans (1978–2012), 116, 2011.

Orszag, S.: Transform method for the calculation of vector-coupled sums: Application to the spectral form of the vorticity equation, Journal of the Atmospheric Sciences, 27, 890–895, 1970.

Robert, A.: A stable numerical integration scheme for the primitive meteorological equations, Atmosphere-Ocean, 19, 35–46, 1981.

Robinson, A. R. and Lermusiaux, P. F.: Overview of data assimilation, Harvard reports in physical/interdisciplinary ocean science, 62, 2000.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model, Journal of Climate, 19, 3771–3791, 2006.

Romanova, V., Lohmann, G., and Grosfeld, K.: Effect of land albedo, CO2, orography, and oceanic heat transport on extreme climates, Climate of the Past, 2.{www. clim-past. net/2/31/2006/}, 31, 2006.

Roscher, M., Stordal, F., and Svensen, H.: The effect of global warming and global cooling on the distribution of the latest Permian climate zones, Palaeogeography, Palaeoclimatology, Palaeoecology, 309, 186–200, 2011.

Schirber, S., Klocke, D., Pincus, R., Quaas, J., and Anderson, J. L.: Parameter estimation using data assimilation in an atmospheric general circulation model: From a perfect toward the real world, Journal of Advances in Modeling Earth Systems, 2013.

Schmittner, A., Silva, T., Fraedrich, K., Kirk, E., and Lunkeit, F.: Effects of Mountains and Ice Sheets on Global Ocean Circulation*, Journal of Climate, 24, 2814–2829, 2011.

Schwartz, S. E., Smith, T. M., Karl, T. R., and Reynolds, R. W.: Uncertainty in climate models, Science, 296, 2139–2140, 2002.

Severijns, C. and Hazeleger, W.: Optimizing parameters in an atmospheric general circulation model, Journal of climate, 18, 3527–3535, 2005.

Simmons, A., Hoskins, B., and Burridge, D.: Stability of the semi-implicit method of time integration(for atmospheric temperature data), Monthly Weather Review, 106, 405–412, 1978.

Spall, J. C.: Implementation of the simultaneous perturbation algorithm for stochastic optimization, Aerospace and Electronic Systems, IEEE Transactions on, 34, 817–823, 1998.

Spall, J. C.: Adaptive stochastic approximation by the simultaneous perturbation method, Automatic Control, IEEE Transactions on, 45, 1839–1853, 2000.

Stammer, D. and Wunsch, C.: The determination of the large-scale circulation of the Pacific Ocean from satellite altimetry using model Green's functions, Journal of Geophysical Research: Oceans (1978–2012), 101, 18 409–18 432, 1996.

Sugiura, N., Awaji, T., Masuda, S., Mochizuki, T., Toyoda, T., Miyama, T., Igarashi, H., and Ishikawa, Y.: Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations, Journal of Geophysical Research: Oceans, 113, 2008.

Sumata, H., Kauker, F., Gerdes, R., Koeberle, C., and Karcher, M.: A comparison between gradient descent and stochastic approaches for parameter optimization of a sea ice model, Ocean Science, 9, 609–630, 2013.

Uppala, S. M., Kållberg, P., Simmons, A., Andrae, U., Bechtold, V., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G., et al.: The ERA-40 re-analysis, Quarterly Journal of the Royal Meteorological Society, 131, 2961–3012, 2005.

Wunsch, C.: The ocean circulation inverse problem, Cambridge University Press, 1996.

Yao, L. and Sethares, W. A.: Nonlinear parameter estimation via the genetic algorithm, Signal Processing, IEEE Transactions on, 42, 927–935, 1994.