Analysing Recent Spatial and Temporal Atmospheric Methane Variations using Forward and Inverse Modelling

Christopher James Wilson

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Earth and Environment
September 2011

Declaration of Authorship

The candidate confirms that the work submitted is his/her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapters 4, 6 and 7 contain simulations of sulfur hexafluoride, methyl chloroform and methane which use parameters provided by the TransCom methane model inter-comparison. This model study provided standardised emission inventories, destruction fields and initial concentration fields which are detailed in **Patra, P. K., S. Houweling, M. Krol, P. Bousquet, D. Belikov, D. Bergmann, H. Bian, P. Cameron-Smith, M. P. Chipperfield, K. Corbin, A. Fortems-Cheiney, A. Fraser, E. Gloor, P. Hess, A. Ito, S. R. Kawa, R. M. Law, Z. Loh, S. Maksyutov, L. Meng, P. I. Palmer, R. G. Prinn, M. Rigby, R. Saito and C. Wilson, 2011: TransCom model simulations of CH_4 and related species: linking transport, surface flux and chemical loss with CH_4 variability in the troposphere and lower stratosphere.** ***Atmospheric Chemistry and Physics Discussions*****, 11, 18767-18821, doi:10.5194/acpd-11-18767-2011**. The candidate provided the results of simulations of methane, methyl chloroform and sulfur hexafluoride, among other species, to the model intercomparison described in the above paper.

Stratospheric observations of sulfur hexafluoride were provided for comparison with model simulations as described in **Brown, A. T., M. P. Chipperfield, C. Boone, C. Wilson, K. A. Walker and P. Bernath, 2011: Trends in atmospheric halogen containing gases since 2004.** ***Journal of Quantitative Spectroscopy and Radiative Transfer*****, 112, 2552-2566**. The candidate provided results of a simulation of tropospheric sulfur hexafluoride for use as a boundary condition for further model simulations.

Abstract

The TOMCAT 3-D chemical transport model (CTM) has been used to investigate the cause of recent variations in global atmospheric methane (CH_4), focusing on examining changes in the balance of sources and sinks of the species. The chemical loss, transport and emissions of methane have been studied and a new 4D-Var inverse version of TOMCAT has been created.

The accuracy of the TOMCAT model transport was investigated by simulating the distribution of the long-lived species SF_6. A range of model grid resolutions, boundary layer schemes and advection schemes were tested. New retrievals from the Atmospheric Chemistry Experiment (ACE) satellite instrument were used to test the model in the upper troposphere and lower stratosphere. The standard CTM simulated the observed distribution and growth of SF_6 well. However, based on comparison with ground-based data, the interhemispheric transport in the TOMCAT model was found to be approximately 20% too slow, with too little temporal variation in southern hemisphere transport. On the whole, however, tracer transport in the CTM using its standard set-up was accurate. As a basis for the inverse model simpler advection and boundary layer (BL) schemes were tested. The advection scheme which conserved only up to first-order moments (rather than second-order moments) did not significantly reduce the accuracy of the model transport. However, use of a local boundary layer mixing scheme rather than a non-local scheme did degrade the quality of the transport by reducing the speed of vertical mixing out of the BL.

A number of currently used CH_4 emission inventories were used with the forward TOMCAT model in order to examine the effect they have on the global CH_4 budget, and two different estimates of the OH sink were also tested. A published OH field derived from global CH_3CCl_3 and a chemical box model was found to be more consistent with OH observations than the field from the full chemistry TOMCAT model. Although both OH fields produced global CH_4 lifetimes consistent with published estimates, the TOMCAT OH field yielded model CH_4 which was up to 100 ppb higher than observations at the surface. Data assimilation was used to improve the estimate of the stratospheric sink of CH_4. Although this sink is small overall, it needs to be represented realistically in order to accurately reproduce global CH_4 to within 10 ppbv.

A new adjoint version of the TOMCAT model was produced by explicit coding, and was thoroughly tested. This was incorporated into a new 4D-Var inverse model which can be used to produce updated CH_4 surface flux estimates which are constrained to agree with atmospheric observations. The inverse model was used to investigate emissions in the Arctic where the forward TOMCAT model and standard emissions revealed a seasonal cycle out of phase with surface CH_4 observations. It was found that northern hemisphere summertime wetland emissions were overestimated in the GISS inventory by up to 100% for the period 2000-2006, and that this was likely due to the estimates of emission rates and thaw period used when producing that inventory. It was also found that increased Asian emissions suggested in the EDGAR V4.0 inventory are not consistent with observations unless mitigated by a corresponding drop in emissions elsewhere.

Acknowledgements

I would first like to express my greatest thanks to my supervisors, Martyn Chipperfield and Manuel Gloor, for their help, guidance, patience and time, without which I would not have been able to produce any part of this thesis.

I would like to express my thanks to Frederic Chevallier at the Laboratoire des Sciences du Climat et l'Environment for taking the time to teach me how to write an inverse model, and for replying to my numerous emails with both politeness and haste.

I have had plenty of help from the members of Team TOMCAT past and present, so I would like to thank Ryan, Mike, Tom, Sandip, Wuhu, Beatriz and Steve. I'd particularly like to thank Lara and Nige for teaching me how to use IDL and for extensive useful work-based discussions and inane conversions about other stuff along the way. Among the wider community I'd like to thank Annemarie Fraser and John Miller for their help at various points.

Kev and Ian were the best house-mates I could have wished for, and I'd like to thank anyone who helped keep me sane by playing sport with me: the whole lot of the five-a-side guys, anyone who played squash with me and especially Mark for going climbing every week. Everybody in the School of Earth and Environment is friendly and encouraging, and I'd like to express my thanks to them all for that.

I would also like to acknowledge the financial assistance I received from the UK National Environmental Research Council.

Special thanks goes to Sarah, without whom, for a thousand reasons, I wouldn't have gotten close to finishing this thesis.

Finally, my mum, my dad and my sister Jade have provided me with unconditional support of every imaginable form, not only throughout the past four years, but always, and I would never have made it even to the start line without them. Thank you.

Contents

List of Figures

List of Tables

Abbreviations

ACE	**A**tmospheric **C**hemistry **E**xperiment
ADM	**AD**joint **M**odel
CTM	**C**hemical **T**ransport **M**odel
EDGAR	**E**mission **D**atabase for **G**lobal **A**tmospheric **R**esearch
ESRL	**E**arth **S**ystem **R**esearch **L**aboratory
FOM	**F**irst **O**rder **M**oments
FTS	**F**ourier **T**ranform **S**pectrometer
HALOE	**HAL**ogen **O**ccultation **E**xperiment
MODIS	**MO**derate **R**esolution **I**maging **S**pectroradiometer
NOAA	**N**ational **O**ceanic and **A**tmospheric **A**dministration
PBL	**P**lanetary **B**oundary **L**ayer
RMSD	**R**oot **M**ean **S**quare **D**ifference
SOM	**S**econd **O**rder **M**oments
TLM	**T**angent **L**inear **M**odel

Chapter 1

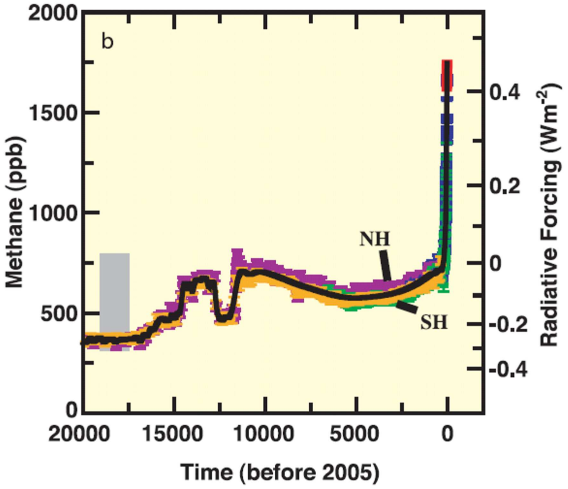
Introduction

1.1 Motivation

In recent years, our planet's climate system has been placed under greater scrutiny than ever before due to the growing concern that human activity has influenced its composition, and hence its temperature and dynamics, to an unprecedented extent. Anthropogenic activities, such as industrial processes, agriculture and land use change have resulted in rising atmospheric concentrations of greenhouse gases such as carbon dioxide (CO_2), methane (CH_4) and water vapour (H_2O), amongst others. Greenhouse gases are known to absorb and re-emit long-wave infra-red radiation which would otherwise escape the Earth's atmosphere, producing a warming effect upon the planet. As atmospheric concentrations of greenhouse gases rise, this greenhouse effect on the planet also increases, and may lead to global mean surface temperatures rising by 4 ± 2 °C by the year 2100 (Meehl et al., 2007).

It is clearly vital that our understanding of the Earth's atmospheric processes and composition is as comprehensive and as detailed as possible in order that we may attempt to control our influence upon the atmosphere and mitigate for climate change. However, the complex and turbulent nature of atmospheric processes, along with the atmosphere's dependence upon the planet's oceanic and bioterrestrial systems, means that there currently remain large uncertainties in our comprehension. Attempts to improve our understanding of the composition of the atmosphere take a variety of forms, including direct measurement of atmospheric concentrations, laboratory experiments of atmospheric species, and model studies which take advantage of recent improvements in computational power in order to simulate the atmospheric system. While atmospheric modelling is a powerful tool in reproducing or predicting the composition of the atmosphere at a given time, every model requires some prior knowledge of certain parameters, such as surface emission distributions and atmospheric reaction rates, in order to function accurately and gaps still remain in

Figure 1.1 The concentration (ppb) and associated radiative forcing (W m^{-2}) of CH_4 over the last 20,000 years reconstructed from Antarctic and Greenland ice and firn data (symbols) and from direct atmospheric measurements (red line). The grey bar displays the reconstructed CH_4 concentration range over the past 650,000 years. Taken from Jansen et al. (2007).

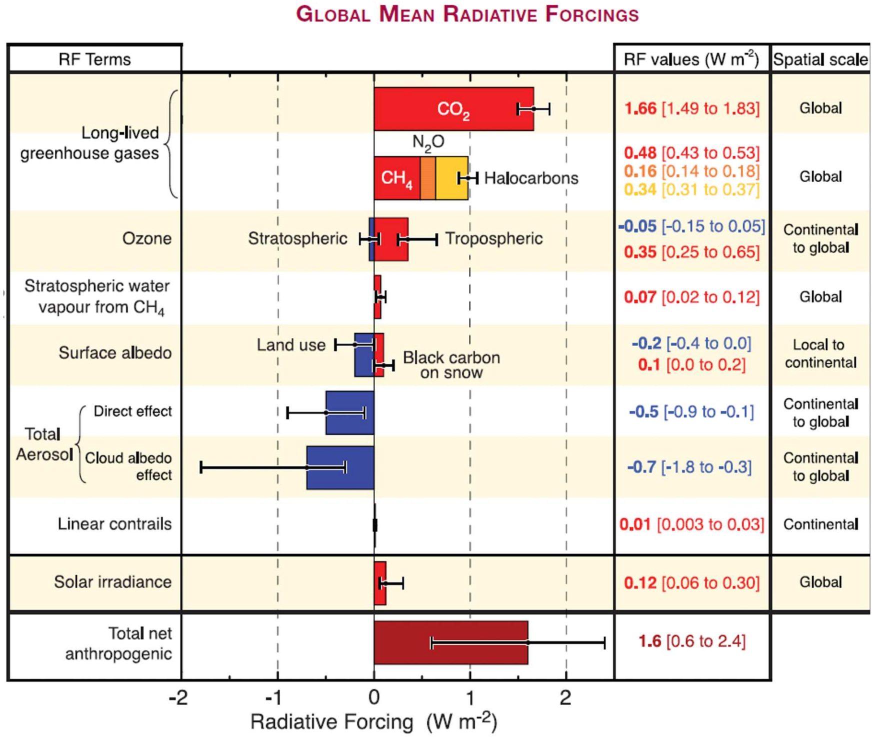


our knowledge of such parameters. However, atmospheric models allow us to produce accurate three-dimensional estimates of atmospheric species without the need for extensive measurement campaigns, and allow us to predict the effect that anthropogenic activities may have upon the future climate of our planet.

1.2 Tropospheric Methane

Methane (CH_4) is a greenhouse gas which is emitted from a variety of anthropogenic and natural sources. Figure 1.1 shows the global mean concentration of CH_4 over the past 20,000 years. Prior to the industrial revolution, which was a period of dramatic technological and manufactural growth throughout the 18th and 19th centuries, the mean atmospheric CH_4 concentration varied between 400 parts per billion (ppb) during glacial periods and 700 ppb in inter-glacial periods (Spahni et al., 2005). Since industrialisation, however, the mean atmospheric concentration of CH_4 has climbed dramatically, at an unprecedented pace, with an increase of 1000 ppb during the 20th century (Ferretti et al., 2005). The IPCC report of 2007 (Forster et al., 2007) used measurements from surface sampling sites to place the the global mean CH_4 concentration at approximately 1774 ppb, a value unprecedented in at least the preceding 650,000 years (Spahni et al., 2005), and since 2007 the global atmospheric CH_4 concentration has continued to rise (Rigby et al., 2008; Dlugokencky et al., 2009).

Figure 1.2 Summary of the principal components of the radiative forcing of climate change. Values represent the forcings in 2005 relative to the start of the industrial era (around 1750). Errors correspond to the level of uncertainty around the respective value. Taken from Forster et al. (2007).



This increase in the atmospheric concentration of methane has a direct effect upon the Earth's climate due to the high radiative forcing produced by CH_4. Radiative forcing is a measure of how much a factor, such as the changing concentration of a greenhouse gas, influences the energy balance of the Earth's climate. Figure 1.2 displays the radiative forcing of various atmospheric constituents, showing that CH_4 is the second most important greenhouse gas in terms of radiative forcing (Denman et al., 2007), behind only CO_2 in its influence on the climate. In fact, on a per-molecule basis, the radiative efficiency of CH_4 is approximately 26 times higher than that of CO_2 but the atmospheric concentration of CH_4 is far lower, meaning that its total radiative forcing is approximately 0.48 ± 0.05 W m^{-2}, compared to a value of 1.66 ± 0.17 W m^{-2} for CO_2. This means that CH_4 accounts for approximately 30% of the planet's total radiative forcing due to anthropogenic activity (Forster et al., 2007).

1.3 Inverse Modelling

Surface emissions of CH_4 originate from a range of contributing biogenic sources, including wetlands, rice agriculture and livestock, and non-biogenic sources such as the fossil fuel industry and biomass burning. It is important that the nature, location and magnitude of these emissions are fully understood in order for policy-makers to be able to reduce anthropogenic contribution to climate change, but also so that atmospheric models may use such inventories as boundary conditions in order to accurately reproduce atmospheric concentrations of the species. However, there are currently large uncertainties in our emission estimates for CH_4, especially from natural processes such as biogenesis from wetlands or those with varying geographical distributions such as biomass burning (Forster et al., 2007).

The rate of CH_4 emission from a given process can generally be estimated using three methods. The first of these involves the extrapolation of direct small-scale flux measurements up to a regional scale, while the second method consists of modelling the underlying processes of CH_4 emissions. While each of these 'bottom-up' methods may be accurate up to a limited point, the large spatial and temporal variation in CH_4 emission rates may lead to inaccuracies when these estimates are extrapolated upwards. The third available option is a 'top-down' approach known as inverse transport modelling, which assimilates observed atmospheric concentrations into atmospheric models in order to constrain emission rates. While top-down methods may help to overcome the weaknesses in the bottom-up approaches, inverse models are dependent on good quality, high density observations and a good understanding of errors in the observation and modelling process in order to produce accurate results. Ideally, the two approaches work in tandem to create the most robust emission inventories possible.

One method of inverse transport modelling which is becoming increasingly prevalent due to to improvements in computational speed and memory is known as the Four-Dimensional Variational (4D-Var) inverse modelling method. This method uses an adjoint version of an atmospheric transport model to find the sources of discrepancies between model simulations and observations the distribution of atmospheric species. The aim of the work described in this thesis is to create a new 4D-Var inversion system for the TOMCAT/SLIMCAT model, henceforth known as the TOMCAT model. This is a Chemical Transport Model (CTM) which has been used in a variety of studies in the past to simulate the transport and chemistry of active atmospheric trace gases (e.g. Chipperfield et al. (1993); Arnold et al. (2005); Breider et al. (2010); Hossaini et al. (2010)), and to use the new inversion system to create updated emission rates for CH_4 which are more consistent with observations, providing information which can help to improve our understanding of the physical processes of CH_4 emission into the atmosphere. An essential step in the model development is an evaluation of the forward model for the transport of long-lived gases such as CH_4.

1.4 Aims of this Thesis

The main aim of this thesis is to improve our understanding of the growth-rate of atmospheric CH_4 and the underlying processes contributing to its emission into the atmosphere using observed atmospheric CH_4 concentrations to constrain global surface emission rates using the TOMCAT CTM. Specifically, this thesis will address the following:

1. **Investigate recent variations in the atmospheric CH_4 budget using the TOMCAT model and a range of state-of-the-art CH_4 emission inventories.** Investigate the ability of the TOMCAT model to accurately simulate the global CH_4 distribution. Assess the accuracy of a variety of CH_4 emission and destruction inventories. Assess the impact that differences in the emission/destruction inventories have on the simulated CH_4 distribution. Investigate the sources of any discrepancies between the forward model and the observations.
2. **Develop a 4D-Var inversion system using the TOMCAT CTM capable of producing improved global emission inventories of atmospheric species on a model-grid scale through the assimilation of observed data.** Assess the accuracy of the simulated atmospheric transport processes in the TOMCAT model. Produce and validate the adjoint version of the TOMCAT model. Incorporate the adjoint model into a 4D-Var optimisation routine.
3. **Use the TOMCAT 4D-Var system together with observed CH_4 data in order to provide an updated emission inventory for CH_4.** Trace the geographical and temporal locations of discrepancies produced using the inversion model, and assess their physical processes which are responsible. Deduce possible reasons for the inconsistencies and use this knowledge to infer the causes of variations in the atmospheric CH_4 budget.

1.5 Thesis Layout

This thesis consists of seven further chapters. Chapter 2 discusses the dynamics and chemistry of the troposphere with respect to CH_4, before describing in detail the various sources and sinks of the species. Chapter 2 also details recent variations in the atmospheric CH_4 budget and discusses previous studies which have attempted to explain these variations with respect to changes in the net CH_4 flux. Chapter 3 describes and evaluates different methods available to inverse modellers in order to estimate the fluxes of atmospheric constituents, and briefly discusses the results of previous inverse modelling studies. Chapter 3 also describes the TOMCAT CTM which is used in this study. Chapter 4 evaluates the accuracy of the transport in the TOMCAT model using observations of a chemically inert tracer, and discusses this model's suitability for use as the basis of an inverse transport model. Chapter 5 documents the development of a new 4D-Var inversion system for the TOMCAT model, giving detailed descriptions of both the methodology used for an inversion and the extensive testing process undergone by the inverse model in order to validate its accuracy. Chapter 6 investigates recent variations in tropospheric CH_4 by examining the sensitivity of the TOMCAT model to a variety of estimates of both sources and sinks of methane. Chapter 7 foucuses on the seasonal cycle of CH_4 in the Arctic, and examines the results of an inverse model simulation which aims to produce updated global emission estimates of CH_4 during a specific year, before examining the physical implications of the results of the inversion. Chapter 8 summarises the results and conclusions of the research in the thesis.

Chapter 2

Methane in the Atmosphere

2.1 Introduction

This chapter is intended to give background on all aspects of atmospheric CH_4, as it is clearly necessary to understand the underlying processes of atmospheric transport, chemistry, emission and destruction of the species in order to accurately simulate its global concentration in an atmospheric transport model. Section 2.2 describes the structure of the atmosphere and gives a brief description of the underlying transport processes which atmospheric CH_4 is subject to, while Section 2.3 describes the chemical destruction processes which CH_4 undergoes in the atmosphere. Section 2.4 gives details of the wide range of sources from which CH_4 is emitted, along with the inherent problems encountered when attempting to compile CH_4 emission budgets due to difficulties in quantifying these processes. Finally, Section 2.5 describes recent variations in the global mean atmospheric CH_4 concentration, and discusses previous attempts by other authors to link these variations to changes in surface emission rates and distributions.

2.2 The Structure of the Atmosphere

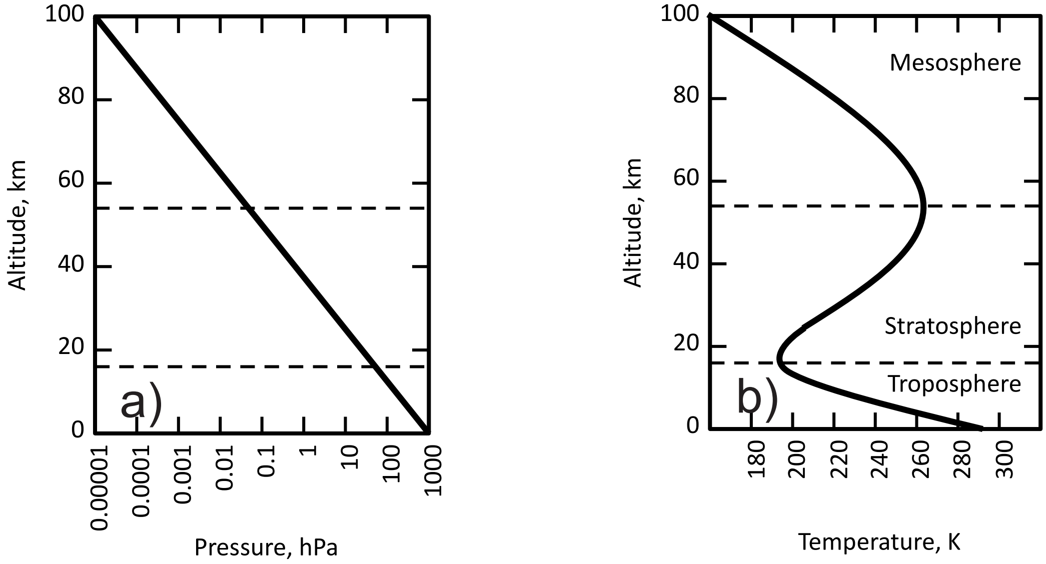
The atmosphere is a layer of gases surrounding the Earth's surface which maintains the planet's ability to sustain life. The atmosphere is made up of a huge number of gaseous species, although most exist only in trace amounts. Only nitrogen (78%), oxygen (21%), argon and water vapour ($<$1% each) contribute a substantial proportion to the composition of the atmosphere. Other gases present in the atmosphere, known as trace gases, may have relatively small atmospheric concentrations, but many of them play a large role in the maintenance of the planet's climate.

Atmospheric pressure is the term given to the force exerted by the weight of the atmosphere, which decreases exponentially with altitude due to the effects of gravity and temperature. Figure 2.1(a) shows the typical vertical profile of atmospheric pressure. The temperature of the atmosphere has a more complex vertical structure, as shown in Figure 2.1(b). The atmosphere is partitioned vertically into domains separated by reversals of the temperature gradient, with each domain having substantially different compositions and physical and chemical properties. The troposphere is loosely defined as the lowest region, where temperature decreases with height. The tropopause, which marks the top boundary of the troposphere, sits at around 8 - 18 km above the Earth's surface, the altitude being highly dependent on latitude and season. The troposphere contains around 85% of the total mass of the atmosphere. Close to the surface is a section of the troposphere known as the planetary boundary layer (PBL), which extends upwards to just a couple of kilometres. The PBL is characterised by turbulent and rapid atmospheric mixing due to the orography and temperature of the surface. The stratosphere sits directly above the tropopause and extends up to an altitude of around 50 km above the surface of the Earth. The temperature of the stratosphere increases with altitude due to the absorption of solar radiation by ozone. Together, the stratosphere and troposphere contain 99.9% of the total mass of the atmosphere, and the majority of the chemical processes necessary for life on Earth occur in these two sections of the atmosphere, although the two domains are chemically and dynamically different. Above the stratopause is the mesosphere, where temperature decreases with altitude again, although not as steeply as in the troposphere, and the thermosphere sits above the mesopause. The temperature of the thermosphere rises sharply up to between 500 K and 2000 K due to the strong absoption of solar UV radiation by N_2 and O_2. This thesis focuses on variations of CH_4 in the troposphere only, but the exchange of air between the stratosphere and troposphere plays an important role in the tropospheric concentration of CH_4, so both domains must be modelled and understood. The effect of the mesosphere and thermosphere on tropospheric CH_4 is minimal.

2.2.1 Atmospheric Transport

Atmospheric transport is the term encompassing all large- and small-scale air movement, which carries trace gas molecules through the atmosphere. This means that regardless of the location of their emission, trace gases are eventually mixed throughout the atmosphere. Advection, convection and boundary layer mixing are three forms of atmospheric transport which are considered in the TOMCAT model, and the nature of each of these three processes is described here.

Figure 2.1 *(Left)* Typical vertical pressure profile and *(Right)* typical vertical temperature profile in the Earth's atmosphere at 30°N, March. Adapted from Jacob (1999).



Advection

Transport of atmospheric air, and therefore of species such as CH_4 contained in the air, due to winds is known as advection, and is responsible for most of the large-scale transport of air around the globe. Wind is the flow of air from areas of high pressure to areas of low pressure, which over a long time period will mix atmospheric constituents throughout the atmosphere. Advection is responsible for the transport of trace gases away from the area where they emitted. For example, since the majority of CH_4 sources are in the northern hemisphere (NH), advection is responsible for transporting atmospheric CH_4 into the southern hemisphere (SH). Advection of an atmospheric tracer at a fixed point satisfies the continuity equation, which states that;

$$\frac{\partial f}{\partial t} = -\underline{\nabla}(f\underline{u}) \tag{2.1}$$

where $f(x_1,x_2,x_3)$ is the tracer mass and $\underline{u} = (u_1,u_2,u_3)$ is a three-dimensional wind velocity field. This means that, assuming no sources or sinks, tracer mass is conserved throughout time.

Convection

Convection is responsible for the rapid vertical transport of air in the atmosphere. There are two types of convection, known as free convection and forced convection. Free convection occurs due to the solar heating of the Earth's surface, which causes a vertical air density gradient known as a thermal, forcing air to rise. Due to the temperature dependence of free convection, its effect is generally greatest in tropical latitudes, and it has less of an effect towards the poles. Forced convection can occur for a number of reasons. It may be due to the orography of the land surface, weather fronts or the convergence of horizontal winds. This effect causes air to rise in mountainous areas, for example.

PBL mixing

The Earth's boundary layer is characterised by turbulent and rapid mixing of air close to the surface due to changing factors such as orography, soil heat capacity and atmospheric heat conductance. Species that are emitted from the Earth's surface quickly become mixed through the boundary layer before they move into the free troposphere at a slower pace. The height of the PBL changes according to heating of the Earth's surface, which affects the rate of venting into the free troposphere. Deposition of some trace species out of the air and into the Earth's biosphere also occurs in the boundary layer, although methane is not a deposited species.

2.3 Atmospheric Chemistry Relating to CH_4

Once emitted into the atmosphere, CH_4 either remains resident in the atmosphere, is destroyed through chemical reactions with other atmospheric species, or is oxidised by the methanotrophic bacteria in soils. The major sink of CH_4 is reaction with the hydroxyl radical, OH, which is responsible for approximately 90% of the removal of CH_4 from the atmosphere, and occurs in both the troposphere and in the stratosphere (Denman et al., 2007). Through a number of further reactions, this process eventually results in the formation of CO and CO_2, giving rise to the indirect radiative forcing potential of CH_4 due to its role in the formation of these other greenhouse gases. The oxidation process of CH_4 in the troposphere to eventually form CO and CO_2 can occur through a number of pathways, which are mainly dependent upon the local presence of nitrogen oxides (NO_x) in the atmosphere. NO_x, which consists of NO and NO_2, is mostly released into the atmosphere through anthropogenic activity such as fossil fuel combustion and biomass burning, although it is also released in relatively small amounts from soils and is produced in the atmosphere due to lightning. NO is oxidised to form NO_2 in a number of tropospheric reactions, such as that with O_3;

$$NO + O_3 \longrightarrow NO_2 + O_2 \tag{2.2}$$

and NO_2 is quickly photolysed back to NO during the daytime;

$$NO_2 + h\nu \xrightarrow{O_2} NO + O_3 \tag{2.3}$$

This means that cycling between NO_2 and NO takes place on a time-scale of minutes in the troposphere, which means that the NO_x budget is generally considered as a whole. In regions with high local concentrations of NO_x, an alternative oxidation pathway is available for CH_4, and both pathways are described here.

When CH_4 reacts with OH, it forms the methyl radical CH_3, which reacts rapidly with O_2;

$$CH_4 + OH \longrightarrow CH_3 + H_2O \qquad [k_{OH} = 2.45 \times 10^{-12} exp(-1775/T)] \tag{2.4}$$

$$CH_3 + O_2 + M \longrightarrow CH_3O_2 + M \tag{2.5}$$

At this point, the presence of NO_x is critical, as the methylperoxy radical CH_3O_2 will react with NO if it is present, or with HO_2 otherwise;

$$CH_3O_2 + NO \longrightarrow CH_3O + NO_2 \tag{2.6}$$

$$CH_3O_2 + HO_2 \longrightarrow CH_3OOH + O_2 \tag{2.7}$$

Methylhydroperoxide (CH_3OOH), which is formed when NO_x levels are low, may react with OH (with two possible reaction branches), or be photolysed;

$$CH_3OOH + OH \longrightarrow CH_2O + OH + H_2O \tag{2.8}$$

$$CH_3OOH + OH \longrightarrow CH_3O_2 + H_2O \tag{2.9}$$

$$CH_3OOH + h\nu \longrightarrow CH_3O + OH \tag{2.10}$$

The methoxy radical (CH_3O), which is formed during reactions 2.6 and 2.10, reacts rapidly with O_2;

$$CH_3O + O_2 \longrightarrow CH_2O + HO_2 \tag{2.11}$$

Formaldehyde (CH_2O) can react with OH or photolyse (with two photolysis branches);

$$CH_2O + OH \longrightarrow CHO + H_2O \tag{2.12}$$
$$CH_2O + h\nu \xrightarrow{O_2} CHO + HO_2 \tag{2.13}$$
$$CH_2O + h\nu \longrightarrow CO + H_2 \tag{2.14}$$

and the CHO radical reacts rapidly with O_2 to form CO;

$$CHO + O_2 \longrightarrow CO + HO_2 \tag{2.15}$$

When CO is formed, it reacts with OH to give CO_2;

$$CO + OH \longrightarrow CO_2 + H \tag{2.16}$$
$$H + O_2 + M \longrightarrow HO_2 + M \tag{2.17}$$

In this way, CH_4 is eventually oxidised to form CO and CO_2, with differing products depending on the presence of NO_x in the atmosphere. In the high-NO_x case, the net reaction is therefore;

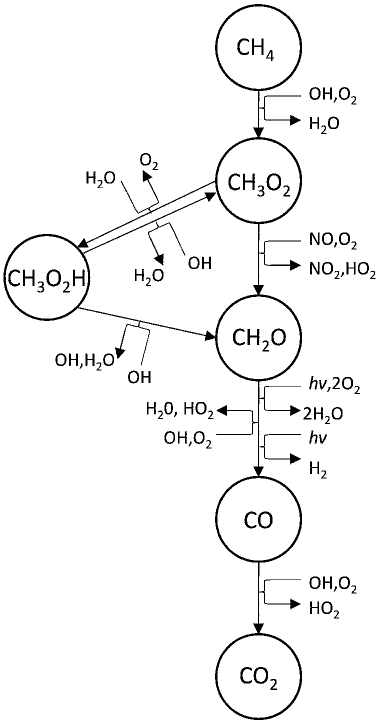
$$CH_4 + 10O_2 \longrightarrow CO_2 + H_2O + 5O_3 + 2OH \tag{2.18}$$

while in case where NO_x is low, the net reaction is;

$$CH_4 + 3OH + 2O_2 \longrightarrow CO_2 + 3H_2O + HO_2 \tag{2.19}$$

The reaction chain is summarised in Figure 2.2. Equations 2.18 and 2.19 indicate the importance of CH_4, along with the presence of NO_x in the atmosphere, in determining the concentration of tropospheric OH. In the presence of NO_x, the oxidation of CH_4 eventually produces two OH molecules and 5 molecules of O_3, which may be photolysed to form OH. Meanwhile, low atmospheric NO_x concentrations lead to the destruction of three OH molecules and the production of a molecule of HO_2. This process is known as HO_x recycling, where the HO_x family is comprised almost entirely of OH and HO_2, and means that the balance of CH_4, CO, CO_2, O_3 and NO_x are responsible for the maintenance of the atmospheric concentration of OH.

Figure 2.2 Chemical processes describing the oxidation of CH_4 in the atmosphere to eventually form CO and CO_2.



CH_4 is also destroyed through reactions with excited oxygen atoms, $O(^1D)$, and with atomic chlorine (Cl). These reactions, detailed in Equations 2.20 and 2.21 respectively, occur only in the stratosphere, and are responsible for approximately 5% of the total loss of CH_4 from the atmosphere.

$$CH_4 + O(^1D) \longrightarrow CH_3 + OH \qquad [k_{O(^1D)} = 1.5 \times 10^{-10}] \tag{2.20}$$

$$CH_4 + Cl \longrightarrow CH_3 + HCl \qquad [k_{Cl} = 7.3 \times 10^{-12} exp(-1280/T)] \tag{2.21}$$

The remaining 5% of atmospheric CH_4 loss is through sequestration into the soils, which is discussed in more detail in Section 2.4.

2.4 Surface CH_4 Fluxes

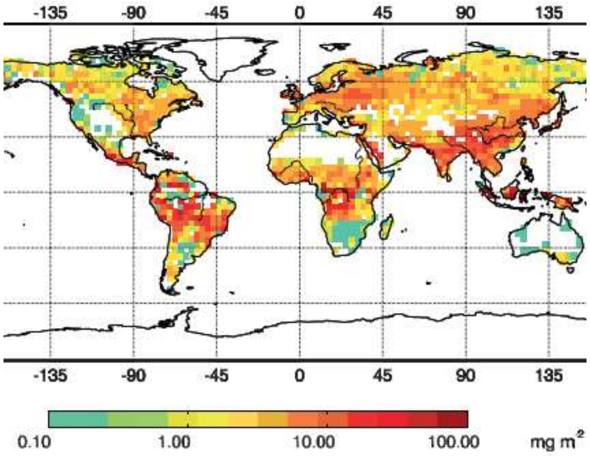
Surface CH_4 emissions have a variety of natural and anthropogenic origins, which stem from both biogenic and non-biogenic processes. Anthropogenic sources include fossil fuel mining and use, landfill sites, biomass burning, rice agriculture and the increased farming of ruminant animals

Table 2.1 Ranges of CH_4 source and sink estimates provided by various contributors to the IPCC 2007 report (Denman et al., 2007), in which the value for total sinks is estimated, and the total magnitude of the sources is determined from the atmospheric imbalance.

Process	Range of Estimates (Tg yr^{-1})
Natural Sources	**145 - 260**
Wetlands	100 - 231
Termites	20 - 29
Ocean	4 - 15
Hydrates	4 - 5
Geological sources	4 - 14
Wild animals	15
Wildfires	2 - 5
Anthropogenic sources	**264 - 428**
Energy	74 - 77
Coal mining	32 - 48
Gas, oil, industry	36 - 68
Landfills and waste	35 - 69
Ruminants	76 - 91
Rice agriculture	39 - 112
Biomass burning	14 - 88
C3 Vegetation	27
C4 Vegetation	9
Sinks	
Soils	26 - 34
Tropospheric OH	445 - 507
Stratospheric Loss	30 - 45
TOTAL SOURCES	**582**
TOTAL SINKS	**581**
IMBALANCE	**+1**

while natural sources include wetlands, termites, the oceans, geological sources, wild animals and wildfires. There are large uncertainties, however, in the nature, size and distribution of these emissions for a variety of reasons which will be discussed later in this section. In total, annual CH_4 emissions are estimated to be approximately 582 Tg yr^{-1} for the years 2000 - 2004 (Denman et al., 2007). Details of the genesis, distribution and size of each of these sources is given here, and a full CH_4 budget, including partitioning of the total source into the various emission processes as estimated in the IPCC report (Denman et al., 2007), is given in Table 2.1. Denman et al. (2007) is henceforth referred to as IPCC07. Note that in the period 2000 - 2004, the total sources and sinks of methane were thought to be almost in balance, giving a net atmospheric increase of just 1 Tg yr^{-1}.

Figure 2.3 Wetland daily emissions of CH_4 (mg m^{-2} d^{-1}) inferred from fitting a temperature-groundwater wetland model to SCIAMACHY CH_4 concentrations averaged over 2003-2005. Taken from Bloom et al. (2010).



Wetlands

Wetlands, such as swamps, peatlands, bogs, dry tundra, and Arctic thermokarst (thaw) lakes are together the largest single source of CH_4, and this flux component has the largest uncertainty. Wetlands are distributed across the globe, with large contributions from the Amazon Basin, swamps in the USA, equatorial Asia and Central Africa, and Arctic tundra and thaw lakes in Canada, Scandinavia and Russia. Figure 2.3 shows the distribution of wetland emissions inferred by Bloom et al. (2010) from a wetland model over the period 2003 - 2005, displaying high wetland sources throughout tropical areas, and also in the high-latitude NH. Methane is emitted from wetlands due to methanogenic bacteria in the soils under anaerobic conditions (Conrad, 1989), and the rate of emission is dependent upon a number of factors including soil temperature, water-table depth, the availability of substrate for bacteria to feed on and the transportation process into the atmosphere. The dependence of wetland CH_4 emissions upon climate-based factors such as soil temperature and water-table depth could lead to a feedback process in which rising temperatures or precipitation levels due to climate change may lead to increased CH_4 flux from wetlands (Gedney et al., 2004; Bohn et al., 2007).

Early estimates placed total CH_4 wetland emissions at around 11 - 57 Tg yr^{-1} (Matthews and Fung, 1987), while IPCC07 included total wetland emission estimates between 100 Tg yr^{-1} (Wuebbles and Hayhoe, 2002) and 231 Tg yr^{-1} (Hein et al., 1997; Mikaloff Fletcher et al., 2004), indicating the extent of the uncertainty surrounding the size of this CH_4 source. Until recently, many estimates of wetland CH_4 emission rates came from small-scale measurements which were extrapolated to regional scales (e.g. Matthews and Fung (1987); Zimov et al. (1997)), or from process-based models which attempted to simulate the processes of wetland methanogenesis (e.g. Cao et al. (1996); Christensen et al. (1996); Walter and Heimann (2000)). However, each of these techniques is limited by our knowledge of the spatial and temporal distribution of wetlands, and of the diversity of CH_4 emission rate between and within wetland types. For example, Walter et al. (2007) found that estimates of CH_4 emission from Arctic lakes often did not take account of small lakes, leading to an underestimation of lake area by a factor of more than 2. Recently, the availability of high resolution satellite measurements have allowed more accurate estimation of wetland size and location. Bloom et al. (2010) used gravity anomaly observations from the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) instrument as a proxy for water table depth, found that surface temperature is the major factor in determining the emission rate of Arctic wetlands, while the water table depth is the determining factor in tropical and SH wetlands. Bousquet et al. (2006) suggested that wetlands were responsible for much of the interannual variability of the atmospheric CH_4 budget throughout the 1990s and early 2000s.

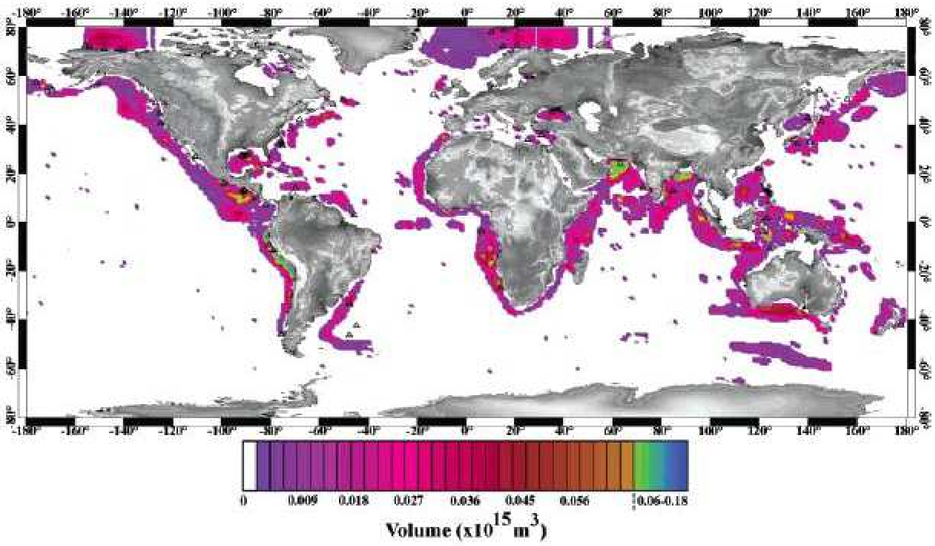
Rice Agriculture

Anthropogenic CH_4 is emitted from rice paddies due to the same methanogenic bacterial processes that are responsible for wetland emissions, and our estimates of the total flux are similarly dependent upon the accuracy of our determination of the spatial and temporal variations in the emission rate. The management system of a particular rice paddy, which may vary greatly from location to location, plays a major role in its methanogenic properties. The majority of CH_4 emissions due to rice agriculture originate in East and South-East Asia. IPCC07 estimated that the annual emission rate due to rice agriculture ranges from 39 (Scheehle et al., 2002) to 112 Tg yr^{-1} (Chen and Prinn, 2005). Li et al. (2002) and Kai et al. (2010) suggested that emissions from rice paddies in China have decreased since 1982 due improved management techniques and a reduction in the area used for rice production. Due to the similarity in the emission process involved, wetland and rice production emissions are often modelled as one (e.g. Bloom et al. (2010)). It is estimated that 70% of emissions from wetlands and rice paddies come from sources in the tropics and SH.

Oceans and Hydrates

Methane is supersaturated in surface ocean waters, where it can outgas into the atmosphere. Some of this excess CH_4, especially around coastal regions, comes from land-based sources such as petroleum production (Sackett and Brooks, 1975) and methane-rich sediments (Bernard et al.,

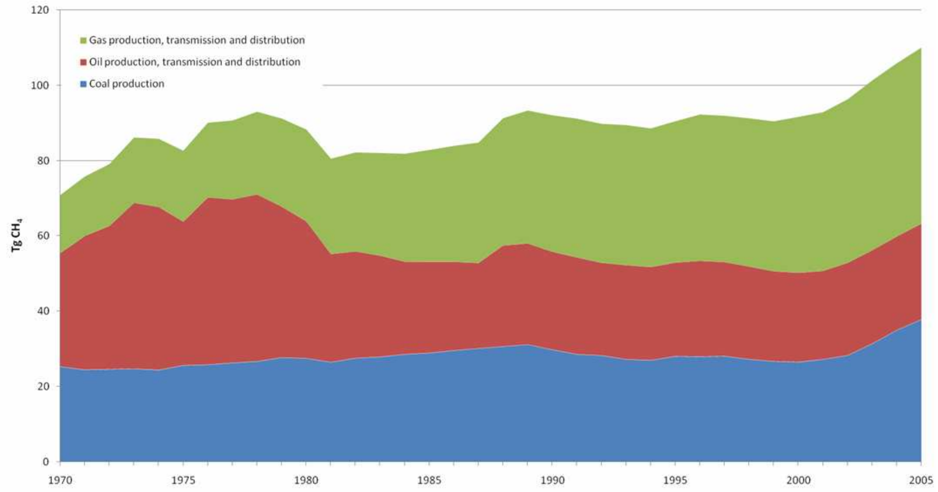
Figure 2.4 Global volume of distribution of methane in hydrate at standard temperature and pressure on a 1° grid on all seafloor locations. Taken from Klauda and Sandler (2009).



1978). However, there must also be an ocean source of CH_4, since transport alone cannot account for supersaturated CH_4 in some parts of the open ocean (Scranton and Brewer, 1977). This source is likely to be due to methanogenic bacteria (Brooks et al., 1981) which reside in the deep ocean, since the surface waters are highly oxygenated, prohibiting methane production (Karl and Tilbrook, 1994).

Marine gas hydrates are structures composed of crystalline water molecules which can trap gaseous molecules such as CH_4 in cage-like configurations, formed under low temperature and high pressure in the planet's oceans. If the temperature rises or the pressure decreases, however, the trapped molecule may be released and emitted into the ocean, and further into the atmosphere. CH_4 hydrates are the most common from of hydrate (Archer, 2007), but knowledge of the extent to which they occur globally is 'very incomplete' (O'Connor et al., 2010). Figure 2.4 shows the global volume distribution of methane in hydrate on the sea floor. CH_4 hydrate concentration is greatest in coastal regions in the tropics and in the NH at high latitudes. Like wetlands, methane hydrates may be subject to a feedback loop due to the temperature dependence of the release of the trapped gas. Archer (2007) carried out a model study of hydrate emissions of CH_4 and estimated that there are 1600 - 2000 Pg of carbon in the ocean in the form of CH_4 hydrates, and, citing a study by Dickens (2003), considered that the hydrate source may be a 'slow but ... irreversible tipping point in the Earth's carbon cycle'. CH_4 emission rates from the oceans are small in comparison to some other sources. IPCC07 estimated that ocean sources ranged between 4 and 15 Tg yr^{-1}, while the emission rate due to hydrates was estimated to be approximately 5 Tg yr^{-1}.

Figure 2.5 Anthropogenic fossil fuel emissions of CH_4 (Tg year^{-1}) for the period 1970-2005. Taken from edgar.jrc.ec.europa.eu/part_CH4.php.



Fossil Fuels

CH_4 is emitted during the extraction processes and burning associated with fossil fuels. Since industrial processes such as these are generally well documented, the magnitude and distribution of our estimates are assumed to be fairly well constrained. IPCC07 placed total anthropogenic emissions due to coal mining, gas, oil and industry at 82 - 106 Tg yr^{-1}. Figure 2.5 shows the total emissions of CH_4 from fossil fuels as estimated by the EDGAR version 4.0 inventory for the period 1970 - 2005. The EDGAR inventory estimated that coal and oil emissions showed no overall trend over the period 1970 - 2000, while CH_4 emissions due to the natural gas industry rose steadily throughout that period. There was relatively small interannual variation associated with each of the three industries between 1980 and 2005. However, Cofala et al. (2007) estimated that, considering the legislation at the time, anthropogenic CH_4 emissions due to industry would increase by up to 35% by 2010, mostly due to increase emissions from Africa and the Middle East and from Centrally Planned Asia (e.g. Vietnam, North Korea and Mongolia) and China. EDGAR estimated that coal and gas emissions had increased by approximately 20 Tg yr^{-1} between 2000 and 2005. Aydin et al. (2011), however, estimated that fossil fuel emissions had been decreasing since the 1980s and were still doing so in 2010.

Landfills

CH_4 is emitted from anthropogenic soil-covered landfill sites due to biodegradation in anaerobic conditions of waste. The size of this source was estimated by Bingemer and Crutzen (1987) to be

30 - 70 Tg yr^{-1}, and this estimate has changed little, with estimates in IPCC07 ranging from 35 - 69 Tg yr^{-1}. As a biogenic source, this reaction is dependent on soil temperature, indicating a possible future feedback process.

Ruminants and Termites

Methane is a by-product of the microbial breakdown of carbohydrates such as cellulose in the digestive systems of most herbivores, with ruminants producing the highest ratio of CH_4 production (Blaxter and Czerkaws, 1966; Wolin, 1981). This source of CH_4 appears to be fairly well constrained thanks to agricultural data provided by each country and published methane yield estimates for different species, and IPCC07 estimates that the annual CH_4 emission due to ruminants ranges between 76 (Scheehle et al., 2002) and 91 Tg yr^{-1} (Mikaloff Fletcher et al., 2004). Russia, Brazil, Western Europe, Africa, India and the US produce the highest CH_4 emissions due to cattle agriculture (Johnson and Ward, 1996), and the distribution due to all other ruminants is similar. Animal manure also produces small methane emissions, especially if stored (Bogner et al., 1995).

CH_4 is emitted from the digestive systems of termites. This source is difficult to model and estimate, due to the differences in emission rates from various termite families (Sugimoto et al., 1998) and the uncertainties in the location of major termite colonies. In fact, the emission rate from workers may vary from colony to colony even within the same species (Wheeler et al., 1996). IPCC07 estimates that the global methane source due to termites sits between 20 (Wuebbles and Hayhoe, 2002; Houweling et al., 1999) and 29 Tg yr^{-1} (Mikaloff Fletcher et al., 2004).

Biomass Burning and Wildfires

Anthropogenic biomass burning and natural wildfires are a source of a number of atmospheric trace gases, including methane. Although burning in oxygen-rich environments releases mostly CO_2, the smouldering phase of fires releases CH_4, CO and other hydrocarbons (Lobert et al., 1991). Depending on the fuel burnt, 2.3 - 10.7 g CH_4 are emitted per kilogram of dry matter burned (Andreae and Merlet, 2001), and IPCC07 placed estimates of biomass burning CH_4 emissions between 14 (Scheehle et al., 2002) and 88 Tg yr^{-1} (Mikaloff Fletcher et al., 2004), while wildfires produced 2 - 5 Tg yr^{-1}. Around 85% of CH_4 released from biomass burning is from tropical regions, with small contributions from temperate and boreal regions (Hao and Ward, 1993). Due to the difficulties in estimating burned areas and vegetation types, and the ever-changing distribution of emissions, there are large uncertainties in our estimates of biomass burning and wildfire CH_4 emissions, and recent studies such as Duncan et al. (2003) and Edwards et al. (2006) attempt to constrain trace gas emissions using satellite observations of fires. Duncan et al. (2003), who studied the distribution of CO emissions due to biomass burning from 1979 to 2003, found that although there was no significant trend in biomass burning emissions over that time period, there was large interannual variation on a regional and global scale. van der Werf et al. (2006) confirmed this over the period 1997 - 2004. As indicated by the range of estimates provided in IPCC07, biomass burning is one of the less constrained sources of CH_4, and further study is necessary in order to quantify its influence.

Other sources

There are other, possibly large, natural sources of CH_4 which are not well understood. IPCC07 estimated that geological sources of methane contribute between 4 and 14 Tg yr^{-1}. CH_4 is formed in the Earth's crust, mainly by bacteria and thermogenic processes (Denman et al., 2007), and released into the atmosphere through faults, fractured rocks and mud volcanoes (Milkov et al., 2003; Kvenvolden and Rogers, 2005), and the total geological source of CH_4 may be as high as 40 - 60 Tg yr^{-1} (Denman et al., 2007). Kvenvolden and Rogers (2005) estimated the geological source to be approximately 45 Tg yr^{-1}, which is accounted for as part of the fossil fuel budget. Many of these geological sources have been subject to detailed measurements only in recent years. For example, measurements from mud volcanoes began in 2001 (Etiope and Milkov, 2004), and the number of mud volcanoes may range between 100 and 10,000 (Milkov et al., 2003). Geologic sources require far greater investigation before we can begin to understand the extent of their influence upon the atmospheric methane budget.

Another natural source which is poorly understood is the natural flux from some plants, which was suggested only recently by Keppler et al. (2006), who found that living C3 and C4 category plants, such as broadleaf trees, sugarcane and maize, release methane in situ. This study estimated that the global source from living plants is 62 – 236 Tg yr^{-1}, mostly from the tropics, although the top limit of this estimate was revised down to 85 - 125 Tg yr^{-1} by Houweling et al. (2006) using a transport model. Total CH_4 emissions from plants was most recently estimated at 20 - 69 Tg yr^{-1}, with a best estimate of 36 Tg yr^{-1}, by Butenhoff and Khalil (2007), who suggested that around 24% of these emissions may already be accounted for in the wetland emission budget. However, this reasoning still meant that approximately 28 Tg yr^{-1} CH_4 were unaccounted for in the methane budget. IPCC07 allowed for 27 Tg yr^{-1} and 9 Tg yr^{-1} for C3 and C4 category plants respectively, in agreement with Butenhoff and Khalil (2007).

Soil Sink

The oxidation of CH_4 by aerobic methanotrophic bacteria in the soils occurs in most types of oxic soils. Most estimates of the soil sink from process models (Ridgwell et al., 1999) and top-down studies (Bousquet et al., 2006) indicate that it accounts for approximately 30 Tg yr^{-1} CH_4, or about 5% of the total methane sink.

2.4.1 Isotopic Composition of CH_4 Emissions

^{13}C is an isotope of carbon containing an extra neutron, and atmospheric methane containing ^{13}C can be monitored independently from the vast majority of CH_4, which contains ^{12}C. This provides a useful diagnostic for locating the origin of atmospheric CH_4, since different methane sources produce different ratios of ^{13}C to ^{12}C. Finding this ratio for an air sample containing methane can therefore constrain the nature of its emission. The value δ^{13}C for an air sample provides a measure of the ^{13}C to ^{12}C ratio, and is defined as;

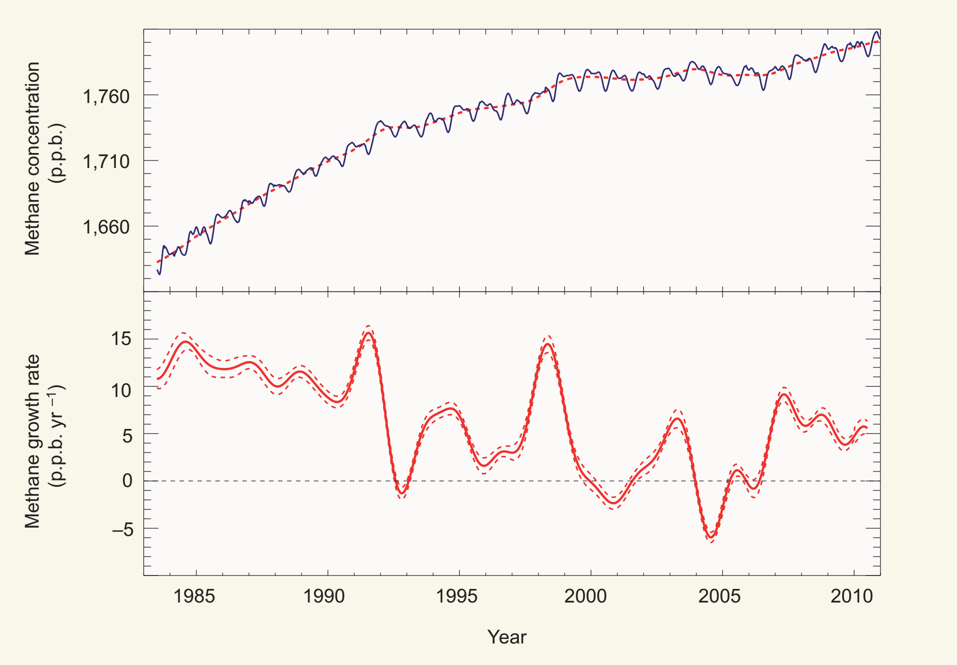
$$\delta^{13}C = \left(\frac{\left(\frac{^{13}C}{^{12}C}\right)_{sample}}{\left(\frac{^{13}C}{^{12}C}\right)_{standard}} - 1 \right) \times 1000‰ \tag{2.22}$$

While ambient background air typically has δ^{13}C≈-47‰, CH_4 formed by combustion is generally enriched in ^{13}C while CH_4 from biogenic sources is depleted. δ^{13}C can also change according to the source location. For example, δ^{13}C for methane emitted from wetlands at high northern latitudes varies between -70‰ and -60‰, while tropical wetlands produce CH_4 with δ^{13}C varying from -60‰ to -50‰ (Dlugokencky et al., 2011). Although monitoring of δ^{13}C for methane is currently limited to not much more than a very small number of locations in the Arctic, a more extensive measurement network would be a powerful tool in future investigation into the methane emission budget. For example, Mikaloff Fletcher et al. (2004) used isotopic measurements as further constraints in an inverse model study for CH_4 emissions in 1998.

2.5 Recent Variations in the Atmospheric CH_4 Budget

Since the 1980s, the rate of increase of the atmospheric CH_4 burden has been slowing down dramatically. Figure 2.6 shows the global mean CH_4 mixing ratio for the period 1983 - 2006 and the instantaneous CH_4 growth rate. The growth of CH_4 since the 1980s has four distinct phases, with some interannual variability within these phases. The first phase is from the start of the observation period until approximately 1992. During this period, the growth rate of CH_4 is reasonably high - approximately 10 ppb yr^{-1} - but decreasing. After a sharp rise in 1991 followed by a significant drop in 1992, which will be discussed later in this section, the second phase begins, in which the rate of increase remains small until 1998. A large increase in 1998 preceded the third phase, in which the net increase was approximately zero (≈0.4 ppb yr^{-1}) between 1999 and 2007. Finally

Figure 2.6 *(Top)* Global mean tropospheric CH_4 (ppb) derived from surface sites operated by NOAA/GMS for the period 1983-2011. (Bottom) Annual growth rate (ppb yr^{-1}) in global atmospheric abundance. Taken from Heimann (2011).



the fourth phase, which covers the period between 2007 and the present day, in which concentrations have again started to increase. The global mean CH_4 increase was 8.3 ± 0.6 ppb in 2007 and 4.4 ± 0.6 ppb in 2008 (Dlugokencky et al., 2009), a significant increase on that of the preceding years, and the atmospheric burden has continued to increase up until the present day (Heimann, 2011).

The atmospheric CH_4 burden depends on a delicate balance between the sources and sinks, and the observed deceleration of the atmospheric CH_4 budget increase has implications for our understanding of CH_4 emission and destruction. Before the recent renewed increases in the budget, there was fierce debate as to whether the period of stability was a temporary pause in the increase or a new steady state for atmospheric CH_4 (Dlugokencky et al., 1998; Simpson et al., 2002; Wuebbles and Hayhoe, 2002; Dlugokencky et al., 2003). However, the recent return to rising global methane levels is a possible indication that the levelling-off was only a temporary respite. A number of attempts have mean made to link the variations in the global methane budget to changes in the sources and sinks of the species, and a brief summary of some of the hypotheses is given here.

The sudden drop in the increase rate in 1992 came the year after two major global events – the collapse of the former Soviet Union (fSU) and the eruption of Mount Pinatubo, a volcano in South Asia. It has been argued (e.g. Dlugokencky et al. (1994b)) that these two events were responsible for decreased CH_4 emissions in 1992, which led to a small decrease in the atmospheric CH_4 budget that year ($\approx$ -1.5 ppb). During the peak years of the fSU, government calls for large-scale, rapid extraction of gas from the West Siberian basin meant that long gas pipelines were quickly built with little regard for reducing leakage, emitting 29 to 50 Tg yr^{-1} in ‘lost’ CH_4 (Reshetnikov et al., 2000). The fall of the fSU in 1991 meant a stabilisation in the amount of gas extraction and a fall in the amount of coal mining in the area, while repairs had been carried out on the pipelines since 1989. These factors may have contributed to a decrease of up to 37 Tg yr^{-1} (Dlugokencky et al., 1994a), although this estimate was later revised down to 10 Tg yr^{-1} (Dlugokencky et al., 2003). The eruption of Mount Pinatubo on June 15, 1991, may also have had a suppressing effect on CH_4 emissions. The eruption, which was the second largest of the 20th century, reduced the surface air temperature in the NH by up to 0.7 °C (Dutton and Christy, 1992). Since the biogenic emission of CH_4 from sources such as the northern wetlands is temperature dependent, this may have further decreased emissions in 1992 by around 2 Tg (Dlugokencky et al., 1994a). It was also found that the eruption may have been responsible for decreased tropospheric OH in the short term due to increased sulfur and aerosol emissions, which may explain the sharp increase in the 1991 growth rate of atmospheric CH_4 before the decline in 1992 (Dlugokencky et al., 1996).

After the reduction of fossil fuel emissions produced by the fall of the fSU, global CH_4 concentrations continued to rise slowly until 1998, when there was an anomalously high atmospheric growth rate of approximately 12.7 ppb yr^{-1}, compared to an average of 3.9 ppb yr^{-1} in 1995 - 1997 (Dlugokencky et al., 2001). This was due to the strong El Niño meteorological event in 1998 which made that year the warmest on record (Hansen et al., 1999) and produced increased precipitation in the high northern latitudes (north of 30°). Dlugokencky et al. (2001) found that these conditions increased CH_4 emissions from wetlands and biomass burning in the NH by up to 31 Tg that year. Bousquet et al. (2006) suggested that decreased OH concentrations due to reaction with increased CO emissions from biomass burning may have also produced increased CH_4 concentrations.

Aside from the anomalous CH_4 growth rates in 1991-92 and 1998, the period 1983 - 2000 was characterised by a gradual decrease in the rate of growth for atmospheric CH_4. During the third phase, which lasted from 2000 until 2007, the net year-on-year increase of the atmospheric CH_4 budget was approximately zero, meaning that the sources and sinks of methane must have been in balance. An exception is in 2002-03, when increased biomass burning due to an El Niño event produced a year of growth (Simpson et al., 2006). Whether the stability meant that emissions and sinks had both stabilised at a constant value, or that the two were varying in unison, was questioned (e.g. Dlugokencky et al. (2003)), and the key issue in discerning the cause of the stability of the global concentration of CH_4 is whether there is significant variation in the concentration of OH. The OH radical is difficult is measure in situ, and its concentration must often be found using a proxy such as methyl chloroform (MCF). Constant OH concentrations throughout the period imply that the global methane source was also near constant, although the partitioning and distribution of the source may of course have varied.

Bousquet et al. (2006), investigating the interannual variability of atmospheric CH_4 using an inversion model, confirmed that the slowdown of the growth rate through the 1990s was due to decreasing fossil fuel emissions. The study found that as emisisons due to fossil fuels began to rise again after 1999, especially in North Asia, the effect was masked by decreasing wetland emissions until 2003, keeping the system in balance. This study also suggested that OH had been slightly decreasing throughout the period. The increase in North Asian emissions is likely due to the booming Chinese economy, and agree with the estimates in the EDGAR inventory. The decrease in wetland emissions between 1999 and 2003 was likely to be due to dryer conditions in the NH and tropics. However, Bloom et al. (2010) estimated that wetland emissions increased between 2003 and 2007 by 7%, mostly in the NH mid-latitudes (45°N - 67°N). Since the global atmospheric CH_4 budget did not increase during this period, the increase in wetland emissions must have been mitigated by some other factor. Fiore et al. (2006) and Monteil et al. (2011) both suggested the OH concentrations were rising moderately throughout the previous two decades (by a level of 1.2% and less than 5% per decade respectively), and this finding agreed with that of Dentener et al. (2003), who claimed that OH had a positive trend of 2.4% per decade for the period 1979 - 1993. Prinn et al. (2005) suggested the OH concentrations were rising between 1998 and 2003, but the study period ended before 2004. If OH concentrations were rising in this period, it may help to account for the stability during the period in which wetland emissions were reportedly rising. Montzka et al. (2011b) agreed that the interannual variability of the OH radical in the atmosphere was small, but claimed that OH concentrations were in fact decreasing in the period 2004 - 2007. If, as is claimed by Olivier (2002) and Bousquet et al. (2006), fossil fuel emissions were also increasing over this period, it is difficult to reconcile the stable atmospheric budget with increased wetland and fossil fuel emissions and a decrease in the major CH_4 sink.

In the final phase, which runs from 2007 until the present day, the atmospheric budget is again rising. As an aside, a study by Khalil et al. (2007) hypothesised that large increases in the atmospheric CH_4 budget (up to 8 ppb) occurred every 7.7 years – in 1984-85, 1991-92 and 1999-2000 – but that these three examples did not form a large enough sample from which to draw the conclusion of periodicity. It is interesting that the next increase of a comparable size came in 2007-08, continuing the cycle. Rigby et al. (2008) examined the most recent increase in terms of two scenarios; the first in which the OH concentration did not change in 2007 and 2008, and the second in which OH variability was included. With constant OH, it was found that a substantial increase in emissions in both the NH and the SH between 2006 and 2007 was necessary to explain the increase. If a small decrease in the OH radical occurred, however, only an NH increase was necessary. It was reasoned that this was the more reasonable scenario. Dlugokencky et al. (2009), however, suggested that anomalously high temperatures in the Arctic and increased tropical precipitation led to increased wetland and biomass burning emissions in this period, and that a change in the OH radical was not necessary to explain the increase. That study, like Dlugokencky et al.

(2011), warned that although the most recent increases did not appear to be influenced by the large CH_4 store in Arctic hydrates, that this may become a possibility in the future.

2.6 Summary

Although our understanding of the overall atmospheric budget and growth rate of CH_4 is well constrained, understanding of the specific contributions of different source and sink processes is limited, and this impedes our ability to predict the effects of future changes in climate and anthropogenic activity on the global CH_4 concentration. Major uncertainties surround a number of the parameters of CH_4 emission, especially those from natural sources such as wetlands, rice paddies and biomass burning. It has been suggested that wetland emissions are responsible for much of the interannual variability of the atmospheric CH_4 budget (Bousquet et al., 2006), but our knowledge even of the size and distribution of global wetlands is limited, and emission rates may vary greatly depending on the temperature, water table depth and wetland type. Likewise, biomass burning emissions vary greatly in time and location, and are generally poorly constrained. Even anthropogenic sources, such as fossil fuel emissions, have uncertainty surrounding their partitioning. The main questions surrounding the methane budget are therefore;

- How is the decrease, pause and subsequent rise of the atmospheric CH_4 growth rate since the 1980s related to variations in the emission and destruction rates of the species?
- What is responsible for the large interannual variability in the growth rate of CH_4?
- What are the location and size of natural, geographically varying CH_4 sources such as wetlands and fires?

Bottom-up inventories are therefore limited by our understanding of the underlying processes. However, top-down emission inventories are not restricted by our estimation of emission processes, and inverse modelling can provide emission inventory which is constrained by observations. Inverse methods can provide an accurate geographical distribution of emissions, which together with bottom-up inventories can increase our understanding of variability in CH_4 emission rates. The next chapter describes some of the inverse methods available for source estimation, and describes the findings of previous top-down CH_4 source studies.

Chapter 3

Inverse Modelling Techniques and TOMCAT Model Description

3.1 Introduction

This chapter provides a background to the field of inverse transport modelling, which comprises a variety of different techniques with a common goal; to find an optimised constraint on a parameter, such as the surface emission of an atmospheric species, through the assimilation of observations into a model which maps between the parameter and the observed data. This chapter also describes the TOMCAT model, which is further developed and used throughout this work. Section 3.2 describes the rationale behind inverse transport modelling, and briefly discusses the various techniques available, examining the advantages and disadvantages of each. Section 3.3 discusses the findings of previous inverse modelling studies. Section 3.4 describes the TOMCAT model, theth atmospheric transport model which is the basis for the inversion system developed in this thesis.

3.2 Inverse Modelling Techniques

The term ‘inverse modelling’ actually encompasses a number of techniques, with each technique providing a different route to the same goal of assimilating observations into a model in order to constrain a parameter controlling the model representation of the observations. Three techniques which have been used in published studies to constrain surface emissions of atmospheric gases are the synthesis inversion, the Kalman Filter inversion and the 4D-Var inversion. Each of these methods is discussed in turn here. A discussion of the advantages and disadvantages of each methodology is provided in Section 3.2.5.

3.2.1 General Inversion Methodology and Notation

This section briefly outlines the generalised methodology of atmospheric inversions and describes the notation used to describe each specific inversion process in the following sections. An inversion process assimilates a set of m observations, $\mathbf{y}$, of the atmospheric concentration of a trace gas, and has the aim of optimising the state vector $\mathbf{x}$, which in this case is the surface flux of the species and has dimension n. This is generally done by making changes to an initial estimate of the fluxes, known as the *a priori* estimate. The a priori is generally the previous best estimate of the surface fluxes, and is divided spatially and temporally depending on the methodology, the required accuracy and the available computational power. The a priori has an associated $n \times n$ error matrix $\mathbf{B}$, which contains the uncertainty associated with the fluxes along the diagonal, and the correlation between these errors elsewhere. An analogous $m \times m$ error matrix $\mathbf{R}$ contains the uncertainty associated which each observation, and also the correlation between them. Regardless of the inversion methodology, the aim is to minimise the difference between the observations $\mathbf{y}$ and the representation of the observations produced using a CTM by altering the fluxes used in the model. Model errors may also be included in the observation error matrix $\mathbf{R}$. The optimised state vector is known as the *a posteriori*, labelled $\mathbf{x}_p$. In each inversion, the results from model simulations are compared with the observations, $\mathbf{y}$. In this case the atmospheric model is described as $T(\mathbf{x})$, and an observation operator H maps the 3D output from the model onto the location and time of the observations, so that $H(T(\mathbf{x}))$ has the same dimension as $\mathbf{y}$. The following sections will provide an outline of three specific inversion types.

3.2.2 Synthesis Inversion Technique

The synthesis inversion is derived from Bayesian theory, and was described by Tarantola and Valette (1982a,b). The flux field is decomposed into a number, S, of individual components, $x_j,\ j = 1, ..., S$, such as source regions or emission processes, where the source strength of each component is normalised to a unit strength. These are known as basis functions (Kaminski et al., 2001). Individual model simulations are then carried out for each basis function, in which the spatial and temporal distributions within each are prescribed. The atmospheric mixing ratio of the trace gas produced by the model can therefore be described as a linear combination of the concentrations produced using the basis functions, i.e.;

$$H(T(\mathbf{x})) = \sum_{j=1}^{S} \alpha_j H(T(\mathbf{x}_j)) \tag{3.1}$$

where α_i is the individual regional/process source strength. If $H(T(\mathbf{x}))$ is linear, it is now possible to write;

$$H(T(\mathbf{x})) = \mathbf{G}\mathbf{x} \tag{3.2}$$

where $\mathbf{G}$ is an $m \times S$ matrix and G_{ij} is the contribution of source j to observation i. The state vector is estimated by minimising the cost function J;

$$J(x) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_a)\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_a) + \frac{1}{2}[\mathbf{y}-H(T(\mathbf{x}))]\mathbf{R}^{-1}[\mathbf{y}-H(T(\mathbf{x}))] \tag{3.3}$$

The cost function measures the difference between the model output produced using the current flux estimate and the observations, summed with the difference between the current flux estimate and the a priori, each weighted by its associated error matrix. From Tarantola and Valette (1982a), using Equations 3.2 and 3.3, we therefore have;

$$\mathbf{x}_p = \mathbf{x}_a + \{\mathbf{G}^T\mathbf{R}^{-1}\mathbf{G} + \mathbf{B}^{-1}\}^{-1}\mathbf{G}^T\mathbf{R}^{-1}(\mathbf{y}-\mathbf{G}\mathbf{x}_a) \tag{3.4}$$

which can be solved through Singular Value Decomposition (SVD) as described in Tarantola and Valette (1982a), or by Cholesky decomposition. This technique has been used for inversions of CO by Bergamaschi et al. (2000), for CO_2 by Bousquet et al. (1999) and for CH_4 by Fung et al. (1991), Hein et al. (1997) and Bousquet et al. (2006).

3.2.3 Kalman Filter Technique

The Kalman Filter was originally described by Kalman (1960), and can be extended to a wide range of data assimilation problems in oceanography, meteorology and inverse modelling. The Kalman Filter inversion method assumes that if $\mathbf{H}$ is an observation operator which maps the fluxes $\mathbf{x}$ onto the observations $\mathbf{y}$, usually consisting of a CTM simulation and some form of interpolation or averaging onto the observation field, then the Kalman Filter method attempts to minimise the cost function of the problem J as described in Equation 3.3. Kalman (1960) showed that the state vector $\mathbf{x}_p$ which minimises J is;

$$\mathbf{x}_p = \mathbf{x}_a + \mathbf{K}[\mathbf{y}-\mathbf{H}(\mathbf{x}_a)] \tag{3.5}$$

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T[\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}]^{-1} \tag{3.6}$$

where $\mathbf{K}$ is known as the Kalman gain matrix and $\mathbf{H}$ is the Jacobian matrix of the observation operator. Finding the values in $\mathbf{K}$ creates difficulties for two reasons. The first is that the Jacobian matrix **H** is difficult to calculate explicitly, and the second is that if the matrix $[\mathbf{HBH}^T + \mathbf{R}]$ has large dimension, its inversion becomes computationally prohibitive.

One way to circumvent these problems is to use a method similar to that used in the synthesis inversion. The size of the state vector is reduced by aggregating emissions into regions, and the Jacobian matrix **H** may be explicitly produced by running a number of individual sensitivity simulations, equal to the number of basis functions, over the experiment time period. This technique is useful only for a relatively small number of observations and basis functions, otherwise the number of sensitivity simulations and the size of the inversion becomes too great. This technique was used to examine CO_2 fluxes by Bruhwiler et al. (2005) and in a study of CH_4 emissions by Chen and Prinn (2005).

The methodology which is commonly used for large-scale atmospheric inversion problems is known as the Ensemble Kalman filter (En-KF), and produces optimised surface fluxes by running a number, P, of perturbed forward simulations (known as the ensemble) in unison and updating them to better represent observations at the end of each assimilation time step. This reduces the burden of the Kalman Filter technique firstly as it allows the use of the forward model to find **H** implicitly rather than explicitly, and secondly due to the fact that the covariance matrix **R** is represented by the spread of the ensemble, but the number of simulations P may be much smaller than the dimension of **R** while still being representative. This allows the assimilation of large amounts of data, which is not possible without the ensemble. This method still requires the aggregation of the surface emissions, however. The a priori error matrix can be approximated using an ensemble of perturbation states $\Delta X = [\Delta x_1, \Delta x_2, ..., \Delta x_P]$ which are representative of the background error matrix, i.e.;

$$\mathbf{B} = \Delta X (\Delta X)^T \tag{3.7}$$

This means that the Kalman gain matrix **K** in Equation 3.6 can be approximated by an ensemble gain matrix $\mathbf{K}_e$;

$$\begin{aligned} \mathbf{K}_e &= \Delta X (\Delta Y)^T [\Delta Y (\Delta Y)^T + \mathbf{R}]^{-1} && (3.8) \\ \Delta Y &= H(\mathbf{x}_a + \Delta X) - H(\mathbf{x}_a) && (3.9) \end{aligned}$$

This approach means that the Jacobian matrix **H** does not need to be explicitly calculated, and also has the advantage that the error covariance of the updated state vector is updated by the assimilation. The En-KF technique was used in a study by Peters et al. (2005) in a CO_2 flux study which assimilated surface station flask measurements, and by Feng et al. (2009) to estimate surface CO_2 fluxes using synthetic satellite measurements, a problem with high dimension.

3.2.4 4D-Var Inversion Technique

The 4D-Var technique, which has been chosen as the inversion method for use with the TOMCAT model in this thesis, is described in full in Section 5.2, but a brief description is given here for comparison with other methodologies. The 4D-Var method iteratively minimises the cost function given in Equation 3.3, improving the estimate of the state vector $\mathbf{x}$ with each iteration. The cost function is minimal when its gradient, $\nabla J(\mathbf{x})$, is equal to zero, and it is at this point that the state vector is optimal with respect to matching both the observations and the a priori estimate. This method requires the creation of the adjoint model H^T, which is described in full in Section 5.2. The development of a linear version of the transport model may also be necessary if the model is not already linear, since the adjoint model is defined as the transpose of the Jacobian matrix of a linear model. Also, the adjoint model requires trajectories produced in the forward model, which generally means that large amounts of data must be saved when running the forward model. The 4D-Var method has been used previously to for CO_2 fluxes by Chevallier et al. (2005), and for CH_4 fluxes by Meirink et al. (2008a), Bergamaschi et al. (2009) and Bousquet et al. (2011).

3.2.5 Evaluation of Inversion Methods

The three methods described here each have advantages and disadvantages which make them suited to different purposes. This section evaluates the strengths and weaknesses of each method, and justifies the use of the 4D-Var method for the work carried out in this study.

The key component to the synthesis inversion method is that the matrix $[\mathbf{G}^T\mathbf{R}^{-1}\mathbf{G}+\mathbf{B}^{-1}]^{-1}$ shown in Equation 3.4 must be invertible, which means that both the number of observations and the number of basis functions must be relatively small, and this fact brings both advantages and disadvantages. The main advantage is a reduction in the computational burden. While this technique does require a number of forward simulations equal to the number of basis functions, the extra computational burden that this brings is more than offset by the fact this method does not require multiple iterations of the simulation and that there is no large memory requirement for saving data, both of which are necessary for the 4D-Var method. However, this necessarily large size for the source regions brings aggregation errors, which are explored in Kaminski et al. (2001) and Engelen et al. (2002). The source distribution within each basis function must be prescribed, which allows for no spatial or temporal variation within the aggregated source region. This means that observations have an artificially large region of influence. Despite this, the synthesis inversion method has the advantage that it allows for simple inversions in terms of individual source processes. The synthesis inversion technique can provide a good first estimate of emission rates from different source processes on a regional basis, but does not generally allow enough flexibility to produce truly accurate solutions.

Evaluating the advantages and disadvantages of the Kalman Filter method against the 4D-Var method is more complicated, and generally comes down to a matter of practicality. The major advantage of the Kalman Filter method is in its relative simplicity. If the En-KF is used, then the method is capable of assimilating large amounts of data, such as that provided by satellites, and there is no need for the development of any additional model since only the standard forward model is used. While use of the En-KF method may necessitate a large number of ensemble simulations, the extra computational burden that this requires is more than offset by the fact that multiple iterations of the simulations are not required, unlike in 4D-Var. The main problem of the En-KF method is due to the necessary aggregation of the state vector into large flux regions, providing little information of the distribution within these regions. However, if there exist large areas which have only a small influence on the atmospheric distribution of the species in question, such as the oceans for methane, these areas may be treated as single large emission regions so that areas with large spatial differences may be treated separately. In Feng et al. (2009), for example, there were 99 land emission regions and only 44 oceanic regions since the ocean variability was much smaller. The Kalman filter method also produces results that are strongly dependent on the a priori state vector, which is generally only slightly changed by the Kalman gain matrix.

It is true that the 4D-Var method requires both a much higher degree of initial work and a larger computational burden than the other methods described here. The initial development of the linear and adjoint versions of the transport model may take time and resources, and the fact that, for longer simulations, a large number of forward model trajectories must be either saved as model parameters or written to files which are then read back into the adjoint model, often leads to a very large computer memory requirement. However, if these issues can be overcome, the 4D-Var method can easily assimilate large amounts of data and deal with a large state vector without the requirement for large inversions. Temporal variations in the state vector are automatically accounted for by the adjoint model, which can easily be adapted for different atmospheric species once it has been created. It is for these reasons of accuracy, high data content and versatility that the 4D-Var inversion method was chosen for use in this thesis.

3.3 Previous CH_4 4D-Var Inverse Modelling Studies

There have previously been only limited inverse modelling studies with the aim of improving CH_4 inventories. Hein et al. (1997) carried out a synthesis inversion for CH_4 emissions for the year 1987 and Bergamaschi et al. (2000), Bergamaschi et al. (2005) and Bergamaschi et al. (2007) carried out multiple CH_4 synthesis inversions for time periods throughout the late 1990s and early 2000s Bousquet et al. (2006) also used the synthesis inversion technique to look at CH_4 emissions over the previous two decades. Chen and Prinn (2005) studied CH_4 emissions for the period 1996-2001 using the Kalman Filter.

Previous studies of CH_4 emissions which use the 4D-Var inversion method are extremely limited. Meirink et al. (2006) first used the 4D-Var system for the TM4 CTM to constrain CH_4 emissions using synthetic satellite data, concluding that the observation accuracy of 1 - 2% provided by the satellite CH_4 observations was high enough to reduce uncertainty in monthly mean subcontinental source strengths. The same system, by this point being used in the TM5 CTM, was used by Meirink et al. (2008a) in order to assimilate real satellite data from the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) in order to constrain CH_4 emissions for the period September to November 2003. While this is a relatively short period, the zoom capability of the TM5 model allowed for detailed $1° \times 1°$ inversions of South American emissions. This study found that the 4D-Var inversions produced large uncertainty reductions in tropical and sub-tropical regions, and that the results produced by assimilating the satellite data were consistent with independent global surface station observations and Brazilian airborne observations of CH_4. Meirink et al. (2008b) also compared results produced using the TM5 4D-Var system to those of an analogous CH_4 synthesis inversion carried out by Bergamaschi et al. (2007), finding that the two methods produced results with a high degree of consistency when the spatial correlation length of the 4D-Var system was high. However, it was found that the 4D-Var system had much greater potential for reducing aggregation errors by reducing the spatial correlation lengths in the model.

Bergamaschi et al. (2009) continued to use the TM5 4D-Var system in order to produce updated global CH_4 emissions for the year 2004 from SCIAMACHY data, with high resolution $1° \times 1°$ emissions inventories provided for South America, Africa and South Asia. This study also showed the ability of the 4D-Var system to take account of individual source processes. The inversions in this study considerably changed the distribution of emissions compared with the a priori, with large increases in wetland emissions in South America and Africa being produced by the assimilation. Bergamaschi et al. (2010) used the TM5 4D-Var system in order to examine the trend in European CH_4 emissions over the period 2001 - 2006 by assimilating surface station data, following a study by Villani et al. (2010) which concluded that the sensitivity of the model to the European surface station network was high enough to constrain emissions in North West Europe. Bergamaschi et al. (2010) found that European emissions were 21% higher than those in the EDGAR version 4.0 inventory and 40% higher than those reported to the U.N. Framework Convention on Climate Change. The study also concluded that the constraints provided by the

European observational network are high enough to produce updated emission inventories that are independent from bottom-up CH_4 inventories.

Finally, and most recently, Bousquet et al. (2011) compared CH_4 emission estimates produced by a 4D-Var inversion system based on the LMDZt transport model version 4 to those of a synthesis inversion carried out using the same model, finding the two to be consistent on a global scale. The results of the 4D-Var inversion indicated that global CH_4 emissions increased by 21 Tg in 2007 and 18 Tg in 2008, compared with the 1999-2006 period. The increased emissions, found to be located mostly in the tropics and high latitudes, were thought to be mostly due to increased wetland fluxes due to high temperatures, although the level of agreement on a regional scale between the two inversions varied.

The research carried out as part of this thesis is intended to continue from, and add to, the results of the work discussed in this chapter, first by assessing the accuracy of our current CH_4 estimates and then by producing new, updated flux estimates using the 4D-Var inversion method. These aims require the use and extensive development of the TOMCAT CTM, which is described in full in the next section.

3.4 TOMCAT Model Description

The TOMCAT/SLIMCAT model is an Eulerian, grid-point off-line three-dimensional (3D) Chemical Transport Model (CTM) which was created in the early 1990s at the Centre National de Recherches Météorologiques (CNRM), Toulouse, in order to study the polar stratosphere. The first use of the model was described in Chipperfield et al. (1993). The model previously existed as two separate models, known as TOMCAT and SLIMCAT, which were used for modelling the troposphere and stratosphere, respectively. The two models were then combined to form the unified TOMCAT/SLIMCAT model (Chipperfield, 2006b), which has since been used and validated in a number of tropospheric studies (e.g. Monge-Sanz et al. (2007); Breider et al. (2010); Hossaini et al. (2010); Feng et al. (2011)). Since the majority of this thesis focuses on the troposphere, the model will henceforth be referred to as the TOMCAT model. The model meteorology, including winds, temperature and pressure data, is read in from analyses provided by the European Centre for Medium-Range Weather Forecasts (ECMWF, http://www.ecmwf.int) and transformed onto the TOMCAT model grid. The model uses a ‘process split’ method, in which separate advection, convection, boundary layer mixing and chemical routines are carried out in sequence.

3.4.1 Model Grid

The standard horizontal model grid in the TOMCAT model is a variable Gaussian grid in which the longitudinal spacing is regular, although the latitudinal spacing of the grid points may be irregular. The model may use non-Gaussian latitudes if required. Model grids used in this study are associated with typical spectral truncations, ranging from T10 (a grid of 11.25° × 11.25°) up to T106 (1.125° × 1.125°). The forcing meteorological analyses are provided by the ECMWF as spectral coefficients, and can be converted to grid-point fields for any prescribed latitudinal grid by a spectral transform using associated Legendre functions. The analyses may be read into the model at various intervals, usually set to be every six hours. The TOMCAT vertical grid is formulated using $\sigma - p$ vertical coordinates. This means that the model levels vary from purely terrain-following σ coordinates near the surface to pressure levels at high altitudes. A model vertical grid box interface $p_{k+\frac{1}{2}}$ is given by the formula;

$$p_{k+\frac{1}{2}} = Ap_0 + Bp_s \tag{3.10}$$

where p_s is the surface pressure and p_0 is a reference pressure of 100,000 Pa. In this study, vertical grids vary from 29 model levels with a top level at 10 hPa to 60 model levels with a top level at 0.1 hPa. The model dynamical time step is chosen according to the model grid resolution in order to satisfy the Courant-Friedrichs-Lewy (CFL) stability condition;

$$\Delta t \frac{u_i}{\Delta x_i} \leq 1, \quad 1 \leq i \leq 3 \tag{3.11}$$

where u_i is the wind speed in the ith dimension, x_i is the grid box length in that dimension and Δt is the dynamical time step. The model time step in this study ranges from 15 minutes up to 1 hour, depending on the model resolution.

3.4.2 Advection Schemes

Advection schemes in CTMs aim to accurately reproduce the atmospheric advection of tracers by parameterising the tracer mass continuity equation (Equation 2.1), whilst maintaining mass conservation and monotonicity and keeping numerical diffusion and dispersion to a minimum. In practice, however, no advection scheme is able to fulfil all of these properties. This study implements two Eulerian finite volume advection schemes which are available for use in the TOMCAT model. Unlike Lagrangian advection schemes, which consider the trajectories of individual tracer masses, Eulerian advection schemes evaluate the exchange of tracer mass between boxes of a model grid which is fixed in space. The first Eulerian scheme is the conservation of second-order moments (SOM) advection scheme described by Prather (1986), whilst the second scheme is the conservation of first-order moments scheme (FOM), which is equivalent to the ‘slopes’ advection scheme described in Russell and Lerner (1981).

The SOM advection scheme acts separately for advection in the zonal, meridonal and vertical directions, and conserves the total tracer mass in the model (zeroth-order moment), the tracer distribution gradient in each of the three dimensions (first-order moments) and curvature of this gradient, including cross terms (six second-order moments). This means that the scheme has a large computational storage requirement. However, it conserves tracer mass exactly, is stable for large time steps, and was found to produce very low numerical diffusion in comparison with other advection schemes (Prather, 1986), although it is not monotonic. These properties mean that the SOM scheme is the default advection scheme for use in TOMCAT model studies. The TOMCAT model also allows use of the FOM scheme which follows the same principle to that of the SOM scheme, but only conserves zeroth-order and first-order moments. This reduces the necessary computational storage compared with the SOM scheme, but also slightly increases the diffusivity (Prather, 1986). Prather (1986) found that use of the SOM advection rather than the FOM scheme was equivalent to increasing the resolution of the model grid by more than a factor of two while the corresponding increase in computational resources and time incurred due to the use of the scheme were smaller than they would be if doubling the resolution of the grid. In this study, the SOM advection scheme is used for tracer studies using the forward TOMCAT model. However, the FOM scheme’s relative simplicity and smaller computational burden mean that it is, at least initially, the preferred option for use in the TOMCAT 4D-Var inversion scheme. The accuracy of the FOM advection scheme in comparison to the SOM scheme is investigated in Chapter 4.

Vertical advection in the TOMCAT model is diagnosed from the divergence of the horizontal mass fluxes, maintaining continuity in order to eliminate the need for interpolation onto the model vertical grid. This method of determining the vertical advection may lead to differences between the grid box mass calculated on-line in the model and those produced by the ECMWF pressure analyses read in at each meteorological time step. In order to prevent discrepancies arising due to this error, the total grid box mass is overwritten each time the analyses are read in to the model (i.e. every 6 hours), and the tracer mass in each grid box is scaled according to the change from the calculated mass to the analysed mass. This has the effect of preserving tracer mixing ratio rather than tracer mass, and this function may be switched on or off as required. It is important to note that there is no continuity condition imposed on tracer distribution at the interface of the grid boxes in either of the advection schemes.

Difficulties arise when considering tracer transport at the poles when using an Eulerian grid. Due to the small size of polar model grid boxes in the east-west direction, an extremely small time step would be necessary in order for the advection scheme to be stable. Therefore, in both the SOM and the FOM advection schemes, zonal advection is carried out at high latitudes by grouping multiple grid boxes in the same latitude band together so that the CFL condition in Equation 3.11 is satisfied (i.e. Δx is increased). These grouped grid boxes were named 'extended polar zones' in Prather et al. (1987). The number of grid boxes to be grouped together depends upon the latitude, the grid box size, the model time step and the wind strength. After the advection step has taken place, the original grid boxes are reformed. An issue also arises when dealing with transport across the pole, due to the fact that a singularity is created by the convergence of the grid boxes neighbouring each pole. The TOMCAT model overcomes this problem by determining the mass of each grid box which would be transported to its diametric opposite box due to the wind vector. For more details see Chipperfield (2006a).

3.4.3 Convection Scheme

The parameterised moist convection scheme in the TOMCAT model described by Stockwell and Chipperfield (1999), is based upon the scheme developed by Tiedtke (1989), which diagnoses mass fluxes from the large-scale ECMWF meterological fields, although there are some modifications to Tiedtke's scheme. The scheme includes cumulus updraughts in the vertical direction and turbulent and organised entrainment and detrainment. However, mid-level convection and convective down-draughts are not included in the TOMCAT model. The model uses the mass fluxes diagnosed from the meteorological fields to determine the effect of convection on the tracer field. Feng et al. (2011) found that the diagnosed convection rates in TOMCAT are underestimated when compared with archived ECMWF convective flux rates, and that the inclusion of mid-level convection improved the comparison only moderately. Feng et al. (2011) found that the vertical extent of convection was also smaller in the TOMCAT model, reaching only around 200hPa, compared with 100hPa in the archived data. However, Hossaini et al. (2010) deduced from a study of the transport of the short-lived species $CHBr_3$ and CH_2Br_2 that vertical transport in the TOMCAT model is too rapid overall. This indicates that the slow convection rate is more than compensated for by rapid vertical advection rates in the TOMCAT model.

3.4.4 Planetary Boundary Layer Schemes

Two Planetary Boundary Layer (PBL) mixing schemes are available for use in the TOMCAT model, and both are used within this study. The first is the local boundary layer diffusion scheme described by Louis (1979), while the second is a non-local scheme developed by Holtslag and

Boville (1993). The Louis scheme is a first-order local diffusion scheme in which the eddy diffusivity coefficient is based upon the local Richardson number, which is dependent on wind gradients and potential temperature. Due to the local nature of this scheme, however, it may not take full account of large eddy transportation and entrainment into the PBL. The Holtslag and Boville scheme, which was developed for the National Center for Atmospheric Research (NCAR) Community Climate Model Version 2 (CCM2), does take account of non-local effects, and allows for counter-gradient transport. The implementation of the Holtslag and Boville scheme in the TOMCAT model is described in Wang et al. (1999), who used the scheme in a similar CTM, finding that the Holtslag and Boville scheme provided stronger tracer transport out of the PBL and into the free troposphere, giving a better match with observations of radon, methyl chloroform and CFC-11. Stockwell and Chipperfield (1999) found that the TOMCAT model produced large tracer concentrations near the surface and weak vertical mixing when the Louis scheme was implemented. Since the Holtslag and Boville scheme is a far more complicated piece of computer code, however, this made its use in the TOMCAT 4D-Var inversion scheme prohibitive. TOMCAT experiments are carried out in Chapter 4 which investigate the accuracy of the Louis scheme in the TOMCAT model and therefore its suitability for use in the inversion scheme. In order to attempt improve the vertical mixing rate into the free troposphere when using the Louis scheme, it is possible to increase the depth of the model layer closest to the surface, and the effectiveness of this technique is examined in Chapter 4.

3.4.5 Chemistry

The TOMCAT model has the option of a full tropospheric chemistry scheme based on the ASAD chemistry package (Carver et al., 1997), which integrates 62 chemical species, 42 of which are advected. The tropospheric chemistry scheme includes reactions for the O_x, HO_x and NO_x families, along with HNO_3, N_2O_5, HNO_4, H_2O_2, CO, HONO, H_2O, CH_4, HCHO, CH_3OOH, C_2H_6, C_3H_8, CH_3COCH_3, C_2H_5OOH, CH_3CHO, C_2H_5CHO, iCH_3H_7OOH, nC_3H_7OOH, $CH_3CO_3NO_2$ (PAN), $C_2H_5CO_3NO_2$ (PPAN) and CH_3ONO_2. Prior to this thesis, due to the lack of an accurate available CH_4 emission inventory and since CH_4 simulations had previously been short, the CH_4 tracer in the TOMCAT full chemistry scheme used only anthropogenic CH_4 emissions, which were then scaled to a mean surface mixing ratio of 1800 ppb in order to approximately replicate global emission totals. The full chemistry scheme is not used in this study, except in order to produce monthly mean concentration fields for the hydroxyl radical OH for investigation of the impact of the OH distribution on the CH_4 concentration. For more details of the full chemistry scheme and emissions of chemical tracers in the scheme, see Arnold et al. (2005).

The model also allows the option of using simplified parameterised chemistry schemes in place of the full tropospheric chemistry scheme. In the case of the CH_4 simulations carried out in this thesis, this allows the model to read in off-line monthly mean OH distributions and stratospheric CH_4 destruction rates instead of calculating them on-line. This method both increases the speed of simulations and allows us to choose the OH distribution according to preference, instead of relying on the distribution provided by the full chemistry scheme, which is highly dependent on the model performance in estimating concentrations of a large number of tropospheric species.

In order to evaluate the accuracy of the TOMCAT atmospheric transport schemes discussed in this section, it is necessary to compare the performance of the transport model against observations of atmospheric trace gases. This can help to improve our understanding of the model's limitations, and to investigate the variations in results produced using different transport schemes and model grids. The following chapter therefore evaluates the transport in the TOMCAT model by comparing the results of a number of model simulations of an inert atmospheric tracer to observations and to each other, allowing us to justify the model's use as the basis for a 4D-Var inversion system, in which accurate representation of atmospheric transport is key.

Chapter 4

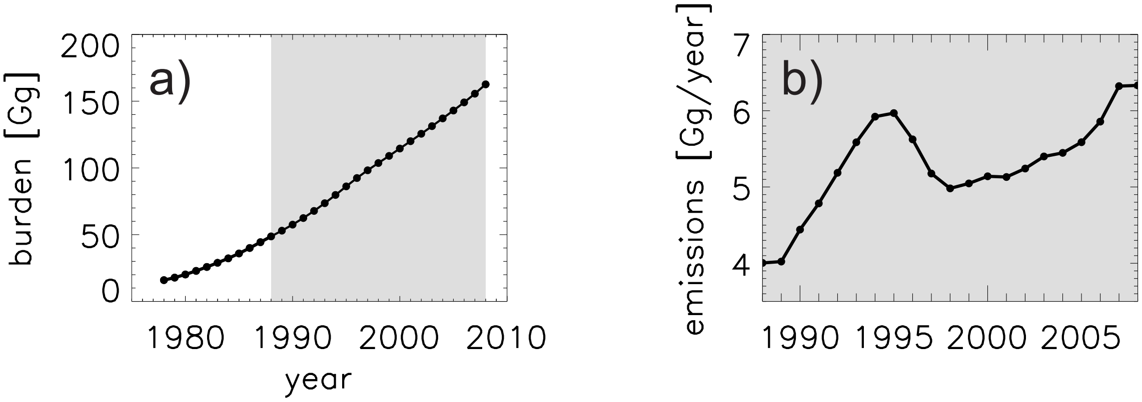
Evaluating Simulated Tropospheric Transport in the TOMCAT Model using SF_6

4.1 Introduction

This chapter examines the ability of the TOMCAT model to accurately replicate atmospheric transport by comparing the model distribution of an inert tracer, sulfur hexafluoride (SF_6), against observations. This is an important step when using a CTM as part of the 4D-Var inversion method, as modelling errors can lead to inaccuracies in the surface flux estimation produced using the inversion method (Tremolet, 2007). An awareness of the nature and size of these modelling errors is also necessary in order to estimate the error reduction of a 4D-Var inversion. This chapter also assesses the impact of changing the resolution of the model grid and the advection and PBL schemes on the model transport. In order to evaluate the tracer transport in the TOMCAT model, and to compare the different transport schemes available for use, a number of simulations of SF_6 were carried out using different model set-ups.

Section 4.2 gives background on the nature of SF_6 emissions and the recent atmospheric budget of the species. Section 4.3 give details of the model set-up and the observational data used for comparison in the chapter. Section 4.4 evaluates the accuracy of the transport in the TOMCAT model using SF_6 comparisons to observed data and investigates the effect of changes in the model grid, PBL scheme and advection scheme, while Section 4.5 summarises the results of this chapter.

Figure 4.1 *(Left)* Total atmospheric burden of SF_6 (Gg) for the period 1978 - 2008, as estimated in Levin et al. (2010) and *(Right)* Total global emissions of SF_6 (Gg year^{-1}) for the period 1988 - 2008 from the EDGAR version 4.0 inventory. The years 1988 to 2008 are shaded grey.



4.2 SF_6 as Diagnostic for Model Transport

Sulfur hexafluoride (SF_6) is a potent greenhouse gas which is inert in the troposphere and stratosphere, giving it an extremely long atmospheric lifetime which has been estimated to be between 800 and 3200 years (Morris et al., 1995; Ravishankara et al., 1993). The only atmospheric sinks of SF_6 are a relatively slow photochemical destruction process and electron capture, both of which only occur in the atmosphere above 60km, therefore having only a small impact on its tropospheric concentration (Reddmann et al., 2001). The low solubility of SF_6 means that oceanic uptake is negligible, and there is no known uptake by soil or plants (Maiss and Brenninkmeijer, 1998). Assuming an atmospheric lifetime of 3200 years means that SF_6 has a Global Warming Potential (GWP) 23,900 times that of CO_2 over 100 years (Solomon et al., 2007), making it the most potent greenhouse gas that the IPCC has evaluated, and its atmospheric concentration, whilst relatively low compared to other greenhouse gases, has increased by around two orders of magnitude since it was first industrially produced in 1953 (Maiss and Brenninkmeijer, 1998). Figure 4.1(a) shows the total global SF_6 inventory for the period 1978-2008, as estimated by Levin et al. (2010), indicating that by 2008 the total atmospheric burden had reached 160 Gg SF_6, more than ten times what it had been in 1978.

SF_6 has many properties which make it a useful tracer for testing the simulated long-term atmospheric transport in CTMs. First, the fact that it is inert in the troposphere and stratosphere mean that there is no need to include chemical processes in the model. The lack of any tropospheric reactions also means that we can accurately infer annual increase in atmospheric burden by measuring the concentration of SF_6 at remote surface sites. Second, the release of SF_6 into the atmosphere is almost entirely anthropogenic in nature. This means that emissions are fairly constant in time within 10% (Levin et al., 2010), with negligible seasonal cycle (Olivier and Berdowski, 2001),

Table 4.1 Details of model grid, time step and PBL scheme used in each TOMCAT SF_6 simulation.

Name	Longitudinal Spacing (°)	Latitudinal Spacing (°)	Vertical Levels	Time step	PBL scheme
T106_L	1.125	1.125	60	15 mins	Louis
T42_L	2.8	2.8	60	30 mins	Louis
T21_L	5.6	5.6	31	60 mins	Louis
T10_L	11.25	11.25	31	60 mins	Louis
T42_HB	2.8	2.8	60	30 mins	Holtslag & Boville
T21_HB	5.6	5.6	31	60 mins	Holtslag & Boville

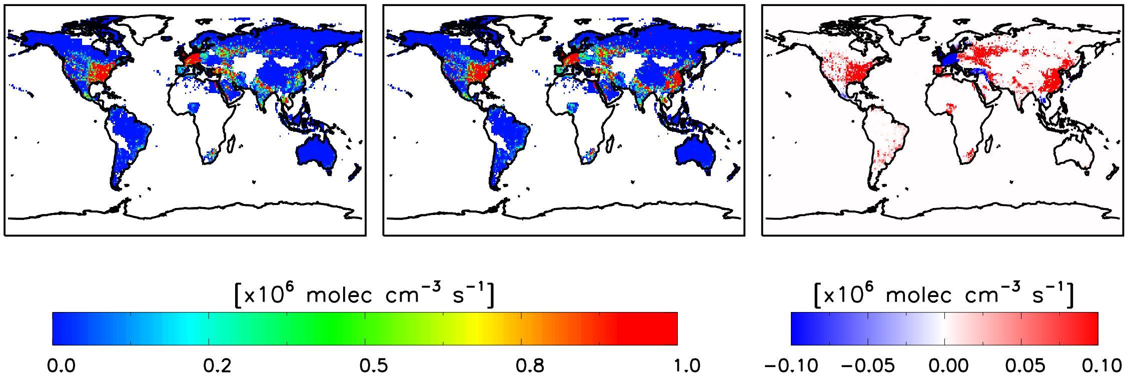
and that we can produce spatially accurate surface emission estimates by distributing national sales numbers spatially within each nation according to electrical energy use (Olivier, 2002). This procedure is discussed in more detail in Section 4.3.1.

In order to test the atmospheric transport in TOMCAT, multiannual simulations of global SF_6 concentrations have been carried out with the model. These simulations provide an indication how well the advection scheme, analysed winds and other transport schemes used in the TOMCAT model represent the long-range transport of the atmosphere, such as interhemispheric transport, zonal transport and seasonal large-scale atmospheric variations.

4.3 Model Set-up

20-year simulations of global SF_6 were carried out using TOMCAT. The 3-D concentration field was initialised on January 1 1988, with values provided as an auxiliary part of the TransCom CH_4 intercomparison (Patra et al., 2011). Two years were then allowed for the model spin-up, after which model data was output every 90 hours. The simulations were carried out using four different model grid resolutions, detailed in Table 4.1. The simulations at all four resolutions used the Louis boundary layer scheme, and were also repeated with the Holtslag and Boville scheme where possible. Winds were forced using ECMWF ERA-Interim meteorological fields with a temporal resolution of 6 hours. The model dynamical time step was chosen to be suitable to the simulation's spatial resolution, and is also given in Table 4.1. Two further simulations, T21_L29 and T21_HB_FOM, were carried out, which will be discussed in Section 4.4.4, which examined the effect of using the model advection scheme which conserves up to first-order moments only, and of changing number of vertical model levels when using the Louis PBL scheme.

Figure 4.2 Mean global distribution of SF_6 emissions (molecules cm^{-2} s^{-1}) for 1988-1995 *(Left)* and for 1996-2008 *(Centre)*. *(Right)* Change in distribution of emissions between the two periods.

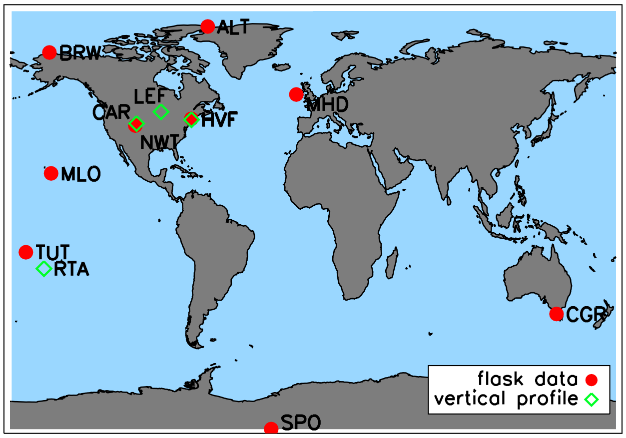


4.3.1 SF_6 Emissions

As discussed Section 4.1, emissions of SF_6 are almost entirely anthropogenic in nature. Trace amounts of SF_6 are produced in the Earth's crust, but these emissions are only enough to sustain atmospheric background levels of up to 0.01 ppt (Harnisch and Eisenhauer, 1998). SF_6 was first emitted anthropogenically in 1953, when it was used as an insulating gas for electrical switch-gear. It has since found further uses in the magnesium production industry and during the manufacture of semi-conductors. Due to leakage and venting during these industrial processes, the atmospheric concentration of the gas has risen from approximately zero in 1953 to a global mean of around 6.7 parts per trillion (ppt) by the end of 2008 (Levin et al., 2010). The nature of these emissions mean that a nation's SF_6 emissions are closely related to its level of industrialisation. After the 1950s, emissions continued to grow until the mid-1990s. In 1995 a voluntary protocol suggested by the United Nations Framework Convention of Climatic Change (UNFCCC) promoted a drop in SF_6 emissions in Europe and Japan, and total emissions decreased by around 17% over the following two years. Since then, however, increased emissions in the USA and newly industrialised countries in Eastern Europe and South-East Asia have led to global emission totals rising again.

This work uses SF_6 emissions provided by the Emission Database for Global Atmospheric Research (EDGAR), Version 4.0 (Olivier and Berdowski, 2001). Global consumption data for SF_6 and its distribution by country was estimated from sales statistics. These totals were spatially distributed over individual countries using information such as electrical consumption, semiconductor production and Chlorofluorocarbon (CFC) usage per country. Surface emission maps are available for 1980, 1990, 1995, 2000 and 2005, and estimates for all other years were scaled from these values. Figure 4.1(b) shows total global SF_6 emissions for the period 1988 - 2008. Total annual emissions reached a high of 6 Gg year^{-1} in 1995, before falling to 5 Gg year^{-1} two years

Figure 4.3 Locations of NOAA ESRL sampling stations (dots) and of vertical aircraft profiles (diamonds) used for SF_6 comparisons in Chapter 4.



later, after which emissions began to increase again. Emissions have recently surpassed their 1995 levels, reaching approximately 6.3 Gg year^{-1} in 2008. Figure 4.2 shows the mean spatial distribution of the EDGAR 4.0 SF_6 emissions for 1988 - 1995 and 1996 - 2008, as well as the difference between the two. Emissions in the early 1990s were mostly localised in Western Europe, Japan and the USA, with smaller contributions from South Africa, India and South-East Europe. After 1995 emissions in Western Europe and Japan decreased, while contributions increased in the USA, Eastern Europe and South-East Asia.

4.3.2 SF_6 Destruction

Sinks of SF_6 are confined to relatively slow destruction processes occurring only in the mesosphere, and are negligible in comparison to its sources. Hall and Waugh (1998) found that ignoring the effect of mesospheric destruction of SF_6 may lead to over-estimation of SF_6 concentration in the high-latitude middle stratosphere (above 30 km), but has only a small effect elsewhere. Therefore in this study, as in most previous SF_6 model studies (Denning et al., 1999; Peters et al., 2004; Gloor et al., 2007), the mesospheric destruction process is neglected, and SF_6 is treated as inert in the atmosphere.

4.3.3 SF_6 Flask Observations

Model output was compared with observed atmospheric SF_6 concentrations from a range of sources. Flask data from a number of surface sites from the National Oceanic and Atmospheric Administration, Earth System Research Laboratory (NOAA ESRL, USA) provide weekly records of SF_6 concentrations from 1995 onwards. The locations of these sites are shown as dots in Figure 4.3,

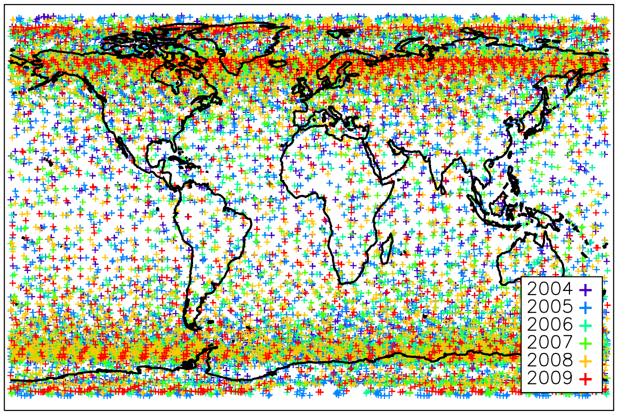
Table 4.2 Details of NOAA ESRL surface and vertical profile sampling sites used for SF_6 comparisons in Chapter 4.

Station Code	**Station Location**	**Longitude (°)**	**Latitude (°)**	**Altitude (m)**	**Observation Type**
ALT	Alert, Canada	62.5W	82.5N	210	surface flask
BRW	Barrow, Alaska	156.6W	71.3N	11	surface flask
MHD	Mace Head, Ireland	9.9W	53.3N	8	surface flask
LEF	Park Falls, WI, USA	90.3W	45.9N	N/A	vertical profile
HVF	Harvard Forest, MA, USA	72.3W	42.9N	340	surface flask and vertical profile
CAR	Briggsdale, CO, USA	104.8W	40.9N	N/A	vertical profile
NWT	Niwot Ridge, CO, USA	105.5W	40.9N	3021	surface flask
MLO	Mauna Loa, HA, USA	155.6W	19.5N	3397	surface flask
TUT	Tutuila, American Samoa	170.6W	14.2S	42	surface flask
RTA	Rarotonga, Cook Islands	159.8W	21.3S	N/A	vertical profile
CGR	Cape Grim, Australia	144.7E	40.7S	94	surface flask
SPO	South Pole	24.8W	90.0S	2810	surface flask

and further station characteristics are given in Table 4.2. The accuracy of such flask measurements is approximately 0.04 ppt. This data can be used in order to test long-range tropospheric transport in the TOMCAT model, such as inter-hemispheric and zonal transport, and seasonal variations. In order to compare model output to flask observations at surface sites, the 3-D modelled SF_6 concentration field is linearly interpolated to the latitude, longitude and altitude of the station site. However, it is necessary to remove the effect of numerical diffusion of European surface emissions on the model estimate at station at Mace Head, Ireland (MHD). Therefore, instead of interpolating along longitude to the station's location, the concentration at the centre of the grid box immediately to the west of that containing MHD is used. Interpolation to the correct latitude and altitude is still carried out, however.

As part of the NOAA ESRL Carbon Cycle Greenhouse Gases (CCGG) group's aircraft project,

Figure 4.4 Sampling locations of the ACE satellite, 2004-2009.

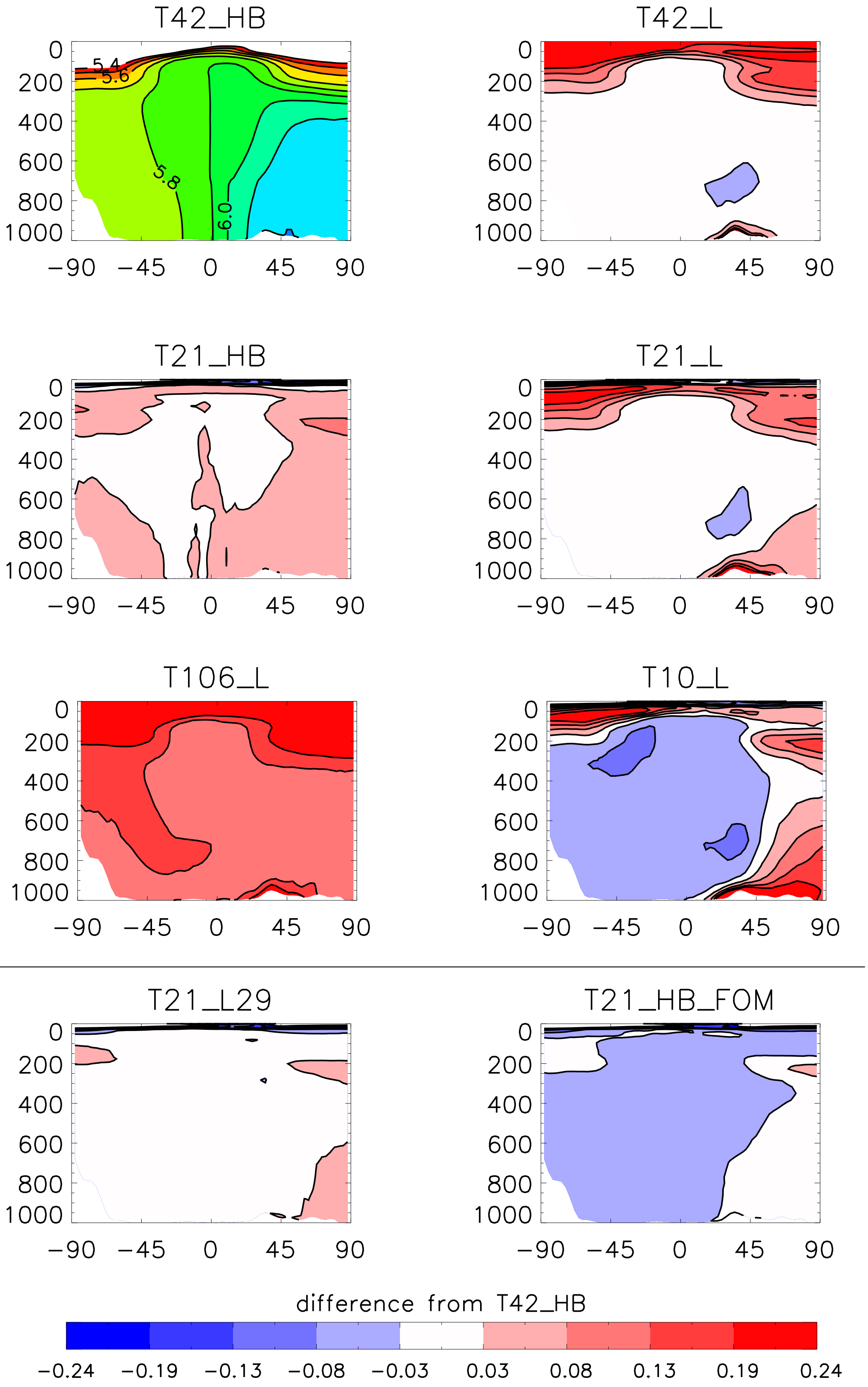


vertical profiles of CO_2, CO, CH_4, H_2, N_2O and SF_6 were measured using aircraft at a number of station sites across the USA and the surrounding oceanic area. Samples were also taken at some SH sites such as Rarotonga, the Cook Islands (RTA). These flask measurements have been taken once or twice a week for various time periods since 1999, with the aim of capturing seasonal and interannual changes in trace gas mixing ratios throughout the boundary layer and free troposphere. Locations of these vertical profile sites are shown as diamonds in Figure 4.3.

4.3.4 SF_6 from the ACE Satellite Instrument

The Atmospheric Chemistry Experiment (ACE) is a satellite mission on-board the Canadian satellite SCISAT-1, which was launched in high-inclination (74°) circular low-earth (650 km from the surface) orbit on August 13, 2003. This orbit gives it a good coverage of polar and mid-latitude regions, with reasonable tropical coverage also provided. The main instrument on-board SCISAT-1 is the ACE-FTS, a high spectral resolution Fourier Transform Spectrometer (FTS), which measures absorption spectra at sunrise and sunset in the limb viewing solar occultation mode. Limb views sample spectra through a slice of the atmosphere, from which vertical profiles of temperature, pressure and trace gases can be estimated. The ACE-FTS has retrieved SF_6 vertical profiles since 2004, and provides a comparison for the model simulations in the upper troposphere and lower stratosphere (UTLS) (10 - 25 km). Figure 4.4 shows the ACE sample locations for SF_6 for the period 2004-2009. These SF_6 retrievals will enable us to examine the model representation of vertical transport through the upper troposphere and lower stratosphere in TOMCAT.

Figure 4.5 Annual zonal mean simulated SF$_6$ (ppt) for 2006 for the eight model simulations described in Tables 4.1 and 4.5. *(Top Left)* shows the annual zonal mean SF$_6$ for the T42_HB simulation, while each of the other plots shows the difference between the labelled simulation and T42_HB (ppt).



4.4 Comparisons of Modelled SF_6 to Observations

Figure 4.5 shows the annual zonal mean distribution of SF_6 in each model simulation for the year 2006, highlighting differences in the vertical and latitudinal transport of the model grids and PBL schemes. In each simulation, the value of the model-observation difference at the SPO station in January 2000 is removed from each simulation in order to remove the effect of variations in the simulation initialisation. The model set-up described by T42_HB simulation is generally the default scheme for tropospheric studies, but for simplicity it is required that the Louis PBL scheme is initially used for the TOMCAT 4D-Var inversion system. Due to the large computational burden of the inversion, it would be favourable to use as low a grid resolution as possible, while still maintaining the integrity of the model results. It is important, therefore, to understand the differences in the model transport produced by changes in the PBL scheme and grid resolution.

Since the majority of sources of SF_6 are in the NH, atmospheric concentrations of the species are greatest there, and this is reproduced in the T42_HB simulation. Concentrations peak at approximately 6.2 ppt in the PBL at approximately 45°N, and decrease with altitude. Since vertical transport is faster than interhemispheric transport, the concentration in the upper troposphere in the NH is greater than the concentration throughout the SH. The modelled concentration of SF_6 is approximately 5.7 ppt throughout the SH, where the species is fairly well mixed. The concentration of SF_6 decreases rapidly in the stratosphere. The T42_L scheme displays the effect of using the Louis scheme rather than the Holtslag & Boville scheme. As discussed by Wang et al. (1999), the Louis scheme mixes emissions through the PBL at a slower rate than the Holtslag & Boville scheme, and this is confirmed by the T42_L simulation. Compared to the T42_HB simulation, concentrations are much greater in the NH emission region (30°N - 60°N) near the surface. Meanwhile, concentrations are a little lower in the NH mid-troposphere since emissions of the species are trapped close to the surface at these latitudes, although differences here are less than 0.1 ppt.

The effect of decreasing the resolution of the model grid is shown by the results of the T21_HB simulation. The decrease in resolution alone has little effect on concentrations throughout the atmosphere, with differences less than 0.1 ppt almost everywhere. The effect of having deeper model layers is seen, however, as vertical mixing in the tropics appears to increase as emissions mix upwards at a more rapid rate due to numerical diffusion through the grid boxes. The T21_L simulation shows the effect of decreasing the model grid resolution coupled with changing the PBL scheme, and shows similar results to the T42_L simulation. Increased concentrations close to the surface in the NH are again seen, with a corresponding decrease in concentration higher up. There are also slight increases in the low to mid troposphere in the NH, which may be due to increased polar transport due to the increased size of the model grid. Comparing the T42_L and T21_L schemes again shows that the main effect of decreasing the resolution is a slight increase in high latitude NH concentrations. The comparison of the T42_HB and T21_L plots shows that if the T21_L set-up were to be used for the TOMCAT 4D-Var inversion system, the main issue would be due slow boundary layer mixing, which has little effect away from emission regions.

The effect of increasing the resolution of the model grid is shown by the T106_L simulation, which shows increased SF_6 concentration throughout the atmosphere. This is likely to be due to a combination of the very small grid boxes and the rescaling of the tracer mass performed by the model every 6 hours. As emissions are released into the surface boxes, they create high mixing ratios due to the size of the grid box. However, each time the wind reanalyses are read into the model, the rescaling of the air mass in each grid box performed by the model to match the analyses conserves tracer mixing ratio rather than tracer mass, creating artificially high concentrations throughout the atmosphere. The very low resolution T10_L simulation shows the usual increased NH concentrations due to the use of the Louis scheme and also an increased effect of large grid boxes spreading the tracer towards to pole at a rapid rate. Concentrations are decreased throughout the SH and tropics, however, as the boundary layer mixing is extremely poor with the low-resolution grid. Although it would be desirable to use the T10_L model set-up for the TOMCAT 4D-Var inversion system, it would appear from Figure 4.5 that the resolution is too low to accurately reproduce atmospheric transport.

4.4.1 Surface Flask Comparisons

Annual growth rate of SF_6

Figure 4.6 shows comparisons of modelled and observed monthly mean SF_6 concentration at a number of station sites for the time period 2000 - 2006. For each model simulation, the model-observation difference at SPO in January 2000 has been subtracted everywhere in order to eliminate the effect of differences in the initialisations of the different simulations. Table 4.3 shows the yearly increase of SF_6 for each of the simulations, estimated by linear regression. The increase at SPO shows the annual SF_6 increase without the effect of local emissions, and each of the simulations is within 0.005 ppb of the observations by this measure. This indicates that the total SF_6 emissions used in the model over the period 2000-2006 are accurate. Table 4.3 also shows the mean annual increase over all of the station sites and the mean total increase over the period. There is more variation between the models by this measure, which is affected more by the transport of emissions than the SPO increase is, but each simulation matches the 6 year increase to within 0.08 ppt. At stations away from SF_6 source regions, such as ALT, BRW, CGR and SPO, the observations show little deviation from the linear increase due to emissions. This is captured well in the model using both PBL schemes. However, at stations close to source regions, such as HVF, a greater range of variation from the linear trend is observed, and this is simulated to some extent using the Holtslag & Boville PBL scheme. When using the Louis scheme, the model predicts

Table 4.3 Rate of increase for SF$_6$ (ppt yr^{-1}) over the period 2000 - 2006 for observations and for each simulation. Growth rate given for the SPO station and averaged over all stations, Mean increase over the whole period is also shown.

	OBS	**T42_HB**	**T21_HB**	**T42_L**	**T21_L**	**T106_L**	**T10_L**
SPO SF$_6$ increase (ppt yr^{-1})	0.224	0.229	0.222	0.229	0.221	0.228	0.219
Mean SF$_6$ increase (ppt yr^{-1})	0.237	0.233	0.226	0.234	0.227	0.236	0.224
Mean SF$_6$ increase 2000-2006 (ppt)	1.422	1.398	1.356	1.404	1.362	1.416	1.344

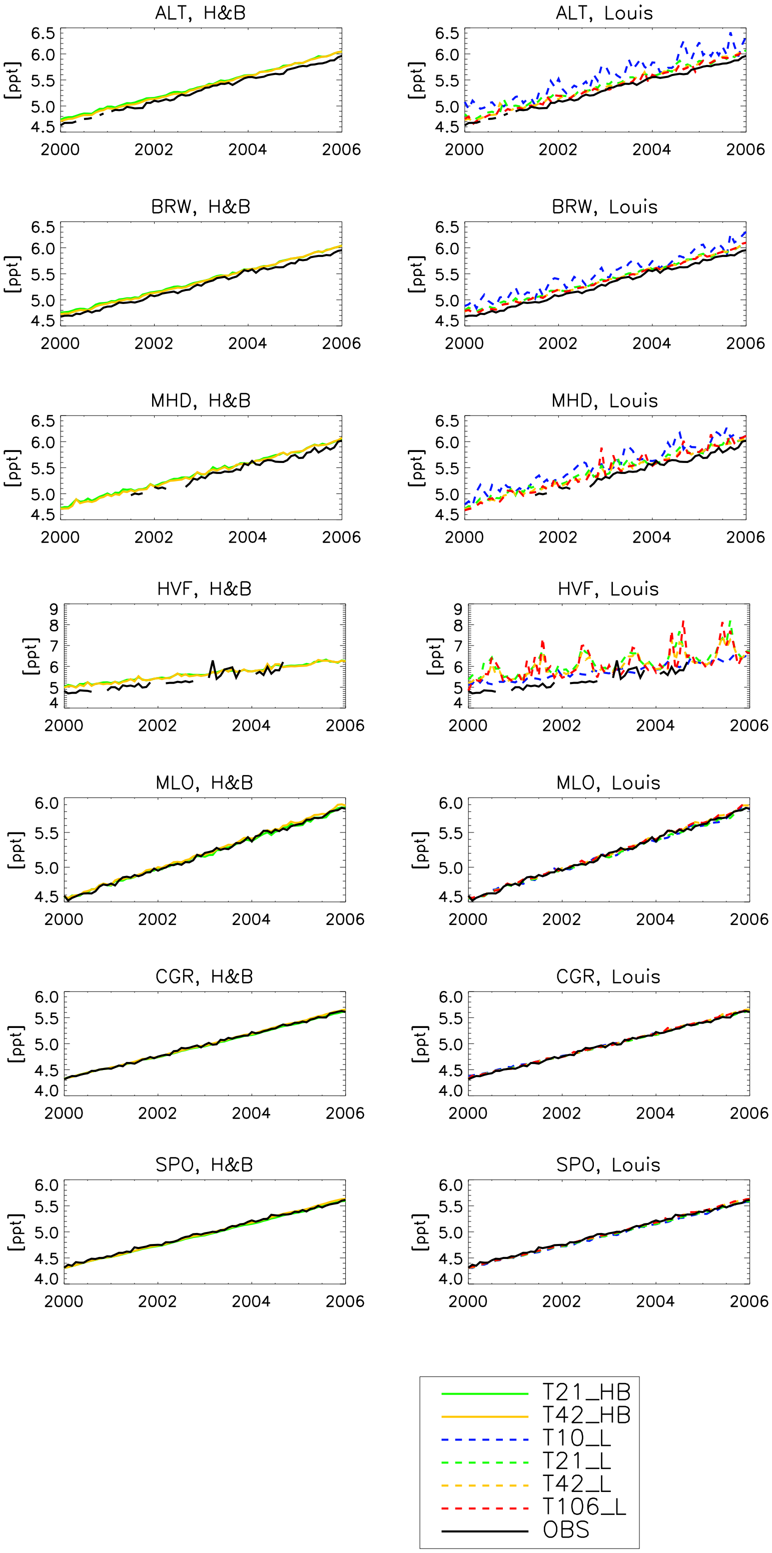
mixing ratios which are far too high in comparison to the observations, due to the limitations of the Louis scheme already discussed. This has a much smaller effect away from the source regions, however, where the two PBL schemes produce similar results. When using the Holtslag & Boville scheme, the choice of resolution had little effect on the SF$_6$ concentration at each station. However, when using the Louis scheme, the varying resolutions produced different results at stations close to the source regions, with the simulated mixing ratio being highly dependent on the depth of the grid-box into which the SF$_6$ emissions are deposited. In fact, the higher resolution runs generally produce higher mixing ratios at these stations, due to the emission of SF$_6$ into the smaller grid boxes producing a high concentration, which is then not mixed out by the PBL scheme.

SF$_6$ Seasonal Cycle

Figure 4.7 shows the observed and modelled seasonal cycle of SF$_6$ at the station sites. In order to display only the seasonal cycle due to transport, the linear trend displayed by SF$_6$ at the South Pole station (SPO), the furthest point from the source regions, was removed from all data. The modelled and observed SF$_6$ was averaged over the years 2003 to 2006, in order to only include years with observations provided for every month, and the mean value at each station over this time period was subtracted. Table 4.4 shows the correlation, r, and root mean square difference (RMSD) between the modelled and observed seasonal cycle at each station for each simulation, where;

$$= \sqrt{\overline{((x_{sim} - x_{obs})^2)}} \tag{4.1}$$

Figure 4.6 Monthly mean SF$_6$ concentration (ppt) at NOAA surface station sites, 2000 - 2006. Solid coloured lines *(Left)* represent TOMCAT simulations using the Holtslag & Boville PBL scheme, while dashed coloured lines *(Right)* represent model simulations using the Louis PBL scheme. Different model resolutions are depicted in different colours, and observations are shown in black.



Here, x_{sim} represents the simulated SF_6 mixing ratio and x_{obs} represents the observed concentration. The over-bar represents the mean. When using the Holtslag & Boville scheme, the model captures the inter-seasonal variation well at the majority of the sites. SH sites such as CGR and SPO have very little seasonal cycle, and the model shows extremely little variation around the mean in the SH. In fact, some SH seasonal variation may be missing in the model, as CGR displays slightly negative and positive anomalies in March and September, respectively, which are not reproduced in the model. This means that correlations and RMSDs are small at these SH stations. TUT displays positive anomalies in December through to March and negative anomalies for the rest of the year due to its position relative to the Intertropical Convergence Zone (ITCZ), which is the physical (rather than the notional 0°) boundary between the NH and SH. As the position of the ITCZ varies throughout the year due to the changing location of the sun's zenith point, TUT alternates between sampling NH and SH air. This oscillation is produced well in each of the model simulations, with correlations greater than 0.8 produced by each, and RMSD less than 0.01 for each resolution apart from T10_L. MLO displays positive anomalies in both MAM and SON, due to the increased influence of SH air at MLO during the NH summer and winter (Lintner et al., 2006), and, excluding the T10 resolution, the model produces the same semi-annual variation, with correlations greater than 0.65 and RMSD less than 0.012. NH high-latitude stations ALT and BRW show an annual seasonality with minimum concentrations produced in NH autumn, and a recovery period throughout the NH winter before concentrations peak around February. The negative anomaly is due to increased convection during JJA, while the recovery period is due to northward transport of SF_6 from NH emission regions during NH winter, increased stability of the boundary layer and the related pollution trapping in the Arctic during this time. The model reproduces the correct seasonal cycle at these Arctic stations when the Holtslag & Boville scheme is used, with positive correlations, although model anomalies are not as large as observed anomalies. The T42_HB simulation performs well at these Arctic stations, but the performance of the T21_L scheme is not as good, with lower correlations and RMSD at this resolution than in T42_HB. The high Arctic concentration of SF_6 during the NH autumn implies that some form of model transport is incorrect during this season. This may indicate strong transport into the Arctic, or slow transport away from it.

Use of the Louis PBL scheme has a small effect at the SH sites and at MLO, where transport is the main influence on the seasonal cycle, and at resolutions greater than 11.25° the seasonal cycle is reproduced by the model at these stations with high correlations. The 11.25° model grid produces seasonal variations different to those found in the observations due to effect of interpolation across the large grid boxes. The Louis scheme produces poor results at the MHD station, especially at low resolutions. This is likely to be due to poor vertical mixing of European emissions out of the PBL. Even at higher resolutions, modelled seasonal anomalies at MHD are up to six times greater than observed anomalies, although the T42_L simulation produces the smallest seasonal variations,

Table 4.4 Correlations and RMSDs (ppb) for monthly mean observed and modelled SF_6 concentration at selected remote surface sites averaged over the period 2003 - 2006 for each of the model simulations. The simulation producing the highest correlation or lowest RMSD at each station is shown in bold.

		Simulation					
		T42_HB	T21_HB	T42_L	T21_L	T106_L	T10_L
ALT:	*r*	**0.82**	0.38	0.06	0.11	-0.31	-0.28
	RMSD (ppb)	**0.0124**	0.0164	0.0214	0.0309	0.0231	0.0811
BRW:	*r*	**0.70**	0.58	0.54	0.33	0.50	-0.04
	RMSD (ppb)	**0.0145**	0.0157	0.0170	0.0218	0.0181	0.0516
MHD:	*r*	**0.44**	0.42	-0.36	-0.04	-0.45	0.06
	RMSD (ppb)	**0.0244**	0.0247	0.0445	0.0709	0.0641	0.0978
MLO:	*r*	0.75	**0.78**	0.70	0.70	0.67	0.06
	RMSD (ppb)	0.0100	**0.0094**	0.0108	0.0107	0.0114	0.0225
TUT:	*r*	0.88	0.87	**0.90**	**0.90**	0.88	0.83
	RMSD (ppb)	0.0085	0.0100	**0.0084**	0.0099	0.0086	0.0211
CGR:	*r*	0.21	0.11	**0.24**	0.10	0.00	-0.15
	RMSD (ppb)	**0.0102**	0.0105	0.0108	0.0107	0.0143	0.0141
SPO:	*r*	0.29	0.25	**0.41**	0.19	-0.03	0.65
	RMSD (ppb)	0.0079	0.0080	0.0075	0.0084	0.0105	**0.0063**

with large positive anomalies only in March and October. Generally the model performs poorly at this station when the Louis scheme is used, with lower correlations and larger RMSD than when the Holtslag & Boville scheme is used, which places some doubt on the model's ability to perform accurately near emission regions when using the local PBL scheme.

Figure 4.8 shows the mean annual interhemispheric difference (IHD) of the background SF_6 level for each year. In this case, the interhemispheric difference was defined as the difference between the NH and SH background SF_6 mixing ratios, where the NH background concentration was represented by the mean of the mixing ratios at BRW and MLO, and the SH background was similarly defined using CGR and SPO. These station sites were chosen as they are far from source regions and had a near-complete observational record over the chosen time period. All model runs produce a IHD which is too large (previously indicated by the modelled NH positive bias). However, the trend of the IHD over time shows good agreement with the observations, with smaller IHD during 1997 - 2004 due to the decrease in emissions during these years (see Figure 4.1(b)). The PBL scheme used in the model makes little difference to the value of the IHD, as the background concentration does not vary much between the two PBL schemes, but the IHD decreases as the resolution of the model grid increases. While there is a large improvement between the T10_L and T21_L resolutions, however, the improvement between T21_L and T42_L is much smaller, and there is no significant improvement in IHD gained by running the simulation at T106_L. Although, as seen in Figure 4.5, the T106_L simulation produces mixing ratios which are consistently higher than the T42_L simulation, the rate of interhemispheric transport appears to be the same.

Figure 4.7 Mean detrended SF_6 seasonal cycle at NOAA surface station sites, 2003-2006 using the Holtslag & Boville scheme *(Left)* anf the Louis scheme *(Right)*. Data is detrended as described in the text, and lines are coloured as in Figure 4.6. Note the different axis scales.

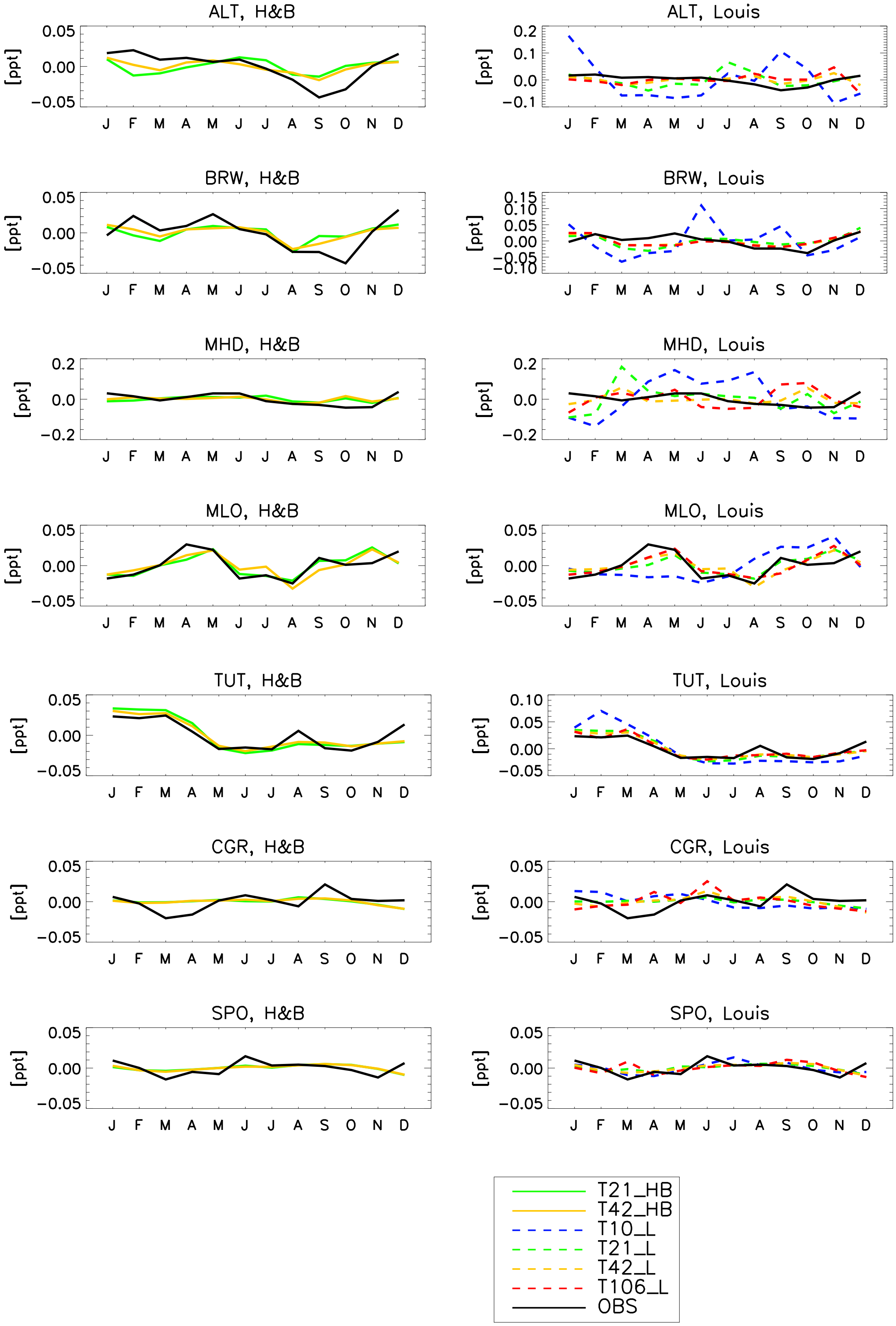


Figure 4.8 Annual mean SF$_6$ interhemispheric difference (ppt) for the period 1995 - 2008 using the Holtslag & Boville scheme *(Left)* and the Louis scheme *(Right)*. IHD is defined as in the text. Lines are coloured as in Figure 4.6

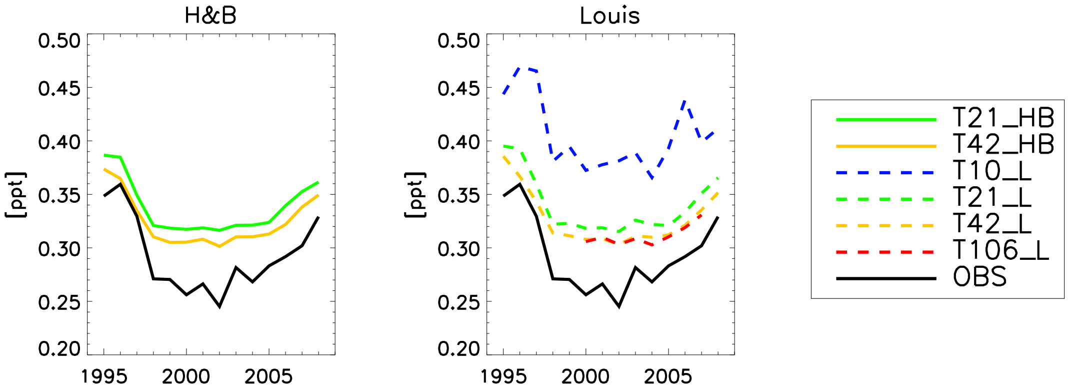


Figure 4.5 showed that interhemispheric mixing was slowest the T10_L simulation and increased as the model grid resolution decreased, but comparison with the observation shows that even at the highest resolution, the IHD is approximately 20% too high. Since the majority of SF$_6$ sources are in the NH, this indicates either that interhemispheric transport time is too slow in the model or that the NH emissions used are too high.

4.4.2 Aircraft Comparisons

Figure 4.9 shows the comparison of the model simulations with monthly mean vertical SF$_6$ profiles described in Section 4.3.1. Note that while LEF, CAR and RTA are sampled during 2003, HVF was compared during 2001 in order to maximise the amount of available flask data. Model profiles are provided for the T42_HB, T42_L, T21_HB and T21_L simulations. The NH bias seen at the surface sites propagates through the lower troposphere at each site, but the distribution of the profiles and the seasonal increase in SF$_6$ is simulated well. At the remote site RTA, the PBL scheme does not significantly alter the model distribution. The resolution of the model also makes little difference to the profiles at this SH station. However, the USA stations LEF, CAR and HVF, which are close to a source region, the slow vertical mixing of the Louis scheme can be seen. At these sites, the T42_L and T21_L profiles show a clear two-level divide, where the poor vertical mixing of SF$_6$ emissions lead to a high mixing ratio up to an altitude of 1-2 km, above which the concentration decreases to a level lower than that produced using the Holtslag & Boville simulations. At LEF, the observations show an almost constant vertical profile throughout the year except during MAM, when concentrations increase in the boundary layer. Both T42_HB and T21_HB reproduce this seasonal cycle. Throughout most of the year at CAR, modelled mixing ratios decrease above 4 km, while observed concentrations remain constant. During MAM, however, observed mixing ratios decrease with altitude across all levels, which is reproduced in the T42_HB simulation but not in T21_HB. At HVF, T42_HB and T21_HB match the observed vertical profiles well during all seasons, although concentrations are slightly high at the surface. In the SH station RTA, the observations display a large amount of variation around the profile, which generally increases with altitude. Generally the vertical transport in the model is good, but once again the Louis scheme produces high surface concentrations.

4.4.3 Comparisons to Satellite Data

For model comparisons with the ACE satellite, SF_6 retrievals were filtered by removing any retrievals with errors larger than 0.5 ppt. Model SF_6 was interpolated to each remaining retrieval location and pressure so that direct comparisons could be made. Only the T42_HB simulations were compared to the satellite data, and the same SPO bias was removed from the data as with the surface flask comparisons. Figure 4.10 shows a comparison between ACE retrieved SF_6 and modelled SF_6 for each retrieval in 2006, divided into latitude bins. At each level, the SF_6 mixing ratios estimated by the satellite display a greater range than those of the model, but the model captures the variation with altitude well. The majority of the satellite measurements are at a pressure level of 100 - 200 hPa, and this is where the agreement is best.

Figure 4.11 shows the latitudinal distribution of the vertical profile of SF_6 in the model and satellite retrievals over the period 2004-2009. The model reproduces the vertical profile of the satellite well. The standard deviation of the satellite retrievals is around 0.5 ppt, and the model falls well within these limits. There a few satellite measurements below 200 hPa, and the agreement is poor below this altitude. Agreement is greatest between 30°N and 30°S above 200 hPa. Brown et al. (2011) compared modelled SF_6 from the SLIMCAT model, which was initialised with the results of this study, to ACE observations in the tropics, finding the two to be in good agreement. The model performance in the UTLS is consistent with the satellite retrievals, indicating that TOMCAT troposphere-stratosphere exchange is robust enough to allow assimilation of observed stratospheric data.

4.4.4 Effect of Changes to the Model Advection Scheme and Vertical Grid

The Louis PBL scheme and the conservation of first-order moments (FOM) advection scheme were chosen for use in the adjoint TOMCAT model due to the fact that they are computationally simpler schemes than the Holtslag & Boville scheme and conservation of second-order moments (SOM) scheme, respectively, and therefore offer a starting point for the development of the adjoint model. In order to understand the implications of using these simpler schemes on tracer transport, a further SF_6 model simulation, T21_HB_FOM , was carried using the FOM advection scheme. Results using the Louis PBL scheme have already been displayed in this chapter, and it was discovered

Figure 4.9 Seasonal mean vertical SF$_6$ profile mixing ratio in ppt compared with NOAA ESRL aircraft flask samples for 2003 (LEF, CAR, RTA) and 2001 (HVF). The data has been averaged into 1km vertical bins. Lines are coloured as in Figure 4.6

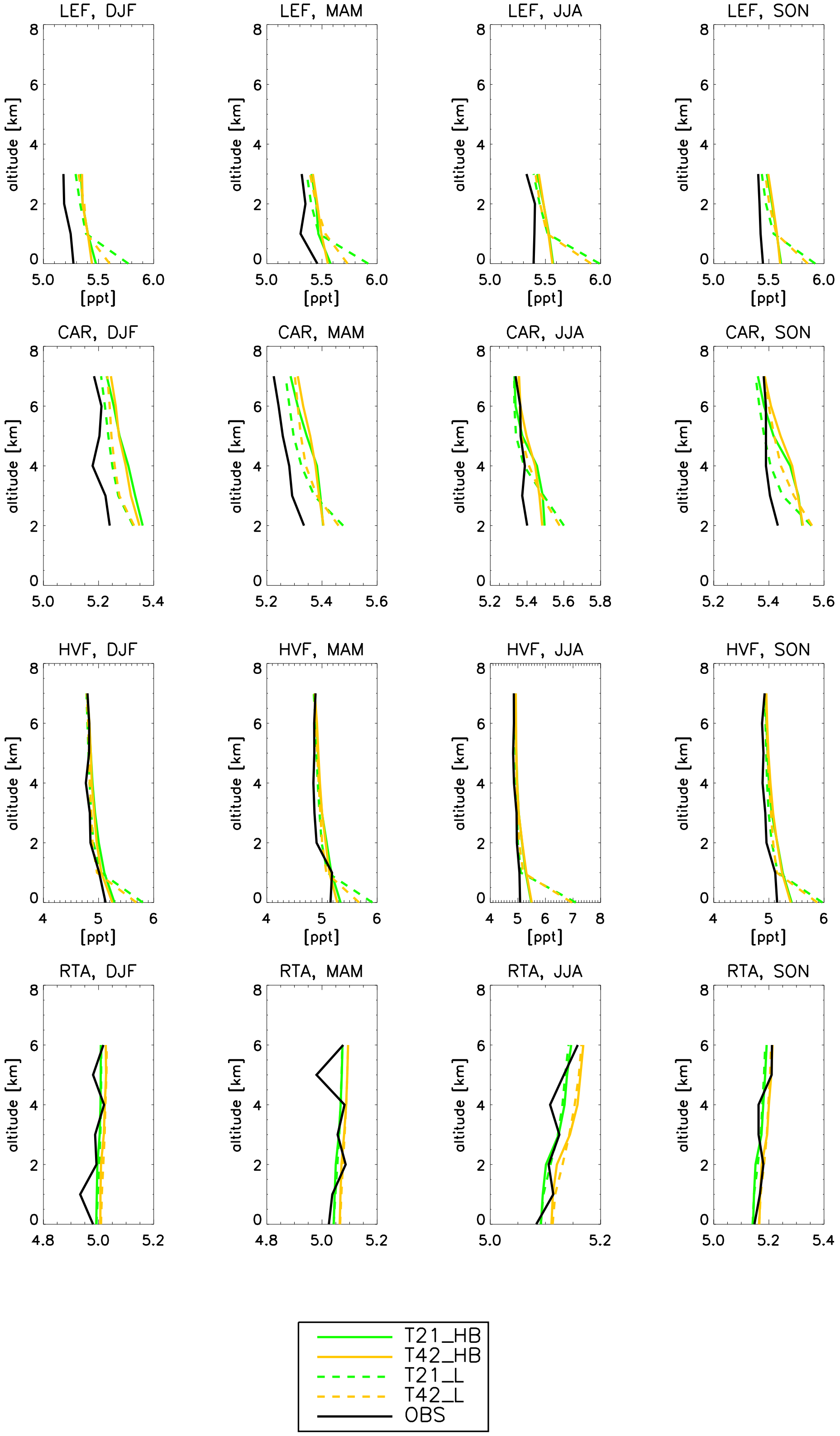


Figure 4.10 Comaprison of TOMCAT T42_HB simulation of SF$_6$ with ACE retrievals in ppt for the year 2006. Retrievals are split into latitudinal bins and coloured to represent the altitude of the retrieval

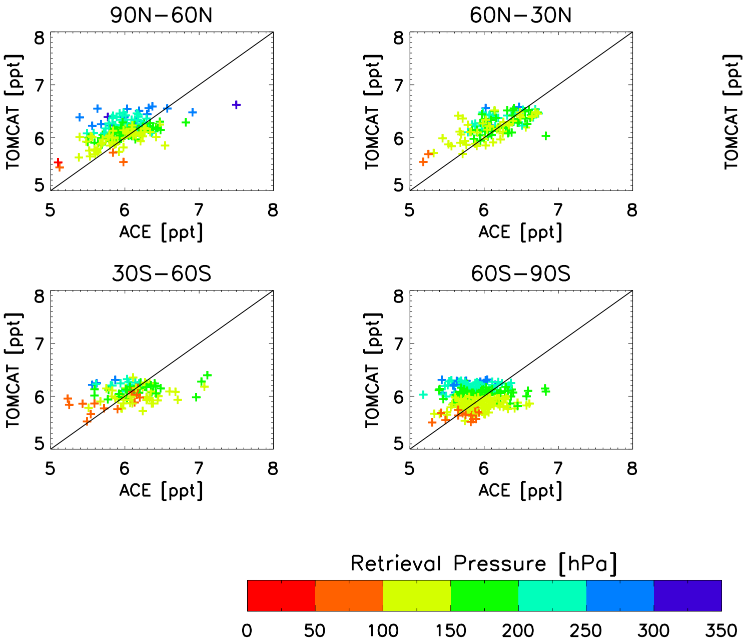


Figure 4.11 Mean vertical profile of TOMCAT T42_HB and ACE SF$_6$ for the period 2004 - 2009 split into latitudinal bins. Colours represent the latitude of the retrieval, dashed lines represent modelled SF$_6$ and solid lines represent satellite retrievals

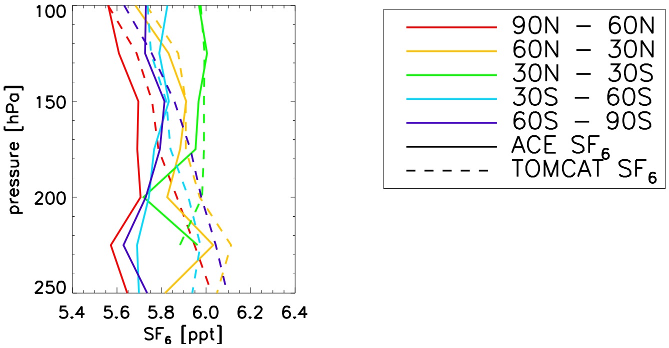


Table 4.5 Details of resolution, PBL scheme and advection scheme of further TOMCAT SF_6 simulations considered in Chapter 4.

Name	**Longitudinal spacing (°)**	**Latitudinal spacing (°)**	**Vertical Levels**	**Time step (mins)**	**PBL Scheme**	**Advection Scheme**
T21_L29	5.6	5.6	29 (bottom three levels of T21_L merged)	60	Louis	Second Order Moments
T21_HB_FOM	5.6	5.6	31	60	Holtslag & Boville	First Order Moments

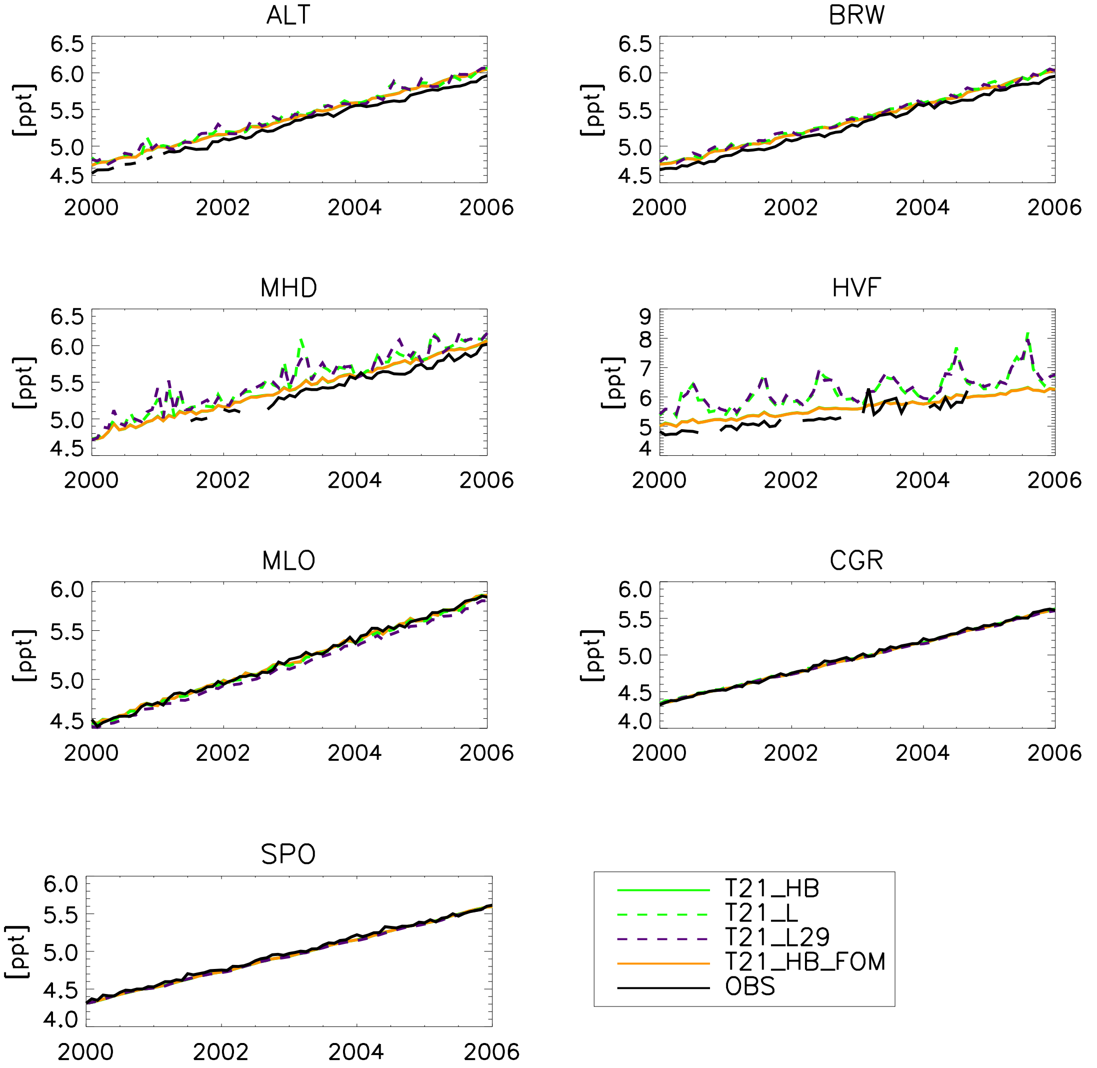
that the vertical mixing out of the boundary layer was not strong enough, which led to the model producing extremely large SF_6 mixing ratios in grid boxes close to emission regions. In an attempt to counter this problem, simulation T21_L29 was carried out in which the bottom three levels of the vertical grid were merged together, providing a larger grid box for emissions to be distributed into and therefore decreasing the mixing ratio. The surface grid box therefore had a depth of ∼200 m rather than ∼30 m. Both of these simulations were carried out using a grid box resolution of 5.6° and further details of these simulations are displayed in Table 4.5.

Figure 4.5 displays the annual zonal mean modelled distribution of SF_6 for the year 2006. Comparing the results of the T21_L29 simulation with those of the T21_L simulation in Figure 4.5 show that the mixing out of the boundary layer does indeed seem to have improved, with the ‘trapped’ emissions at around 30°N not appearing to the same extent as before. There is a slightly higher SF_6 concentration in the high latitude NH, but the effect is not as great as it was in T21_L. Concentrations are slightly lower in the SH than before but on the whole, the zonal mean distribution produced by the simulation agrees fairly well with that produced by the default T42_HB simulation. Use of the FOM advection scheme has produced slightly lower concentrations throughout the troposphere, but the scheme does not appear to have produced a significant change.

Figure 4.12 shows the SF_6 trend for the period 2000 - 2006 at NOAA station sites (as in Figure 4.6) for the T21_L29 and T21_HB_FOM model runs. At the surface, the T21_HB and T21_HB_FOM simulations are almost identical, indicating that the choice of advection scheme makes no difference over this long time period. The T21_L29 simulation does show improvements upon the T21_L simulation at MHD, decreasing model over-estimation by up to 50%. However, modelled SF_6 mixing ratio does not improve at HVF, and at both sites the observations are far more consistent with simulations when using the Holtslag & Boville scheme.

Figure 4.13, like Figure 4.7, shows the detrended seasonal cycle at the station sites. Table 4.6 shows the correlation and RMSD for the four T21 simulations. Again it should be noted that there

Figure 4.12 As Figure 4.6, for model simulations T21_L29 and T21_HB_FOM



is extremely little difference between the two advection schemes and that the T21_L29 simulation offers no improvement over T21_L at most sites. In fact, the agreement with observations at ALT, BRW, MLO and TUT is worse in T21_L29. At MHD, the surface site closest to an emission region, the large positive model anomaly in March is reduced in T21_L29. However, a new positive anomaly is produced in October.

Figure 4.14 shows model comparisons to observed vertical profile data as previously seen in Figure 4.9. Once again, the FOM advection scheme does not alter the performance of the model. In this case, however, the improvement offered by the T21_L29 simulation over the T21_L scheme can be seen. At the USA sites, LEF, CAR and HVF, the SF$_6$ mixing ration in the boundary layer (up to 2 km) and upwards is reduced when using the T21_L29 scheme than in the standard Louis scheme,

Figure 4.13 As Figure 4.7, for model simulations T21_L29 and T21_HB_FOM

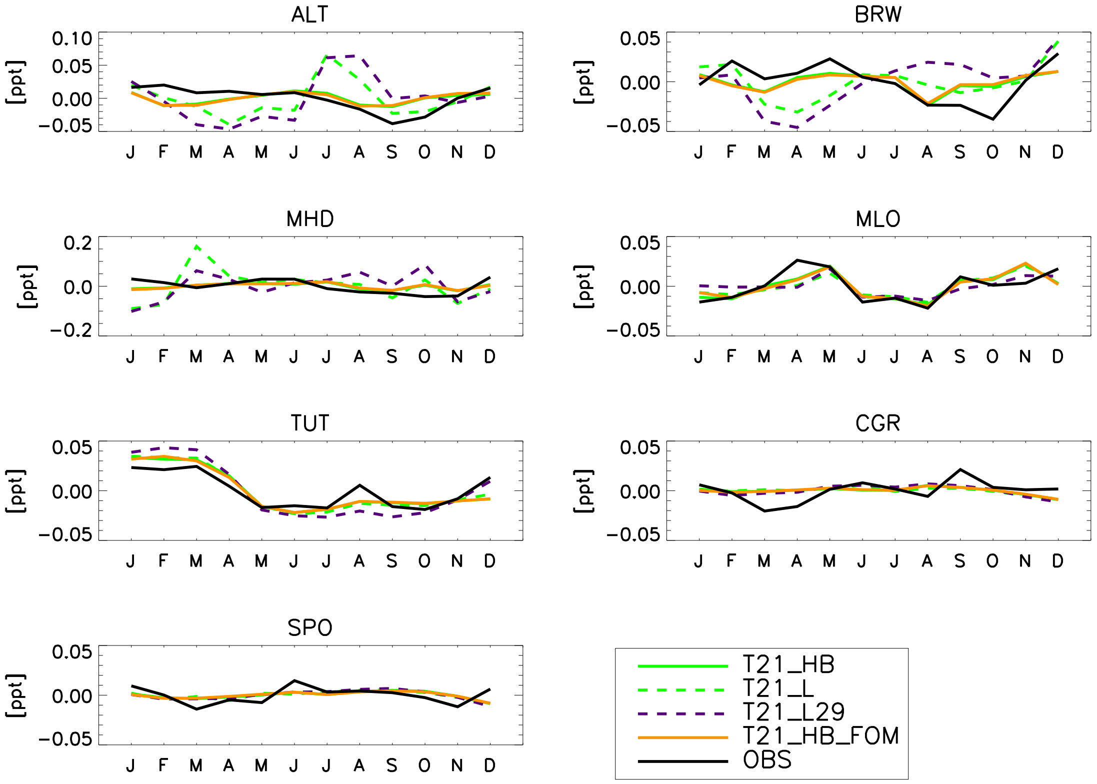


Table 4.6 Further correlations and RMSDs (ppb) for monthly mean observed and modelled SF_6 concentration at selected remote surface sites averaged over the period 2003 - 2006 for each of the model simulations. The simulation producing the highest correlation or lowest RMSD at each station is shown in bold.

	Simulation			
	T21_HB	T21_L	T21_L29	T21_HB_FOM
ALT: *r*	**0.38**	0.11	-0.30	0.35
RMSD (ppb)	**0.0164**	0.0309	0.0432	0.0167
BRW: *r*	**0.58**	0.33	-0.13	0.53
RMSD (ppb)	**0.0157**	0.0218	0.0328	0.0164
MHD: *r*	**0.42**	-0.04	-0.44	0.29
RMSD (ppb)	**0.0247**	0.0709	0.0707	0.0263
MLO: *r*	**0.78**	0.70	0.69	0.74
RMSD (ppb)	**0.0094**	0.0107	0.0111	0.0102
TUT: *r*	0.87	0.90	**0.92**	0.87
RMSD (ppb)	0.0100	**0.0099**	0.0134	0.0100
CGR: *r*	0.11	**0.10**	0.29	0.11
RMSD (ppb)	0.0105	0.0107	**0.0101**	0.0105
SPO: *r*	0.25	**0.19**	0.24	0.21
RMSD (ppb)	**0.0080**	0.0084	0.0084	0.0081

bringing the modelled mixing ratio closer to those observed on the aircraft missions. At the remote site RTA, T21_L29 produces a small decrease in concentration.

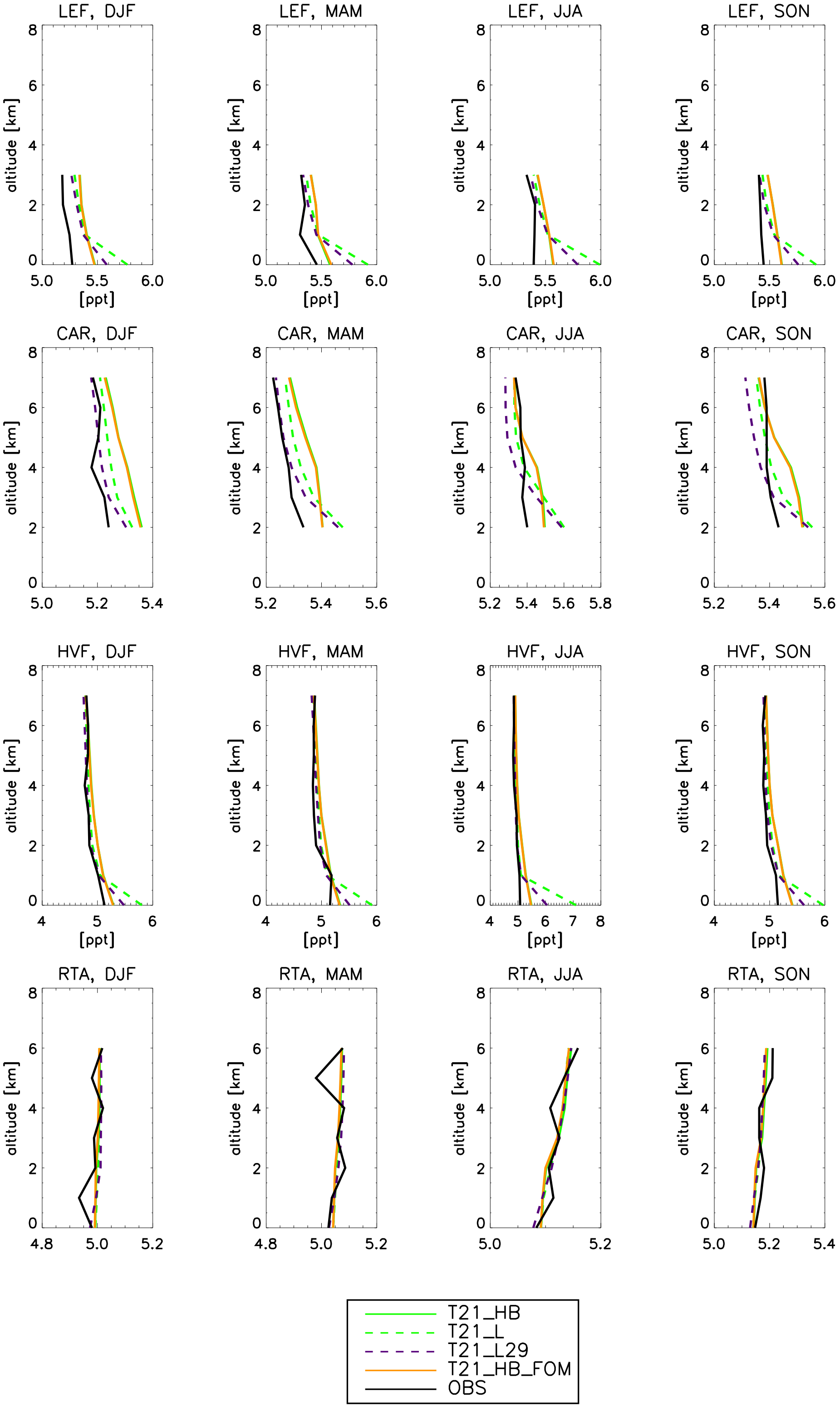
4.5 Summary

TOMCAT model comparisons with SF_6 observations have provided a validation of the atmospheric transport in the model, and indicated its potential for use as part of a 4D-Var inversion process. These comparisons have also provided an understanding of the differences that each transport scheme and model grid resolution available for use in the model produces in terms of large scale transport, vertical mixing and seasonal variations. Since the 4D-Var inversion process uses simplified advection and PBL mixing schemes, it is important to understand the effect that this will have on a surface flux inversion produced using the model. Comparisons of the two different advection schemes indicate that the FOM scheme does not produce significantly different results to the SOM scheme. The Louis PBL scheme, however, does produce different tropospheric distributions to model simulations using the Holtslag & Boville scheme, with slightly slower vertical transport and much less vertical mixing within the boundary layer. This is only partly compensated for by introducing a vertical model grid with a larger surface grid box. The resolution of the model grid can have a large effect on vertical and meridional transport, but the T42 and T21 resolutions perform well in comparison with observations.

When run using its ‘optimum’ set-up, T42_HB, the model performs well in comparisons with observed SF_6. Comparisons with flask samples at stations sites provide a validation of the large-scale transport in the model such as interhemispheric and zonal transport and representation of seasonal large-scale atmospheric variations such as the ITCZ. The T42_HB simulation reproduces the annual increase at each station site compared with observations, and also predicts the phase and amplitude of the seasonal cycle at many stations. Some SH seasonal transport variations are not reproduced in the model, and model transport in the Arctic may not be strong enough during the NH autumn. However, the observation accuracy of 0.04 ppt at these stations is close to the absolute seasonal variation in many places, so the model is always with in the observational accuracy. Any model transport errors may be taken account of when producing the 4D-Var inversion. Vertical transport up to the lower stratosphere is also simulated accurately when using the optimum model set-up.

The 4D-Var inversion uses the T21_L29 resolution with the Louis PBL scheme and the FOM advection scheme. As already discussed, the two advection schemes do not produce significantly different transport in the model, so the FOM advection scheme may be used with confidence. However, use of the Louis scheme may produce underestimation of fluxes, since tracers are not mixed vertically through the boundary layer, even when the size of surface-layer grid boxes is

Figure 4.14 As Figure 4.9, for model simulations T21_L29 and T21_HB_FOM



increased. Vertical transport into the upper troposphere is slower when using the 5.6° model grid than when using the 2.8° grid, which may also lead to underestimation of surface fluxes during using the inversion process. Theses issue will have to be taken account of when the modelling error term of the inversion is created. Modelled interhemispheric transport is slightly slower than indicated by observations, but within the range found by Denning et al. (1999), and the annual variation of the cross-hemispheric transport is reproduced in the model. The large-scale seasonal variation of the ITCZ is also produced in the model.

A key assumption when the 4D-Var inversion method is used in data assimilation is that all errors, including model errors, are unbiased and Gaussian in nature. While the optimum set-up of the TOMCAT model performs consistently with observations, it is clear that the introduction of the Louis PBL scheme produces a systematic error into the model, which cannot be comprehensively handled in any potential 4D-Var inversion performed with this PBL scheme. For the purposes of the development of the TOMCAT 4D-Var inverse model, however, the decision has been taken to persevere with the T21_L29 set-up described in this chapter. While this introduces bias into the model error term of the inversion, the relative simplicity of the Louis scheme compared to the Holtslag and Boville scheme makes it the preferred option in the short-term future. However, the development of the adjoint version of the Holtslag and Boville scheme will clearly be necessary when there is opportunity.

Chapter 5

Inverse Modelling using the 4D-Var Method: Development and Testing for the TOMCAT Model

5.1 Introduction

Chapter 4 assessed the quality of the representation of atmospheric transport in the TOMCAT model in order to validate the model's capability for use in the 4D-Var inversion method of estimating surface fluxes of atmospheric species. The model's performance in simulating the global distribution of the long-lived tracer SF_6 was accurate when using the optimum set-up. Although the quality of the model performance deteriorated due to the use of the Louis PBL scheme, extending the depth of the bottom model layer improved the results, and it was decided that the benefits of the Louis scheme in terms of computational simplicity out-weighed the inaccuracies produced in the results. This chapter therefore introduces the 4D-Var data assimilation method in detail before discussing the development and testing of the TOMCAT 4D-Var system.

Section 5.2 provides details on the modelling steps involved in the 4D-Var method and provides information on its derivation. Section 5.3 describes the development of the TOMCAT 4D-Var system, after which the chapter describes results from a series of idealised model experiments which test the 4D-Var scheme and demonstrate its performance. Section 5.6 details results of a simple inversion performed with the 4D-Var system which replaces observed data with synthetic 'observations' provided by the forward TOMCAT model before Section 5.7 summarises the results of the chapter.

5.2 Background Theory to the 4D-Var Method

The 4D-Var data assimilation technique is derived from Bayesian theory, and uses iterative methods in order to reduce as far as possible the value of a function measuring the difference between model predictions and observations. The method has previously been used in studies discussed in Chapter 3. 4D-Var requires the use of an adjoint model in order to evaluate the sensitivity of model-observation differences to boundary conditions and parameters. Within the scope of this work, this entails the construction of an adjoint version of the TOMCAT model, which integrates model-observation differences in the atmospheric mixing ratio of a species, such as CH_4, backwards in time in order to estimate the sensitivity to its surface fluxes. This section gives an outline of the 4D-Var method, with all notation specified as in Ide et al. (1997), described in Section 3.2.1.

Generally, if $\mathbf{x}$ is a vector requiring optimisation, known as the state vector, and observations $\mathbf{y}$ and observation operator $H(T(\mathbf{x})) : \mathbf{x} \longrightarrow \mathbf{y}$ exist, then the aim of the 4D-Var method is to minimise, in a least square sense, the model-observation difference $[\mathbf{y} - H(T(\mathbf{x}))]$. Since the number of observations over time $[t_0, t_n]$ is usually far less than the dimension of $\mathbf{x}$, the minimisation is further constrained using an *a priori* estimate, $\mathbf{x}_b$, of the state vector. The cost function $J(\mathbf{x})$, a weighted measure of the model-observation difference combined with information from the *a priori*, is therefore defined as

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y} - \mathbf{H}(\mathbf{T}[\mathbf{x}]))^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}(\mathbf{T}[\mathbf{x}])) \tag{5.1}$$

where $\mathbf{R}^{-1}$ represents the inverse of the error covariance matrix of the observations, which can also include model error, and $\mathbf{B}^{-1}$ represents the inverse of the error covariance matrix of $\mathbf{x}_b$. If m is the dimension of the state vector and n is the number of observations, then $\mathbf{H(T)}$ is a matrix of size $n \times m$ representing $H(T)$, which maps $\mathbf{x}$ onto $\mathbf{y}$.

In order to find the optimal state vector, it is necessary to minimise the value of the cost function. The minimum of $J(\mathbf{x})$ is found when the gradient of J with respect to the state vector is equal to zero;

$$\nabla_{\mathbf{x}} J(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \mathbf{T}^T \left[\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}(\mathbf{T}[\mathbf{x}]))\right] = 0 \tag{5.2}$$

In order to find a solution to Equation 5.2 it is therefore necessary to compute $\mathbf{T}^T$, the transpose of the observation operator matrix, known as the adjoint operator. For this work, the state vector, $\mathbf{x}$, is the 2D surface flux field of CH_4 at a given time, which is used in the TOMCAT model, T. H is the observation operator which maps the 3D concentration field produced by T onto the observations, $\mathbf{y}$.

In reality, the observation operator matrix, $\mathbf{H}(\mathbf{T})$ is not found explicitly, but is replaced with the forward model, which outputs model concentration fields at each point in time, i, where observations exist. These fields are then interpolated to the locations of the observations. The adjoint matrix $\mathbf{H}^T$ can be replaced with an adjoint version of the model, which is created directly from the forward model. Further details on adjoint modelling are given in Section 5.2.1. The cost function is minimised through an iterative process which computes the value of $J(\mathbf{x})$ and $\nabla_{\mathbf{x}} J(\mathbf{x})$ at each iteration, a process which entails a large computational burden in terms of both computer memory and processing time. During each iteration, the gradient of the cost function is computed using the adjoint and forward models, and an appropriate descent direction is chosen along which to minimise the cost function. Once the minimum along this direction has been found, the state vector is updated according to the minimisation so far, and the process is repeated using the new state vector as an initial guess. Eventually the value of the norm of the gradient of the cost function will decrease past a threshold value chosen by the modeller, indicating that a close-to-optimal state vector has been produced. This threshold value is known as the convergence criteria. The iterative process of minimisation will be discussed further in Section 5.3.2.

5.2.1 Adjoint Modelling

As discussed in Section 5.2, atmospheric CTMs can be viewed as numerical operators, T, acting on the state vector $\mathbf{x}$. In practice, T consists of parametrisations of various transport processes such as advection, turbulent mixing and convection, and each of these parametrisations is made up of a finite number of mathematical operations, T_j;

$$T = \prod_j T_j \tag{5.3}$$

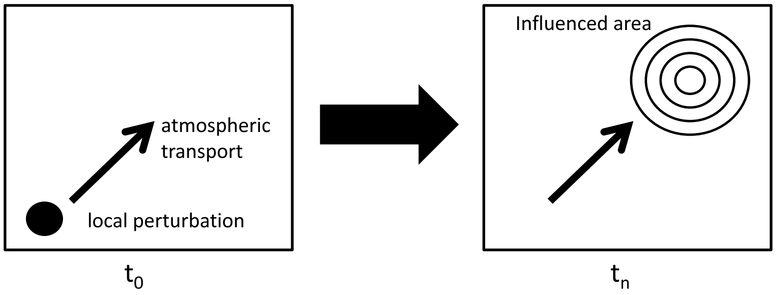
Each T_j may be linear or non-linear, and differentiable or non-differentiable. Assuming that the operator T is differentiable, its first derivative, or Jacobian, can be represented by a tangent linear model (TLM), T'. The TLM is linear, simulates the propagation of perturbations within T and is dependent upon the trajectory of T at which the linearisation took place. If the forward model T is defined such that, for a concentration field $\mathbf{c}$ at time t_i;

$$\mathbf{c}(t_{i+1}) = T[\mathbf{c}(t_i)] \tag{5.4}$$

Then the TLM is defined such that;

$$\delta\mathbf{c}(t_{i+1}) = T'[\mathbf{c}(t_i)]\delta\mathbf{c}(t_i) = \frac{\partial T[\mathbf{c}(t_i)]}{\partial \mathbf{c}}\delta c(t_i) \tag{5.5}$$

Figure 5.1 Atmospheric transport of a local perturbation in the TLM at times t_0 *(Left)* and t_n *(Right)*. Adapted from Giering (2000).



Note that the model operator is differentiated with respect to the concentration field **c**, and not the perturbation $\delta\mathbf{c}$. In practice, this means that the trajectories of the forward model must be stored in order to run the TLM. Each mathematical operation T_j must be individually differentiated to create T'_j;

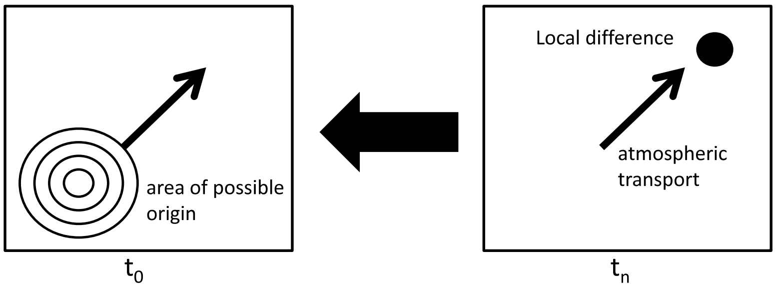
$$\delta\mathbf{c}(t_{i+1}) = \prod_j T'_j[\mathbf{c}(t_i)]\delta\mathbf{c}(t_i) = \prod_j \frac{\partial T_j[\mathbf{c}(t_i)]}{\partial\mathbf{c}}\delta\mathbf{c}(t_i) \tag{5.6}$$

Of course, if the model T is already linear, then the TLM is identical to the forward model. From the TLM, the adjoint model (ADM), T^*, can also be developed. The adjoint model is the representation of the transpose of the matrix representing the TLM and propagates sensitivities backwards through time. The ADM is defined such that, for an inner product $\langle\ ,\ \rangle$ and for vectors **x** and **y**, it holds that;

$$\forall\mathbf{x},\forall\mathbf{y} \quad \langle T'\mathbf{x},\mathbf{y}\rangle = \langle\mathbf{x},T^*\mathbf{y}\rangle \tag{5.7}$$

Figures 5.1 and 5.2 illustrate the difference between the TLM and the ADM. Figure 5.1 shows a local perturbation at time t_0 being transported downwind and broadened by diffusion at time t_n. If the initial perturbation is small, the TLM can therefore be used in order to investigate the impact of this small disturbance later in time. Figure 5.2, meanwhile, shows the process modelled in the ADM. A local difference at time t_n could be due to an anomaly further upwind. The ADM finds the sensitivity of this local difference to other locations at time t_0. Again, the anomaly undergoes diffusion as it travels upwind, indicating that the possible origin of the anomaly is located in a broad region. This adjoint transport process is clearly not the same as integrating the forward model backwards through time.

Figure 5.2 Adjoint atmospheric transport of a local anomaly in the ADM at times t_n *(Right)* and t_0 *(Left)*. Adapted from Giering (2000).



In practice, the forward model consists of numerous lines of computer code, each carrying out a mathematical operation or creating a conditional statement or loop. TLM and ADM codes can therefore be created from the forward code by hand or through the use of automatic code generators. A variety of these are available, such as TAMC (http://www.autodiff.com/tamc) and TAPENADE (http://tapenade.inria.fr), which create an TLM or ADM from a supplied forward model without the need for the time-consuming hand-coding process. Both of these methods may produce errors or inconsistencies in the adjoint code however, whether due to human error or through bugs produced by the automatic coding process (Nehrkorn et al., 2006), and it can often be more efficient to code by hand in the first place.

5.3 Development of the 4D-Var Method for the TOMCAT Model

5.3.1 Creating the Adjoint TOMCAT Model

As discussed in Chapter 3, the TOMCAT model simulates horizontal and vertical advection, moist convection, PBL mixing and chemical processes in order to produce mixing ratios of atmospheric species. A choice of schemes exists for many of the transport processes, and each scheme is contained in a separate model subroutine so that it may be included or omitted as required. Schemes may be linear or non-linear, and in order to create the adjoint version of the TOMCAT model, it is necessary to create the adjoint version of each subroutine independently. For this work, and as a first step in the construction of a full adjoint version of the TOMCAT model, it was important to find a balance between accuracy and simplicity when choosing appropriate transport schemes for the adjoint model. Consequently, the conservation of first-order moments (FOM), or ‘slopes’, scheme was chosen to simulate advection, while the Louis scheme was chosen as the PBL scheme. As seen in Chapter 4, there is little difference between results produced using the slopes scheme and those using the conservation of second-order moments (SOM) advection scheme, and there is little structural difference between the codes of the two schemes. Producing an adjoint SOM scheme would be a small step forwards once the adjoint slopes scheme exists. However, the SF_6 simulations discussed in Chapter 4 also showed that the Louis PBL scheme produces significantly different results to those produced using the Holtslag & Boville scheme on the default verticval boundary layer grid. The decision to use the Louis scheme was taken due to the complex nature of the Holtslag & Boville PBL scheme, which precluded the creation of its adjoint at this stage. It was shown in Chapter 4 that increasing the depth of the model layer nearest to the surface improved results when using the Louis scheme (e.g. Figure 4.14), and this alteration was therefore made to the adjoint TOMCAT model.

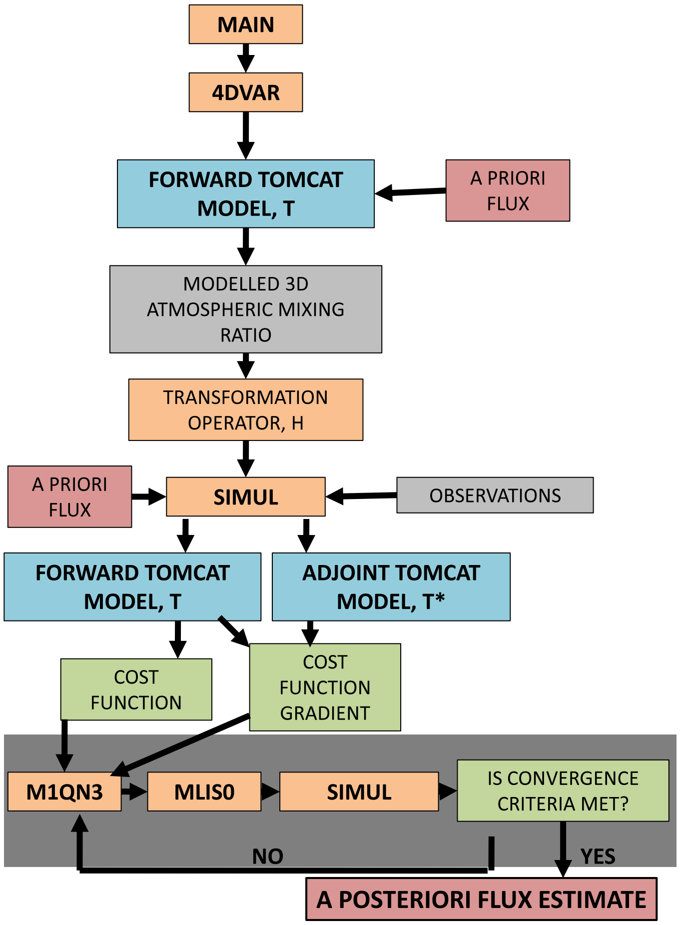
Creation of the adjoint TOMCAT model was carried out by hand instead of using an automatic differentiation program, with each separate adjoint subroutine being coded individually. The slopes advection scheme is made up of three subroutines, named `ADVX1`, `ADVY1` and `ADVZ1`, which carry out zonal, meridional and vertical advection, respectively. Convection and Louis scheme PBL mixing are both carried out in the `CONVEC` subroutine, while the simple chemistry scheme used in the model is part of the `CHIMIE` subroutine. Surface emission is also carried out in `CHIMIE`. In order for the adjoint TOMCAT model to integrate backwards through time, it was also necessary to make changes to initialisation subroutines `INIEXP`, `INICYCL`, `INITER` and output subroutine `FINITER`. Altogether, nine subroutines of the TOMCAT model were required to be rewritten or changed in order to create the adjoint version of the model. Each of these subroutines was found to be completely linear (see Section 5.4.1), so there was no need for a TLM to be created. Adjoint versions of the TOMCAT model and of individual subroutines of the TOMCAT model will henceforth be denoted with the ‘AD_’ prefix, e.g. AD_TOMCAT, `AD_ADVX1`, etc.

5.3.2 The TOMCAT 4D-Var Optimisation Program

The forward and adjoint versions of the TOMCAT model were incorporated into the TOMCAT 4D-Var optimisation program. This program was adapted from a 1D-Var optimisation program provided by F. Chevallier of the Laboratoire des Sciences du Climat et l’Environnement (LSCE), Paris. This program had previously been used for the retrieval of cloud profiles using the RTTOV model (see Chevallier et al. (2002) and Courtier et al. (1994)) for more details). Figure 5.3 shows a flow-chart depicting the 4D-Var optimisation system, which consists of a number of subroutines briefly explained in Table 5.1 and in more detail here.

The `MAIN` outer program initialises all variables in the system, including the error covariance matrices, observations and a priori surface flux estimate, and passes them into `4DVAR`. `MAIN` also finds the eigenvectors and eigenvalues of $\mathbf{R}$, along with its inverse. `4DVAR` then calls the forward TOMCAT model T and the observation operator, H in order to produce $H(T(\mathbf{x}))$, the mapping of the 3D tracer field produced by TOMCAT onto the observations. The program then calls `SIMUL`, which

Figure 5.3 Flow chart depicting the TOMCAT 4D-Var inversion process. Optimisation programs are displayed in orange, while TOMCAT routines are in blue. 2D flux fields are shaded red, atmospheric mixing ratios are light grey and variables relating to the cost function are coloured green. The large dark grey box indicates all the processes which are repeated iteratively until the output is found to fulfil pre-defined criteria. See the main text for further details on each program or variable.



works out the current value of the cost function and its gradient before finally passing these variables through to the minimisation program, `M1QN3`. This program, which is explained in detail in Section 5.3.3, iteratively finds the value of the state vector (the surface flux) which minimises the value of the cost function using the `MLIS0` step-minimisation program which is explained further in Section 5.3.4. After each iteration, `M1QN3` checks the convergence criteria and performs another iteration if the criteria is not yet met. If the solution has converged, the current value of the state vector is output as the a posteriori estimate of the surface flux.

Table 5.1 Details of each subroutine within the TOMCAT 4D-Var optimisation program

SUBROUTINE NAME	**DESCRIPTION**
`MAIN`	• Initialises all variables • Sets up error covariance matrices **R** and **B** , finds $\mathbf{R}^{-1}$ and finds eigenvalues and eigenvectors of **R** • Reads observations and a priori flux estimate
`4DVAR`	• Calls forward TOMCAT model with a priori fluxes to produce 3D atmospheric mixing ratio field • Calls transformation operator to map 3D field onto observations • Calls `SIMUL` to evaluate cost function and cost function gradient
Forward TOMCAT model, T	• Produces 3D atmospheric mixing ratio field from supplied flux field
Transformation operator, H	• Maps 3D tracer field onto observations
`SIMUL`	• Finds value of cost function • Calls AD_TOMCAT • Finds gradient of cost function with respect to flux field
AD_TOMCAT, T^*	• Finds 2D flux sensitivity field to model - observation anomalies
`M1QN3`	• Finds optimal descent direction for minimisation of cost function • Calls `MLISO` to find minimum along this descent direction only • Calls `SIMUL` to evaluate new cost function and cost function gradient • Checks convergence criteria
`MLISO`	• Finds minimum of cost function along chosen descent direction using step-minimisation process

5.3.3 The `M1QN3` Minimisation Program

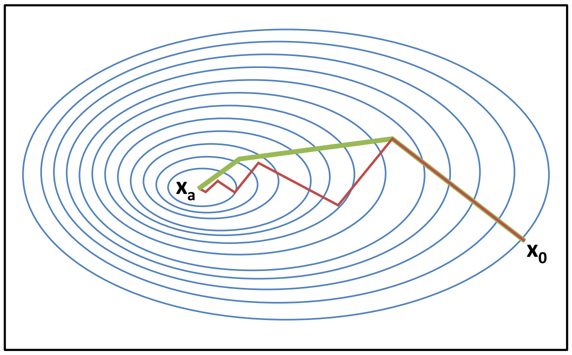
The `M1QN3` minimisation procedure uses a limited memory quasi-Newtonian (QN) method and was developed at Institut National de Recherche en Informatique et en Automatique (INRIA). The procedure will be briefly outlined here, but for more details see Gilbert and Lemarechal (1989). The `M1QN3` routine solves unconstrained minimisation problems of the form;

$$min\{J(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^m\} \tag{5.8}$$

where the function J is supposed continuously differentiable. The method is based upon an algorithm proposed by Nocedal (1980). QN methods are used in optimisation programming as more practical versions of conjugate gradient (CG) methods, which solve minimisation problems through iterative descent using exact line searches. CG methods produce faster convergence than the simple steepest descent (SD) method due to the fact that the descent direction at each iteration in a CG method is chosen to be orthogonal to each of those used previously (Daniel, 1967). Figure 5.4 shows the difference between the SD and CG methods of descent. The SD method requires many steps to converge to the minimum due to the fact that each subsequent descent direction must be perpendicular to the last. The CG method has no such limitations and generally converges within m steps (Shewchuck, 1994). The CG method uses only a small amount of computer memory to find the descent direction, but requires an exact line search in order to find the minimum along each descent vector, which is computationally expensive. QN methods are similar to CG methods but have the converse features, with no need for exact line searches but a large memory requirement (Gilbert and Lemarechal, 1989). QN methods use the value of the Hessian of the function, $\nabla^2 J(x)$, to find the next descent direction, and it has been shown that the QN and CG methods are equivalent if an exact line search is used in each (Nazareth, 1979). If an inexact line search is used, however, the QN produces a faster rate of converge than CG methods (Gilbert and Lemarechal, 1989). The QN method is therefore chosen for use here and details of the method used to find the descent vector is described later in this section.

The `M1QN3` subroutine is first called once the initial value and gradient of the cost function have been found. It is then repeated iteratively until the cost function or its gradient have met the pre-defined convergence criteria, when the program returns the a posteriori state vector. At each iteration k, $k \geq 1$, the program determines a descent direction, $\mathbf{d}_k$, of J at $\mathbf{x}_k$. Newton's equation states that this descent direction has the form $\mathbf{d}_k = -\mathbf{W}_k \mathbf{g}_k$, where $\mathbf{W}_k$ represents the current value of the inverse Hessian, $[\nabla^2 J(\mathbf{x}_k)]^{-1}$, of J at $\mathbf{x}_k$, and $\mathbf{g}_k$ is the gradient of the cost function at $\mathbf{x}_k$. Once this descent direction is chosen, the step-size, α_k, to be taken along this direction is determined by the by the line-search procedure `MLIS0`, described in Section 5.3.4. At the next iteration the state vector has the form $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$.

Figure 5.4 Results of using the steepest descent method (*red lines*) and the conjugate gradient method (*green lines*) in order to find the minimum of a function of which lines of constant value are shown in blue. The conjugate gradient method produces faster convergence than the steepest descent.



This QN method reduces the necessary amount of computational memory by dealing only with the local Hessian — at each step $\mathbf{W}_k$ is found using only

$$\{(\mathbf{y}_i, \mathbf{s}_i) : k - m \leq i \leq k - 1\} \tag{5.9}$$

where m is chosen according to the available computer memory. An approximation of the value of $\nabla^2 J(\mathbf{x}_k)$ is obtained by updating a diagonal matrix $\mathbf{D}_k$ using the inverse of the Broyden - Fletcher - Goldfarb - Shanno (BFGS) formula, given in Equation 5.13, which uses the vectors $\mathbf{s}_i = \mathbf{x}_{i+1} - \mathbf{x}_i$ and $\mathbf{y}_i = \mathbf{g}_{i+1} - \mathbf{g}_i$. $\mathbf{D}_0$ is initialised as $\delta_1 \mathbf{I}$ where $\mathbf{I}$ represents the identity matrix and

$$\delta_k = \frac{\langle \mathbf{y}_k, \mathbf{s}_k \rangle}{||\mathbf{y}_k||^2} \tag{5.10}$$

where the inner product, $\langle\ ,\ \rangle$ is defined in the Euclidean sense as the dot product of the two vectors, i.e;

$$\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T \mathbf{b} = \sum_{i=1}^{n} a_i b_i \tag{5.11}$$

and the vector norm $||\quad||$ is defined such that;

$$||\mathbf{a}|| = \langle \mathbf{a}, \mathbf{a} \rangle^{\frac{1}{2}} \tag{5.12}$$

The inverse BFGS formula is as follows;

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \frac{(\mathbf{s}_k - \mathbf{W}_k\mathbf{y}_k)\otimes\mathbf{s}_k + \mathbf{s}_k\otimes(\mathbf{s}_k - \mathbf{W}_k\mathbf{y}_k)}{\langle\mathbf{y}_k,\mathbf{s}_k\rangle} - \frac{\langle\mathbf{s}_k - \mathbf{W}_k\mathbf{y}_k,\mathbf{y}_k\rangle}{\langle\mathbf{y}_k,\mathbf{s}_k\rangle^2}\mathbf{s}_k\otimes\mathbf{s}_k \tag{5.13}$$

where the tensor product $\otimes$ is defined such that;

$$[a\otimes b]_{ij} = (ba^T)_{ij} = b_i a_j \tag{5.14}$$

The matrix $\mathbf{W}_k$ is not stored in the computer memory, but instead the product $\mathbf{W}_k\mathbf{g}_k$ is formed by an efficient algorithm. This choice of $\mathbf{W}_k$ ensures that the Hessian is positive definite, ensuring that $\mathbf{d}_k$ is always a descent direction.

5.3.4 The `MLISO` Step-Minimisation Process

As discussed in Section 5.3.3, at each iteration of `M1QN3`, the state vector is updated to be $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{d}_k$, where $\mathbf{d}_k$ is the descent direction and α_k is the step length along this descent direction necessary to minimise the cost function along $\mathbf{d}_k$. α_k is given by the `MLISO` line search algorithm, which computes α_k though iterative testing of the Wolfe condtions, which are as follows;

$$J(\mathbf{x}_k + \alpha_k\mathbf{d}_k) \leq J(\mathbf{x}_k) + \omega_1\langle\mathbf{g}_k,\mathbf{d}_k\rangle \tag{5.15}$$

$$\langle\mathbf{g}(\mathbf{x}_k + \alpha_k\mathbf{d}_k),\mathbf{d}_k\rangle \leq \omega_2\langle\mathbf{g}_k,\mathbf{d}_k\rangle \tag{5.16}$$

where it is necessary to have $0 < \omega_1 < \frac{1}{2}$ and $\omega_1 < \omega_2 < 1$. For this study, $\omega_1 = 0.0001$ and $\omega_2 = 0.9$. `MLISO` iteratively finds the maximum value of α_k for which Equations 5.15 and 5.16 hold.

5.4 Validation of the Adjoint TOMCAT Model

The quality of results produced using the 4D-Var method is highly dependent upon the accuracy of the adjoint model used in the inversion process. It is therefore important to thoroughly test the adjoint model once it has been written to make sure that the transport produced by the adjoint is an accurate representation of the transpose of the forward model. This section will provide details of a number of tests performed on AD_TOMCAT in order to validate its use in the TOMCAT 4D-Var inversion process.

5.4.1 Confirming the Linearity of the Forward Model

Each atmospheric transport scheme chosen for use in the adjoint version of the TOMCAT model is linear, which means that the ADM can be directly inferred from the forward model without the need for an intermediate TLM. However, it is necessary to confirm, in a numerical sense, the linearity of the forward model before production of the adjoint model is begun, as any unseen non-linearities in the forward model would lead to errors in the adjoint model. For each individual subroutine of the TOMCAT model intended for use in the adjoint it must be confirmed that;

$$\forall\lambda, \quad \lambda T'(\delta\mathbf{x}) = T'(\lambda\delta\mathbf{x}) \tag{5.17}$$

where λ is a scalar quantity. The result shown in Equation 5.17 was verified individually for the subroutines `ADVX1`, `ADVY1`, `ADVZ1` and `CONVEC`, and also for the full TOMCAT model. The input variables, $\mathbf{X}_{IN}$, for each subroutine consist of `S0`, the mass of tracer in each model grid box, and the first-order moments of `S0` as described in Chapter 3 labelled `SX`, `SY` and `SZ`. The output variables, $\mathbf{X}_{OUT}$, are the updated values of $\mathbf{X}_{IN}$. One iteration of each subroutine was carried out in turn, in which each variable of $\mathbf{X}_{IN}$ was defined as a normally distributed random array. The linearity identity shown in Equation 5.17 was checked using ten different random initialisations and for $\lambda = 1\times 10^n$, where $n = -10, 10$. For every subroutine, including the full TOMCAT model, the value of $\frac{\lambda T'(\mathbf{X}_{IN})}{T'(\lambda\mathbf{X}_{IN})}$ was equal to 1.0 within 8 decimal places, the accuracy of the machine used to carry out the test, irrespective of the value of λ or the initialisation of $\mathbf{X}_{IN}$. This indicated that the TOMCAT model is perfectly linear, and there was no need to create a TLM.

5.4.2 Numerical Validation of the Adjoint Model

After ensuring the linearity of the forward TOMCAT model, the adjoint version of each transport subroutine was produced. Due to the fact that the sensitivities in the ADM are dependent upon the trajectory of the forward model at each time step, a new version of the TOMCAT model was also developed which saved the necessary trajectories at every iteration so that they may be found later by the adjoint model. Since both of these models were coded by hand, there was a possibility of errors being introduced, whether due to mistakes in the transposition of the forward code or merely typographical errors. In order to ensure the accuracy of the ADM, it was verified that each adjoint subroutine satisfied the adjoint identity given in Equation 5.7, i.e. that;

$$\forall \mathbf{x}, \forall \mathbf{y} \quad \frac{\langle T'\mathbf{x}, \mathbf{y}\rangle}{\langle \mathbf{x}, T^{*}\mathbf{y}\rangle} = 1 \tag{5.18}$$

Equation 5.18 was tested for the subroutines `ADVX1`, `ADVY1`, `ADVZ1`, `CONVEC` and for the full TOMCAT model. The input and output variables for the forward subroutines, $\mathbf{X}_{IN}$, were defined as normally distributed random variables as in Section 5.4.1, and in the input variables for the adjoint subroutines were defined to be equal to the output from the corresponding forward subroutine, $\mathbf{X}_{OUT}$. The identity in Equation 5.18 therefore becomes;

$$\frac{||\mathbf{X}_{OUT}||^{2}}{\langle \mathbf{X}_{IN}, M^{*}(\mathbf{X}_{OUT})\rangle} = 1 \tag{5.19}$$

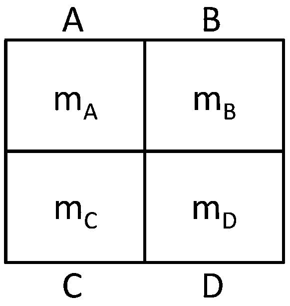
Equation 5.19 was tested for ten different random initialisations for one iteration of each subroutine. For each subroutine and each initialisation, the identity given in Equation 5.19 holds up to the accuracy of the machine used to carry out the tests, strongly indicating that the adjoint model has been accurately coded from the forward model. It is not, of course, possible to check that the identity holds for all input vectors and so the test of the adjoint model is not exhaustive, but the level of accuracy of the results of the tests implies that the adjoint model is likely to be correct. Additional tests of the adjoint model were therefore carried out in order to further examine the model's accuracy, which will be discussed in Sections 5.4.3 and 5.4.4.

5.4.3 Sensitivity conservation in the Adjoint Model

Measurements of sensitivity in the adjoint TOMCAT model are analogous to mass in the forward model. Therefore, if total mass is conserved in the forward model then, in certain cases, the sum of all the sensitivities within the adjoint model must also be conserved. In order to investigate the conservation of sensitivity in the ADM, it is first necessary to fully understand the nature of adjoint transport. This is most clearly illustrated with a simple example in which transport is investigated in a 2D model with four grid boxes, as shown in Figure 5.5.

In the forward model, which transports tracer mass around the model grid, the total tracer mass in the model is conserved over time. The grid boxes, labelled A, B, C and D each contain a mass of tracer at time t_0 labelled $m_A^{t_0}$, $m_B^{t_0}$, $m_C^{t_0}$ and $m_D^{t_0}$ respectively. Similarly, if the forward model is integrated over the period [t_0, t_n], then the mass in each grid box at time t_n is $m_A^{t_n}$, $m_B^{t_n}$, $m_C^{t_n}$ and $m_D^{t_n}$. The model transport over this time period can be seen as a 4×4 matrix which distributes the original tracer mass in each grid box into each of the four grid-boxes, i.e.;

Figure 5.5 Example of the model grid for a simple four-box transport model used to illustrate adjoint transport properties in Sections 5.4.3 and 5.4.4.



$$\begin{pmatrix} m_A^{t_n} \\ m_B^{t_n} \\ m_C^{t_n} \\ m_D^{t_n} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} m_A^{t_0} \\ m_B^{t_0} \\ m_C^{t_0} \\ m_D^{t_0} \end{pmatrix} \tag{5.20}$$

where, assuming no mass is created or destroyed;

$$\sum_{i=1}^{4} a_{ij} = 1 \quad 1 \leq j \leq 4 \tag{5.21}$$

In this case, the total mass in the model at time t_n is equal to;

$$\begin{aligned} m_A^{t_n} + m_B^{t_n} + m_C^{t_n} + m_D^{t_n} &= (a_{11} + a_{21} + a_{31} + a_{41})\, m_A^{t_0} \\ &+ (a_{12} + a_{22} + a_{32} + a_{42})\, m_B^{t_0} \\ &+ (a_{13} + a_{23} + a_{33} + a_{43})\, m_C^{t_0} \\ &+ (a_{14} + a_{24} + a_{34} + a_{44})\, m_D^{t_0} \\ \\ &= m_A^{t_0} + m_B^{t_0} + m_C^{t_0} + m_D^{t_0} \end{aligned} \tag{5.22}$$

and therefore, the total mass is conserved.

The adjoint model, however, does not transport mass around the model grid, but measures sensitivities between grid boxes. Total model sensitivity is not necessarily conserved, and is in fact unlikely to be conserved except in certain circumstances. To provide an example, the simplified four-box model described previously is integrated over the time period $[t_n, t_0]$ and initialised with values $s_X^{t_n}$, where $X = A, D$ and $s_X^{t_n}$ is initialised to be equal to $m_X^{t_0}$, the initial mass placed in the forward model in the previous example. The adjoint transport is defined as the transpose of the forward transport, which means that;

$$\begin{pmatrix} s_A^{t_0} \\ s_B^{t_0} \\ s_C^{t_0} \\ s_D^{t_0} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} & a_{31} & a_{41} \\ a_{12} & a_{22} & a_{32} & a_{42} \\ a_{13} & a_{23} & a_{33} & a_{43} \\ a_{14} & a_{24} & a_{34} & a_{44} \end{pmatrix} \begin{pmatrix} s_A^{t_n} \\ s_B^{t_n} \\ s_C^{t_n} \\ s_D^{t_n} \end{pmatrix} \tag{5.23}$$

where $s_X^{t_0}$ is the sensitivity in grid box X at time t_0. The total sensitivity in the model at the end of the adjoint simulation is therefore equal to;

$$\begin{aligned} s_A^{t_0} + s_B^{t_0} + s_C^{t_0} + s_D^{t_0} &= (a_{11} + a_{12} + a_{13} + a_{14})\, s_A^{t_n} \\ &+ (a_{21} + a_{22} + a_{23} + a_{24})\, s_B^{t_n} \\ &+ (a_{31} + a_{32} + a_{33} + a_{34})\, s_C^{t_n} \\ &+ (a_{41} + a_{42} + a_{43} + a_{44})\, s_D^{t_n} \end{aligned} \tag{5.24}$$

which indicates that total model sensitivity is only conserved either if

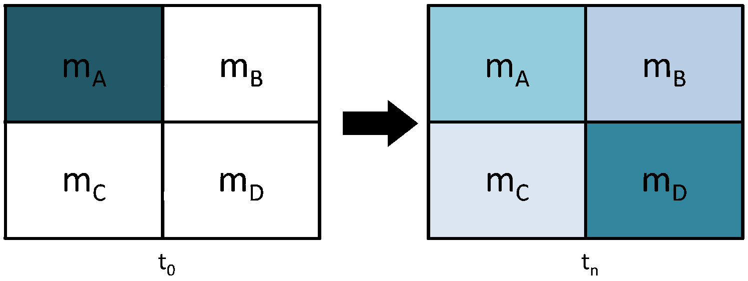
$$\sum_{j=1}^{4} a_{ij} = 1, \qquad 1 \leq i \leq 4 \quad (CASE\ 1) \tag{5.25}$$

or if

$$s_A^{t_n} = s_B^{t_n} = s_C^{t_n} = s_D^{t_n} \quad (CASE\ 2) \tag{5.26}$$

Case 1 is true only if the forward model evenly distributes the tracer among the model grid boxes (the homogeneous transport case), while Case 2 is true if the model is initialised with the same tracer mass in every grid box at the start of the adjoint simulation (the homogeneous initialisation case). In order to test sensitivity conservation of the AD_TOMCAT model, the homogeneous

Figure 5.6 Forward model transport of a tracer mass in simplified model initialised with mass only in grid box *A* time t_0 (*Left*). At time t_n, (*Right*), the tracer is mixed throughout all model grid boxes. The tracer mass in a grid box is indicated by its colour. Darker grid boxes contain more tracer.



initialisation test was used, since the homogeneous transport case does not examine physically realistic atmospheric transport. An adjoint simulation was carried out in which each grid box was initialised with a value of 1 and the model was integrated over one month. The total sensitivity value remained perfectly constant throughout the simulation, implying that the transport in the ADM is consistent with that of the forward model.

5.4.4 Reciprocity of Adjoint Transport

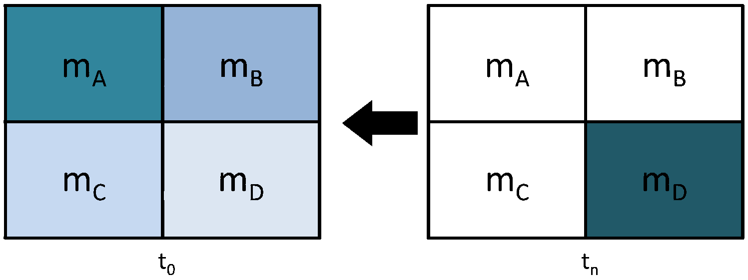
In order to further test the accuracy of the adjoint transport in AD_TOMCAT, the property of reciprocity was investigated. It has been previously discussed (e.g. Hourdin and Talagrand (2006a,b)) that for a linear model, transport in the adjoint model is reciprocal to transport in the forward model. This property is best illustrated using the simplified model described in Section 5.4.3. Consider a simulation using the forward model in which a mass $m_A^{t_0}$ is placed into grid box *A* at time t_0. All other grid boxes are initialised with zero mass. If the tracer transport is represented in matrix form as in Equation 5.20, then the tracer mass in, say, grid box *D* at time t_n is

$$m_D^{t_n} = a_{41}\, m_A^{t_0} \tag{5.27}$$

This process is illustrated in Figure 5.6. Now, if the adjoint model is integrated over the period $[t_n, t_0]$ with values which are zero in each grid box apart from grid box *D*, in which $s_D^{t_n}=m_A^{t_0}$, then it holds that;

$$s_A^{t_0} = a_{41}\, s_D^{t_n} \tag{5.28}$$

Figure 5.7 As Figure 5.6 but for adjoint transport. The simplified adjoint model is initialised with a sensitivity only in grid box D at time t_n (*Right*) and the model is integrated back to time t_0 (*Left*). Note that the sensitivity in grid box A at time t_0 is equal to the tracer mass in grid box D at time t_n in Figure 5.6.



and therefore;

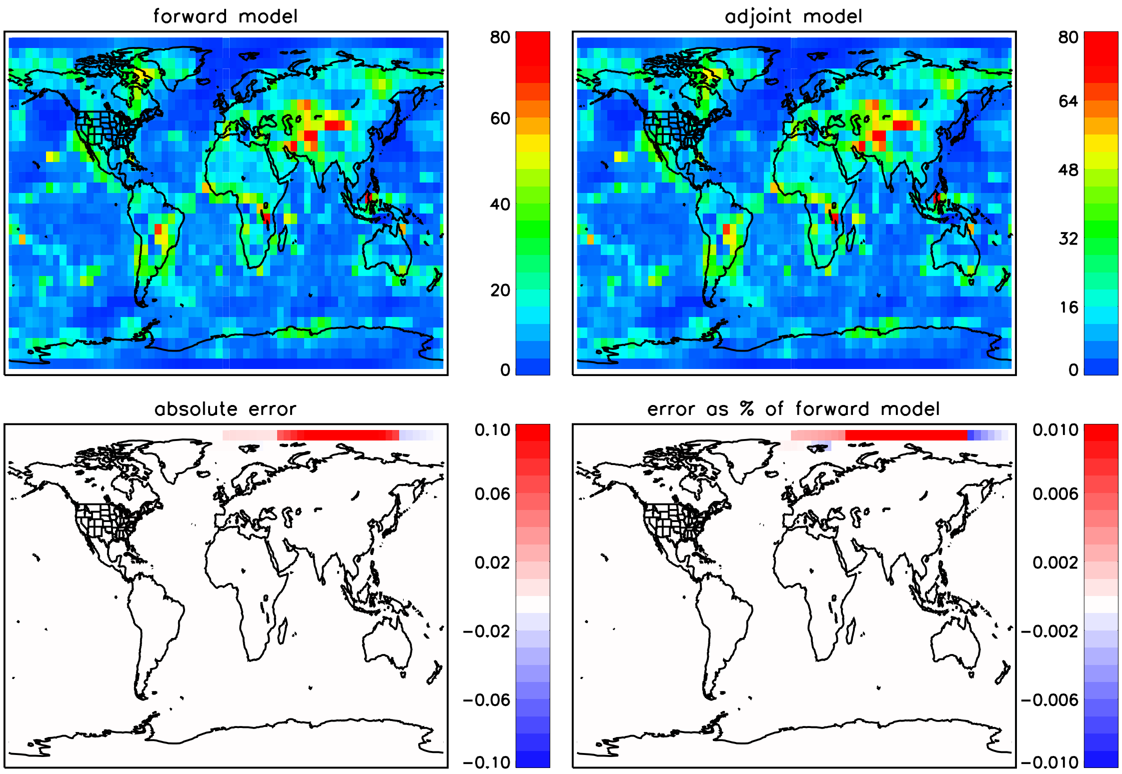
$$s_A^{t_0} = m_D^{t_n} \tag{5.29}$$

i.e. the transport between the two grid boxes is reciprocal. This process is illustrated in Figure 5.7. For more information on this result and its use in inverse modelling, see Hourdin and Talagrand (2006b). This result implies that the accuracy of the adjoint transport in the AD_TOMCAT model may be tested by examining the reciprocity of its transport.

Due to the high computational burden of adjoint modelling and the increased simulation time required to carry out both forward and adjoint simulations, the reciprocity of tracer transport in the AD_TOMCAT model was examined on two different time scales. The adjoint transport over one day was examined from every surface grid box, while longer simulations were carried out which investigated adjoint transport from selected grid boxes only.

In order to test the short-term reciprocity of the ADM transport, separate forward simulations were carried out in which one surface grid box, S, was initialised with an (arbitrary) tracer mass of 100, with zero mass elsewhere. One separate simulation was performed for each surface grid box. After the simulation period of one day was complete, the location D and value m_D of the maximum tracer concentration in each simulation was noted. Figure 5.8(a) shows the value of the maximum tracer concentration associated with each single grid box simulation, which for display purposes has been placed at S rather than D, irrespective of the maximum's actual location. Following this, separate adjoint simulations were carried out in which a value of 100 was placed into each box D and the ADM was integrated backwards over the day. Figure 5.8(b) shows the sensitivity value m_S in each S at the end of the adjoint simulation. Figure 5.8(c) shows the difference between m_D and m_S, indicating that at almost every grid box the short-term reciprocity of the adjoint transport

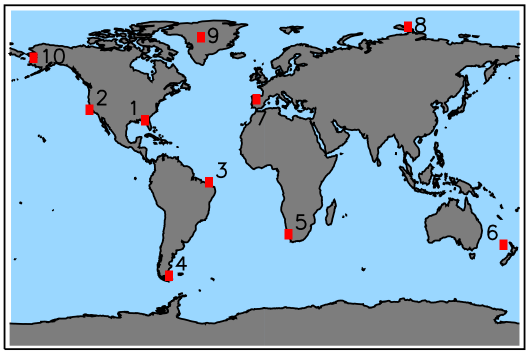
Figure 5.8 (*Top Left*) Value of m_D for each model surface grid box at the end individual forward model simulations in which one grid box had been initialised with a tracer mass of 100, with all others having zero tracer mass, and integrated over one day. For clarity, the value of m_D is displayed in the initial grid box S. (*Top Right*) Value of m_S for each model surface grid box at the end of corresponding adjoint model simulations. Here, the value of m_S is displayed in the initial grid box S. m_D and m_S should be equal if the adjoint transport is correct. (*Bottom Left*) Difference between m_D and m_s. (*Bottom Right*) Difference between m_D and m_s as a percentage of m_D.



is correct. There is a very small difference between m_S and m_D at some high polar latitude grid boxes, which is likely to be due to the fact that the TOMCAT model groups high-latitude grid boxes together during the zonal transport, as discussed in Chapter 3. The error at these grid boxes is less than 0.01% of the expected value, however, so over short time periods at least, this issue is insignificant.

In order to test the reciprocity of adjoint transport in the AD_TOMCAT model over a longer time period, simulations were carried out in which the reciprocity experiment was repeated over a time period of one month instead of one day, for ten surface grid boxes only. This test was carried out only at certain locations due to the computational burden which a more large-scale test would necessitate. The initial locations, S^n, $1 \leq n \leq 10$ are shown in Figure 5.9, and details of the sites are given in Table 5.2. Again, at the end of the forward simulation the grid box D^n with the largest mass of tracer m_D^n was found and chosen to be the initial grid box for an adjoint simulation also

Figure 5.9 Locations of S^n, the initial locations of tracer release in the forward model simulations used for reciprocity testing over one month.



of length one month. Figure 5.10 shows an example of the forward and adjoint mixing over one month for the FLO tracer, which is emitted from Florida (marked S in the figure) at time t_0 in the forward simulation. Figure 5.10(a) shows the mass of tracer on one model level after one month, with the location of the maximum tracer mass, marked D, at the north coast of Africa. The tracer has been transported vertically, and the maximum tracer concentration is at a height of approximately 1300m. The FLO tracer in the adjoint simulation was therefore initialised at location D, and the distribution of the tracer sensitivity at the beginning of the month is shown in Figure 5.10(b). This figure shows that a tracer at location D is most sensitive to areas in the Caribbean and the Eastern USA, including Florida. Table 5.2 shows the values of m_D^n and m_S^n for each site. The results show that even after a month, the tracer transport in the AD_TOMCAT model is consistent with that of the forward model, with m_D^n and m_S^n within 2% of each other at every site. When it is considered that each initial grid box was initialised with a value of 100, the fact that each of the final values is within 0.0015 of its reciprocal suggests that over a long time period the adjoint model is representative of the forward model transport to a high level of accuracy.

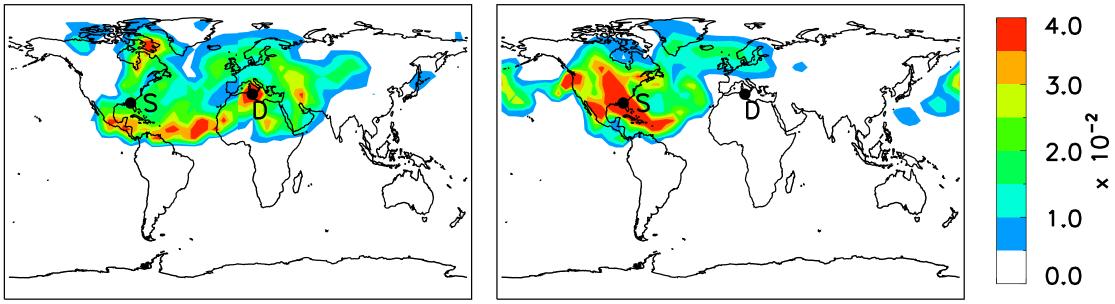
5.5 Validation of the TOMCAT 4D-Var System

In order to validate the TOMCAT 4D-Var system, a number of tests were carried out in which the assimilated observations were replaced with modelled fields produced with known flux fields. This provides an indication of the ability of the system to optimise the state vector to a high level of accuracy and also gives an indication of its efficiency in doing so. The fact that the synthetic ‘observations’ provided by the model are produced using a known flux field allows quantification of the accuracy of the optimised a posteriori fluxes from the resulting inversion. This also provides some indication of the model error, which is necessary when compiling the error covariance matrix $\mathbf{R}$.

Table 5.2 Results of the one-month reciprocity tests, showing the locations $S^n$of the original tracer release, m^n_D and m^n_S and the difference between m^n_D and m^n_S as a percentage of m^n_D. The values of longitude and latitude indicate the centre of the corresponding grid box.

SITE, *n*	**NAME**	**LONGITUDE (°)**	**LATITUDE (°)**	m^n_D $(\times 10^{-2})$	m^n_S $(\times 10^{-2})$	% diff
1	FLO	84.4E	30.5N	6.544	6.616	+1.10
2	CAL	123.8E	36.0N	6.800	6.844	+0.65
3	BRA	39.4E	2.8S	5.869	5.873	+0.068
4	TDF	67.5E	52.6S	2.263	2.259	-0.177
5	SAF	16.9W	30.5S	4.757	4.756	-0.021
6	NZE	168.8W	36.0S	1.426	1.433	+0.491
7	ESP	5.6E	41.6N	7.368	7.482	+1.547
8	RUS	101.3W	80.3N	8.685	8.645	-0.461
9	GNL	45.4E	74.8N	7.693	7.742	+0.637
10	BRW	163.1E	63.7N	9.277	9.233	-0.474

Figure 5.10 (*Left*) Tracer mass at approximately 1300m after one month forward model simulation initialised with a tracer mass of 100 at the surface at location *S*. *D* represents the location of maximum tracer mass. (*Right*) Tracer sensitivity at the surface after one month adjoint model simulation initialised with a sensitivity of 100 approximately 1300m above location *D*.



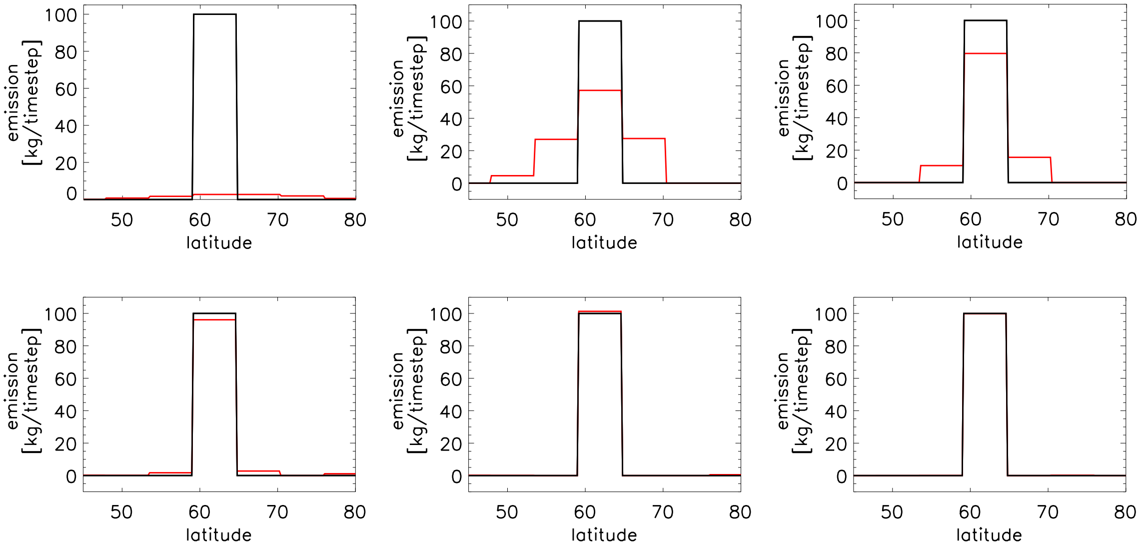
A simple experiment was first carried out in which the synthesised observations were produced from an idealised forward model simulation in which the fluxes, $\mathbf{x}_{true}$, were initialised to be zero everywhere except in model grid boxes centred at a longitude of 62°E, where they were set to be 100 kg grid-box^{-1} timestep^{-1}. Advection was not carried out using ECMWF analyses, with a solid-body rotation (SBR) advection system being used instead. This replaces the analysed horizontal mass flux U with

$$U = \frac{2R_e\pi}{5 \times 86400} \tag{5.30}$$

where R_e is the radius of the Earth. This removes all meridional transport and produces purely zonal advection which transports an air mass around the globe in five days. The forward simulation was run for six hours only, to minimise the impact of model error on the result, the 3D concentration field produced at the end of the simulation then represented the observations, $\mathbf{y}$. The TOMCAT 4D-Var system was then used to attempt to retrieve the original flux band $\mathbf{x}_{true}$ from an a priori, $\mathbf{x}_a$, containing no information about the emission field (in which each grid box is set as zero). The initial 3D model field was also initialised as zero everywhere. The observation error covariance matrix $\mathbf{R}$ was diagonal, with each element along the diagonal being initialised as 1.0×10^{-10} ppb. This small value was partly reflective of the small mixing ratio produced by the forward model and partly in order to help constrain the solution towards the observations rather than the a priori. The a priori error covariance matrix $\mathbf{B}$ was also diagonal, with elements set to be 200% of $\mathbf{x}_{true}$. The optimisation process was carried out until 40 iterations had been completed. Figure 5.11 shows the gradual retrieval of the true fluxes in one dimension at a latitude of 2.8°N, while Figure 5.12 shows the gradual retrieval of the the full 2D flux field. The two figures show that after one iteration of the optimisation process, only a small change to the a priori has been produced, although the approximate location of the flux has been found by the inversion. After five iterations, however, there has been a large improvement to the a priori field, with fluxes of 50-80 kg grid-box^{-1} timestep^{-1} estimated at the correct location, albeit with some noise in adjacent grid boxes. The 2D retrieval displays some variation with latitude, which is likely to be due to factors such as the varying orography and dimensions of the surface grid boxes. After 10 iterations the solution is close to being correct, with the updated flux field estimating 75 - 95 kg grid-box^{-1} timestep^{-1} in the correct location, with a much smaller 'halo' of emissions from surrounding grid boxes. After 20 iterations, the solution has converged to within 2% everywhere, while after 30 and 40 iterations, the solution is refined, with the solution after 40 iterations almost indistinguishable from $\mathbf{x}_{true}$.

Figure 5.13 shows the value of the cost function J and the norm of its gradient $||\mathbf{g}||$ throughout the optimisation process. The value of the cost function reduces by many orders of magnitude as the iterative process continues. Initially, the value of the cost function is approximately 1.0×10^{16}, and is gradually reduced to a value of approximately 3.0×10^{10}. Given the small values in the observation error covariance matrix ($R_{ij}^{-1} = 1.0 \times 10^{10}$), this shows that there is an extremely small difference between the true fluxes and the a posteriori estimate. The norm of the cost function gradient is initially around 1.0×10^{30}, which is extremely large compared to the value of J, indicating that the a priori estimate is far from correct. $||\mathbf{g}||$ decreases dramatically after the first optimisation iteration to approximately 1.0×10^{14}, which is a similar order of magnitude to the value of J. The TOMCAT 4D-Var system performed well in retrieving $\mathbf{x}_{true}$ to a high level of accuracy, reaching a near perfect solution within 40 iterations, although the solution was correct to within 2% after 20 iterations.

Figure 5.11 Retrieval (red line) of $\mathbf{x}_{true}$ of 100 kg grid box^{-1} timestep^{-1} (black line) starting from an a priori of zero after *(Top Left)* 1 iteration, *(Top Centre)* 5 iterations, *(Top Right)* 10 iterations, *(Bottom Left)* 20 iterations, *(Bottom Centre)* 30 iterations, *(Bottom Right)* 40 iterations of the TOMCAT 4D-Var retrieval system



A second experiment was carried out which was identical to the previous retrieval experiment aside from the fact that an a priori flux field $\mathbf{x}_a$ was used which had a flux band in the wrong location, rather than no prior estimate. An incorrect emission band was placed in grid boxes centred at 285°E. This experiment investigated the optimisation program's ability to function given an incorrect starting point. Figures 5.14 and 5.15 show the gradual retrieval of $\mathbf{x}_{true}$ from $\mathbf{x}_a$ in one dimension, at 2.8°N, and two dimensions, respectively. The retrieval process is similar to that produced when no prior information was given, and the incorrect starting point does not appear to hamper the inversion. Again, there is little change after one iteration, but after five iterations the are fluxes of approximately 50 kg grid-box^{-1} timestep^{-1} in the correct location, while the incorrect fluxes have fallen to a similar flux strength. After ten iterations the retrieval is approximately 75% complete, and after 20 iterations the solution is close to being correct. As before, running the inversion to 30 and 40 iterations refines the result. Figure 5.16 shows the values of J and of $||\mathbf{g}||$ for the retrieval, which both display similar statistics to those produced in the previous experiment. The value of J falls from around 2.0×10^{16} to 4.0×10^{10}, while the norm of the gradient falls immediately to a value close to that of J and decreases similarly.

5.6 4D-Var Retrievals with Simulated Data

The value of the cost function at the end of an assimilation provides a valuable diagnostic of the system. As discussed in Dee (1995), if the model used in the inversion is perfect, and all other

Figure 5.12 As Figure 5.11, but in two-dimensions. *(Left)* $\mathbf{x}_{true}$, *(Centre)* retrieval $\mathbf{x}$ of TOMCAT 4D-Var system and *(Right)* $(\mathbf{x}_{true} - \mathbf{x})$ for 1, 5, 10, 20, 30 and 40 iterations.

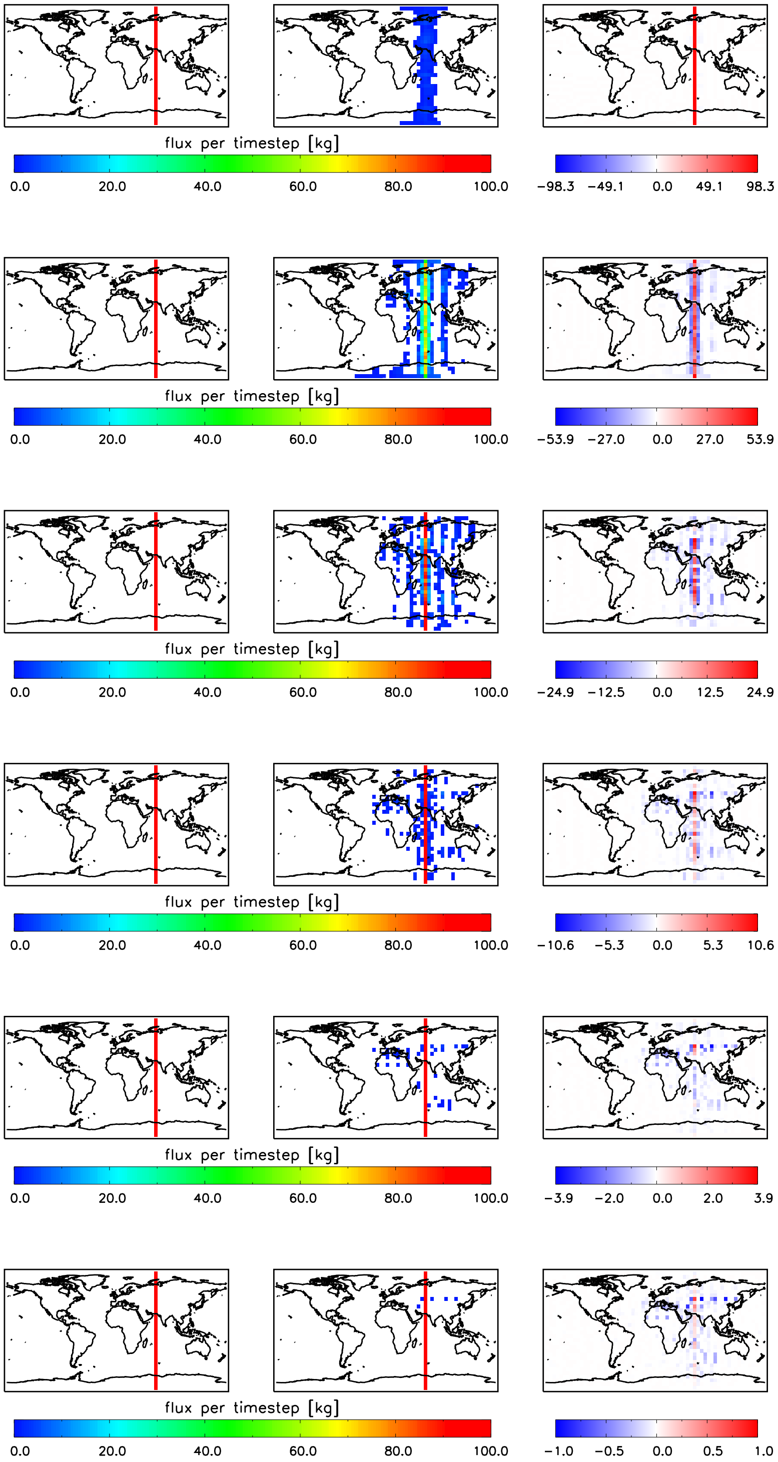


Figure 5.13 *(Left)* Value of the cost function and *(Right)* value of the norm of the cost function gradient for the retrieval displayed in Figures 5.11 and 5.12.

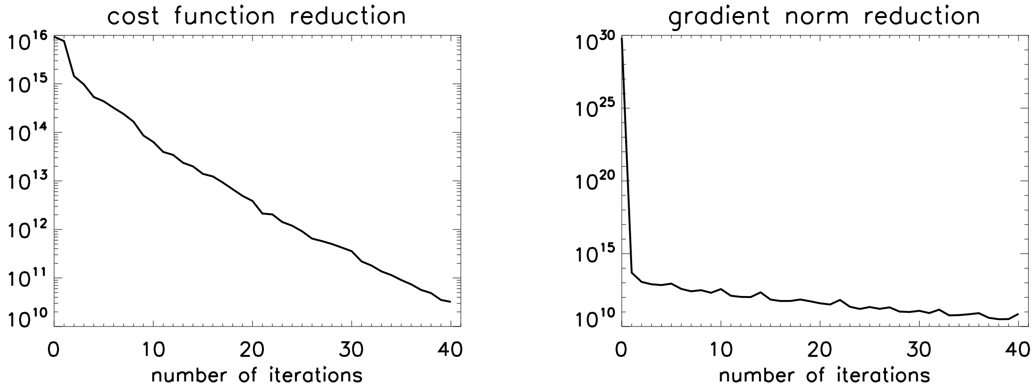
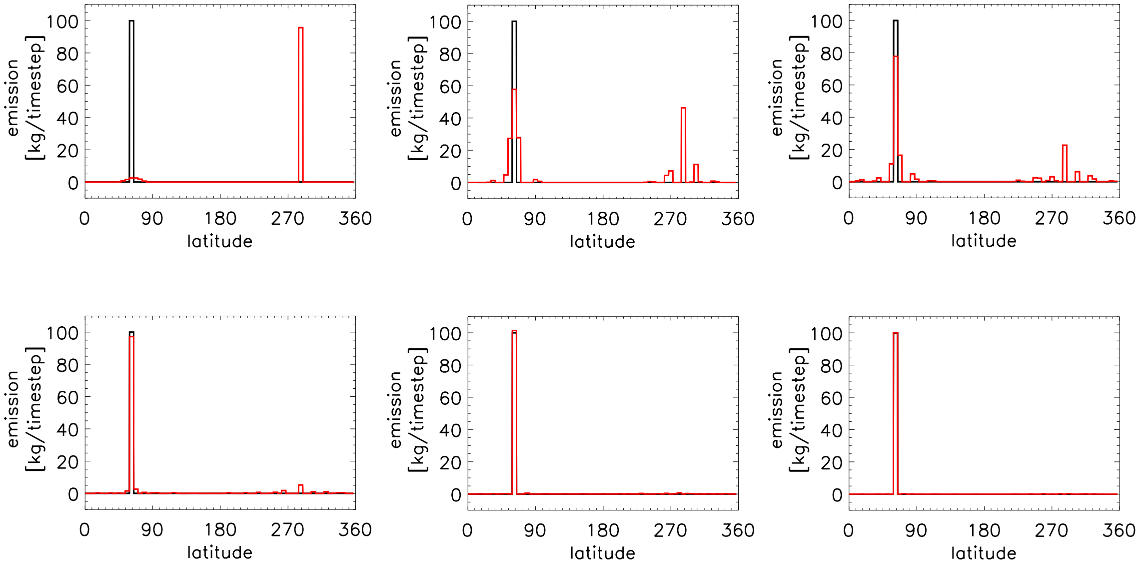


Figure 5.14 Retrieval (red line) of $\mathbf{x}_{true}$ of 100 kg grid box^{-1} timestep^{-1} (black line) starting from an incorrect a priori after *(Top Left)* 1 iteration, *(Top Centre)* 5 iterations, *(Top Right)* 10 iterations, *(Bottom Left)* 20 iterations, *(Bottom Centre)* 30 iterations, *(Bottom Right)* 40 iterations of the TOMCAT 4D-Var retrieval system



errors (in the observations and a priori) are Gaussian and independent, then the expected value of of the observation term of the cost function at the end of the minimisation, known as the residual, r, should be equal to twice the number of observations n_{obs} (since the cost function is multiplied by a factor of $\frac{1}{2}$). Therefore;

$$\frac{2r_{min}}{n_{obs}} = 1 \tag{5.31}$$

However, if model error is present (and also Gaussian) then expression 5.31 becomes;

Figure 5.15 As Figure 5.14, but in two-dimensions. *(Left)* $\mathbf{x}_{true}$, *(Centre)* retrieval $\mathbf{x}$ of TOMCAT 4D-Var system and *(Right)* $(\mathbf{x}_{true} - \mathbf{x})$ for 1, 5, 10, 20, 30 and 40 iterations.

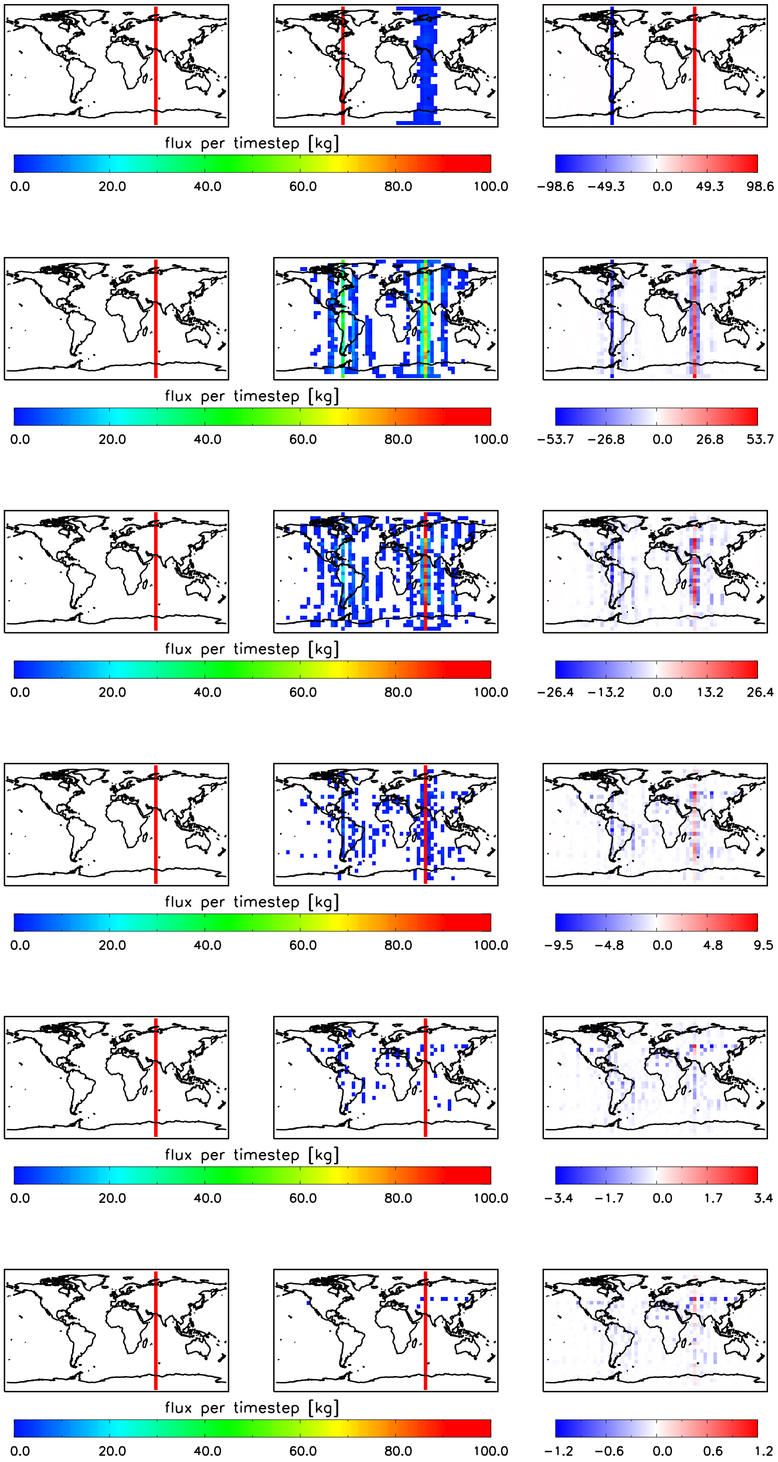
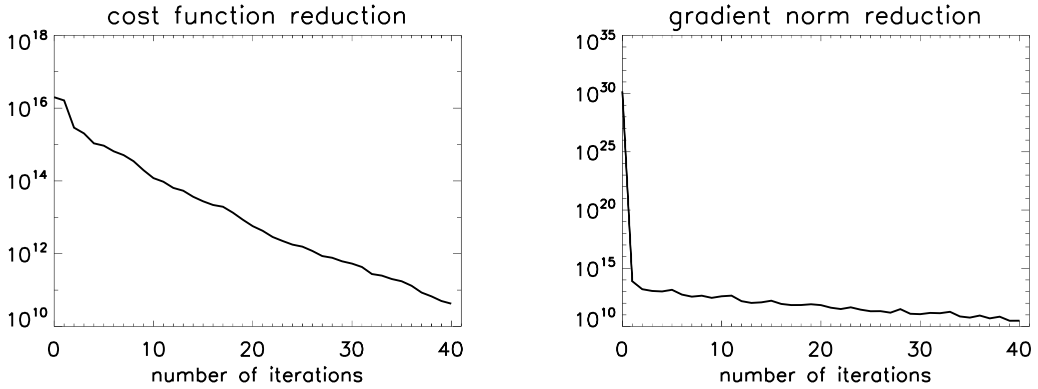


Figure 5.16 *(Left)* Value of the cost function and *(Right)* value of the norm of the cost function gradient for the retrieval displayed in Figures 5.14 and 5.15.



$$\frac{2r_{min}}{n_{obs}} = 1 + s_n \tag{5.32}$$

where s_n represents the scale of the model error relative to the data error. This term is dependent upon the length of the assimilation window, as any model errors propagate and accumulate throughout the window, becoming more significant as time goes on. In order to investigate the effect of model error of the TOMCAT 4D-Var system, an experiment was carried out following a method similar to that of Chevallier et al. (2007), in which the model creates a 3D concentration field, $\mathbf{y}_{clim}$, produced from known surface fluxes $\mathbf{x}_{true}$, both of which are then randomly perturbed in order to create pseudo-observations $\mathbf{y}$ and a priori $\mathbf{x}_a$ as follows;

$$\mathbf{y} = \mathbf{y}_{clim} + \mathbf{V}^T \mathbf{v}^{\frac{1}{2}} \mathbf{p} \tag{5.33}$$

$$\mathbf{x}_a = \mathbf{x}_{clim} + \mathbf{W}^T \mathbf{w}^{\frac{1}{2}} \mathbf{q} \tag{5.34}$$

where $\mathbf{V}$ and $\mathbf{v}$ are the eigenvector and eigenvalue matrices of the observation error covariance matrix $\mathbf{R}$, and $\mathbf{W}$ and $\mathbf{w}$ are the eigenvector and eigenvalue matrices of the prior error covariance matrix $\mathbf{B}$, so that $\mathbf{R} = \mathbf{V}^T\mathbf{v}\mathbf{V}$ and $\mathbf{B} = \mathbf{W}^T\mathbf{w}\mathbf{W}$. $\mathbf{p}$ and $\mathbf{q}$ are realisations of random variables with standard normal distributions which are the size of $\mathbf{y}$ and $\mathbf{x}_a$, respectively. Correlations between all errors are neglected, which means that $\mathbf{R}$ and $\mathbf{B}$ are diagonal matrices, so that Equations 5.33 and 5.34 reduce to;

Table 5.3 Values of n_{obs} and r_{min} at the end of the minimisation, for several lengths of assimilation window.

Simulation	Assimilation Window	n_{obs}	r_{min}	$\frac{2r_{min}}{n_{obs}}$	Difference from A_6
A_6	6 hours (4 cycles)	237568	118417	0.9969	0.0000
A_{12}	12 hours (2 cycles)	118784	58795	0.9899	0.0007
A_{24}	24 hours (1 cycle)	59392	28991	0.9763	0.0206 ($\approx 3 \times$ 0.0007)

$$\mathbf{y} = \mathbf{y}_{clim} + \mathbf{v}^{\frac{1}{2}}\mathbf{p} \tag{5.35}$$

$$\mathbf{x}_a = \mathbf{x}_{clim} + \mathbf{w}^{\frac{1}{2}}\mathbf{q} \tag{5.36}$$

Three inversions, A_6, A_{12} and A_{24}, were carried out in order to constrain the perturbed fluxes $\mathbf{x}$ to the perturbed observations $\mathbf{y}$ by running the model for one day only. The length of the assimilation window in each inversion was set to be 6 hours, 12 hours and 24 hours, respectively. The inversion was carried out for 24 hours during 1 January 2008, using an initial CH_4 field produced from output from the CTL_S TOMCAT forward model simulation, which is described in Section 6.2. Observation errors were set to be 0.1 ppb at all grid points. In order to allow the inversion to alter fluxes as necessary to match the mixing ratio perturbations, the prior error was set to be three times as large as the highest flux rate. The minimisation was continued until the value of the cost function had been reduced 20 times. This number of iterations was enough for the value of J to be falling by less than 3 during each reduction. Table 5.3 shows the evolution of $\frac{2r_{min}}{n_{obs}}$ as the length of the assimilation increases. The fact that the values are less than one indicates that the background errors are overestimated compared to the observation error, but this was intentional. Underestimating the background error is dangerous, as it may lead to the solution drifting away from the observations (Talagrand and Bouttier, 1999). The value of $\frac{2r_{min}}{n_{obs}}$ changes as the assimilation window length increases, although the value of the change is extremely small, indicating that model error does have a small effect on inversion results. From this limited example, the value of s_n, the value of the model error in comparison with the data error, is approximately 0.0031 for a 6 hour assimilation window, with an extra 0.0007 added for every 6 hours added to the assimilation window. This indicates that the model error would be small even for longer assimilation windows, and may be neglected for initial inversions using the TOMCAT 4D-Var system. Tremolet (2007) suggested a method in which the model error is explicitly formulated as part of the control variable, and this may be a consideration in future inversions.

5.7 Summary

The adjoint version of the TOMCAT CTM has been developed as part of this thesis, and has been integrated into the TOMCAT 4D-Var inverse model, which was also developed within this study. Since the TOMCAT model set-up chosen for use in the 4D-Var system was linear, there was no need to construct a tangent linear version of the model. The adjoint model was coded explicitly, and tested thoroughly in order to validate the adjoint transport. The adjoint model passed the numerical identity test for an adjoint model, and the adjoint transport also passed tests of its conservation and reciprocity against the forward model on a long time-scale.

The TOMCAT 4D-Var model was tested in order to examine its ability to retrieve known surface fluxes to a high level of accuracy. The 4D-Var system was able to assimilate synthetic observations produced using the forward TOMCAT model in order to retrieve the fluxes used to produce the observations to within 1%, even when given an incorrect a priori flux estimate. The inverse model was able to reduce the residual of the cost function to the value of the number of observations, as described in Dee (1995). The system does introduce model error, although it this is found to have only a small effect as the assimilation window lengthens. The TOMCAT 4D-Var system was found to be accurate and consistent, and ready to be used to assimilate observations in order to constrain surface emissions of CH_4.

Chapter 6

Investigating Recent Spatial and Temporal MethaneVariations with the TOMCAT Forward Model

6.1 Introduction

This chapter explores how well the TOMCAT model and emission inventories come together in order to accurately reproduce observed concentrations of methane. Assessing the strengths and weaknesses of current CH_4 emission inventories is an essential step in the process of improving their accuracy. Comparison of the simulated atmospheric concentration of CH_4 from TOMCAT with observational data can provide information on areas where the accuracy of the current best emission estimates can be improved. In order to produce an improved flux estimate using the 4D-Var inversion process, it is important that the modeller has confidence in the accuracy of the forward model transport, the atmospheric chemistry scheme used in the model and the quality of the emission inventory which will be used as an a priori during the inversion. The quality of the atmospheric transport in TOMCAT was evaluated in Chapter 4, using the inert tracer SF_6, and this chapter will investigate the effect that the CH_4 chemistry and emissions used in the model have on simulated CH_4 atmospheric concentration and will also compare the model output to observed concentrations.

Section 6.2 gives details of the model set-up used in this work to study tropospheric CH_4 concentrations, and details the different emission inventories used in the simulations. Section 6.3 discusses the reproduction of atmospheric methane chemistry in the TOMCAT model, and tests the validity of different tropospheric OH distributions against observations. Section 6.4 compares the modelled CH_4 concentrations to global observed data, while Section 6.5 summarises the results of this chapter. The simulations of CH_4 carried out as part of this chapter were submitted to the TransCom CH_4 model intercomparison project, in which the results of simulating CH_4 using a standardised set of emissions and chemical loss fields produced from a variety of CTMs were compared to each other and to observations. The results of the comparison are described in detail by Patra et al. (2011).

6.2 Model Set-up

The TOMCAT model was used to produce simulated CH_4 atmospheric distributions for the period 1988 - 2010. The model was initialised on January 1 1988, with CH_4 concentrations provided as part of the TransCom CH_4 experiment (Patra et al., 2011). Two years were allowed for model spin-up, and three-dimensional modelled CH_4 concentration fields were output every 90 hours. The model grid had a horizontal resolution of $2.8° \times 2.8°$, with 60 vertical σ-p levels up to a height above the surface of approximately 60 km. The model time step was 30 minutes, and winds were provided by ECMWF ERA-Interim re-analyses from 1989 onwards. ECMWF ERA-40 re-analyses were used for the year 1988. The simulations used the SOM advection scheme and the Holtslag & Boville PBL scheme.

6.2.1 Emission Inventories

In order to investigate the effect that different emission estimates have on the simulated atmospheric distribution of CH_4, six different emission inventories were used, all of which were supplied as part of the TransCom CH_4 experiment (Patra et al., 2011). This provided monthly mean CH_4 emissions on a $1° \times 1°$ grid for each year in the simulation up to and including 2008. 2008 emission estimates were repeated for the years 2009 and 2010. Each inventory was made up of various contributions of different source estimates, from both bottom-up and top-down estimates. Table 6.1 gives details of the different emission components which contribute to the six TOMCAT emission inventories.

EDGAR 3.2 anthropogenic emissions

These emissions (IAV ANT) are based on the annual mean $1° \times 1°$ emission maps produced by the EDGAR, version 3.2 (Olivier and Berdowski, 2001). EDGAR emission inventories are produced for the years 1990, 1995 and 2000, and inventories for all other years are interpolated

Table 6.1 Details of CH_4 emission components used to compile inventories used in simulations in Chapter 6. IAV indicates emission components which vary year-to-year, while CYC indicates components which do not vary inter-annually.

EMISSION TYPE	DESCRIPTION	TIME RESOLUTION
Anthropogenic emission (IAV ANT)	Based on Edgar 3.2	Annual means; IAV (inter-/extra-polated using 1990, 1995, 2000 emission maps)
Anthropogenic emission (IAV ANT E4)	Based on Edgar 4.0	Annual means; IAV (2005 used for 2006 - 2008).
Natural emission (CYC NAT)	GISS inventory, REAS rice, Ocean and Mud Volcano	Monthly means; CYC. Trends in rice emission is used to scale REAS rice
Wetland emission (IAV WL)	ORCHIDEE model based	Monthly means; IAV for 1994 - 2000 (CYC for other years)
Wetland emission (IAV WLe)	VISIT model simulated wetland and rice	Monthly means; IAV 1988 - 2008
Biomass Burning emission (IAV BB)	GFED version 2; Satellite products based	Monthly means; IAV for 1996 - 2008. (CYC for 1988 - 1996)
Inversion flux (IAV INV)	IPSL optimised flux (Bousquet et al., 2006)	Monthly means; IAV for 1988 - 2005. (CYC for 2006-2008)

or extrapolated from these as in Patra et al. (2009). Emissions are inter-annually varying monthly means.

EDGAR 4.0 anthropogenic emissions

These emissions (IAV ANT E4) are based upon the more recent version of the EDGAR database (version 4.0) (http://edgar.jrc.ec.europa.eu). This provides 1° × 1° emission inventories for each year up until 2005. These 2005 emissions are repeated for 2006 - 2008. Emissions are interannually varying monthly means.

GISS inventory natural emssions

These emissions (CYC NAT) include all natural emissions with an annual cycle such as those from all natural wetlands, from domestic and large-scale biomass burning and from termites, and are based on the GISS inventory (Matthews and Fung, 1987). Emissions from rice paddies are taken from Yan et al. (2009) and all emission components are scaled individually as in Patra et al. (2009) in order to produce an annual total of 273 Tg. Although rice emissions are mostly anthropogenic in nature, they are included in this category as the seasonal cycle is dependent upon seasonal rainfall and temperature. Ocean exchange emissions are distributed over coastal regions (Lambert and Schmidt, 1993; Houweling et al., 1999), and mud volcano emissions are based upon Etiope and Milkov (2004). These emissions do not vary annually between years.

ORCHIDEE wetland emissions

These emissions (IAV WL) are derived from the wetland emission module of the ORganizing Carbon and Hydrology in Dynamic EcosystEms (ORCHIDEE) terrestrial ecosystem model (Ringeval et al., 2010), which uses satellite data to derive area of inundation for the period 1994 - 2000 (Prigent et al., 2007). These emissions are scaled by a factor of 0.76 in order to match the wetland emission component of CYC NAT. An average seasonal cycle is used for all periods not derived by the ORCHIDEE model.

VISIT wetland emissions

These emissions (IAV WLe) are from the Vegetation Integrative Simulator for Trace gases (VISIT) terrestrial ecosystem model (Ito, 2010). They provide a different estimate of global wetland and rice emissions by estimating inundated area through analysed rainfall and temperature (Mitchell and Jones, 2005). The wetland emission estimates are scaled by a factor of 0.69 to match with CYC NAT, while the rice emissions are similarly scaled by a factor of 0.895.

GFED version 2 biomass burning emissions

Biomass burning emissions (IAV BB) are from the Global Fire Emission Database (GFED) Version 2, and mainly represent forest and savannah burning (van der Werf et al., 2006). This dataset is available only after 1997, so an average seasonal cycle is used for the preceding years. This method produces a global emission total around 35% of that in CYC NAT. It is likely that this method may overestimate open burning while underestimating closed burning of biomass.

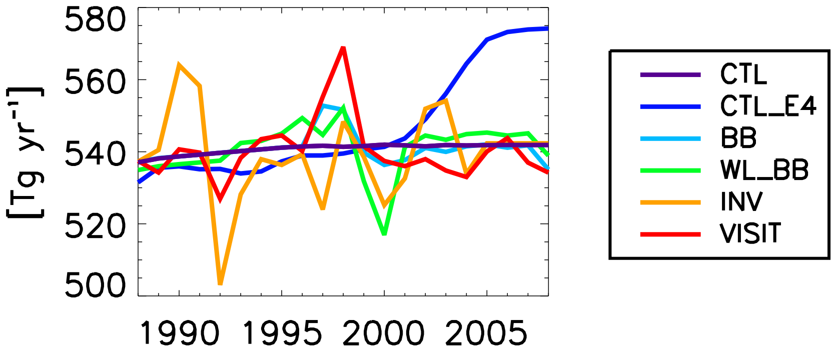
IPSL inversion flux

This complete emission inventory (IAV INV) is obtained through an optimised synthesis inversion of measured CH_4 concentrations using the LMDZ model over the period 1988 - 2005 (Bousquet et al., 2006), previously discussed in Chapter 3. An average seasonal cycle is used for the years not provided.

Table 6.2 Deatils of TransCom emission inventories.

INVENTORY NAME	DESCRIPTION
CTL	CYC NAT + IAV ANT (CYC BB & CYC WL)
CTL_E4	CYC NAT + IAV ANT E4
BB	CTL - 0.35 CYC BB + IAV BB
WL_BB	BB - CYC WL + 0.76 IAV WL
INV	as in Table 6.1
VISIT	BB - CYC WL - Rice + WLe IAV (0.69 Wetland + 0.895 Rice)

Figure 6.1 Total annual CH_4 emissions for six TransCom emission inventories in Tg yr^{-1}.



From the emission processes described above, six complete emission inventories are compiled, and are detailed in Table 6.2. Figure 6.1 shows the total annual emissions for each inventory. Each emission inventory emits 500 - 580 Tg yr^{-1}, with inventories with inter-annually varying wetland emissions displaying the widest range of annual emission totals.

Figure 6.2 shows mean global distribution of methane emissions over the period 1990 - 2008 for DJF and JJA for each emission inventory, while Figure 6.3 shows monthly mean emissions for each inventory for each year individually. The CTL inventory displays high emission rates in DJF in industrialised areas such as Europe and the USA, and in South Asia where high emissions are due to large areas of rice paddies and wetlands. In JJA, these emissions continue, while biomass burning emissions increase in the tropics and wetland emissions begin after the thawing at the NH high latitudes. Emissions peak at approximately 680 Tg yr^{-1} in August from 460 Tg yr^{-1} in DJF. Since cyclic natural emissions are used for this inventory, all inter-annual variability is due to anthropogenic trends, leading to a slight year-on-year increase.

The CTL_E4 inventory displays a similar seasonal variation to that of CTL. However, the total annual emissions are more variable than in CTL. Total global emissions are lower than in CTL until the year 2000, when emissions undergo a rapid and significant increase, reaching a level

Figure 6.2 Mean geographical distribution of CH_4 emissions for six TransCom inventories in molecules cm^{-2} s^{-1} for DJF and JJA for the period 1988 - 2008. Underneath each plot is the difference from the CTL inventory. The left-hand colour bar corresponds to the top two plots for each emission inventory and measures total emissions, while the right-hand colour bar corresponds to the bottom two plots for each inventory and measures the difference from the CTL inventory.

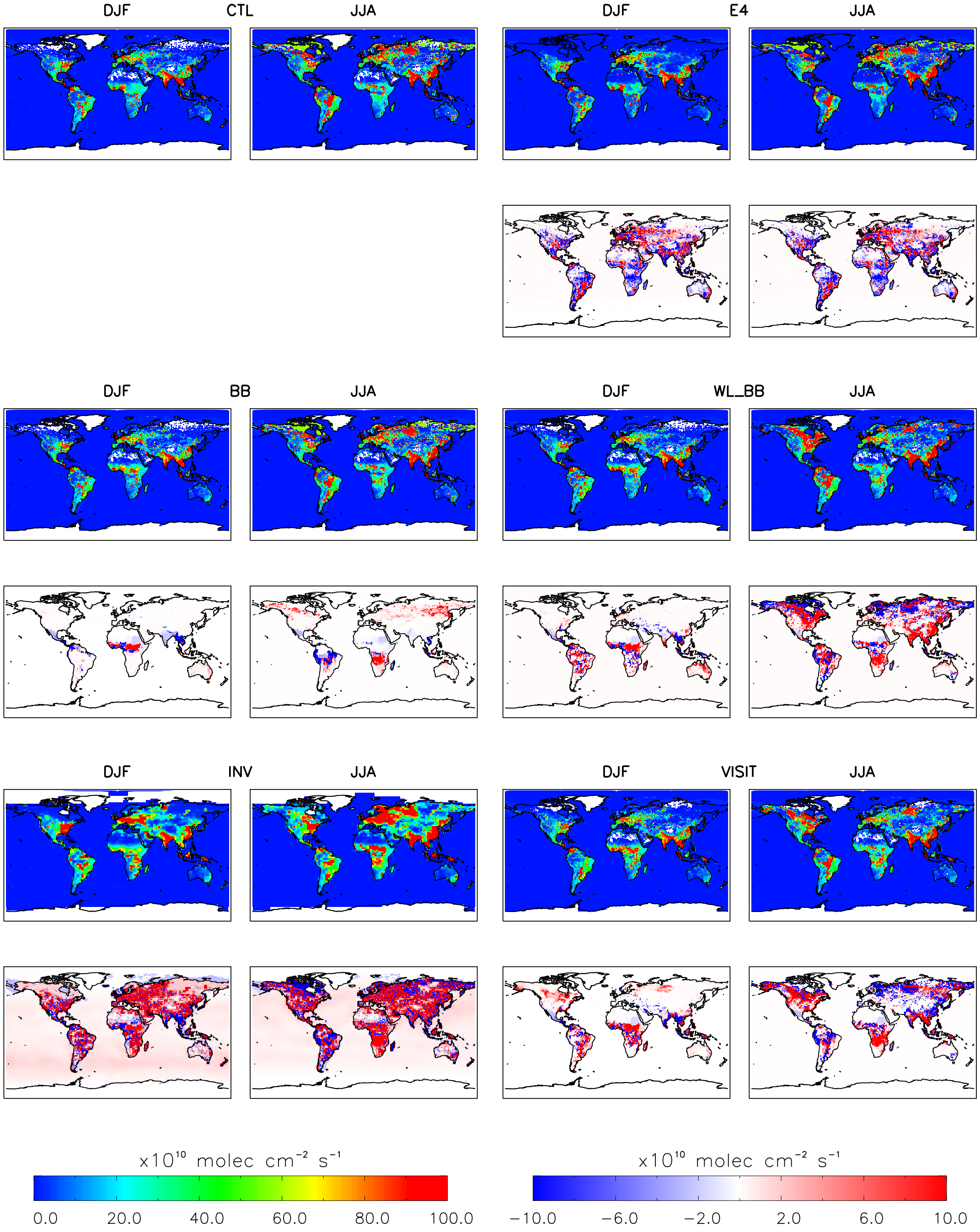
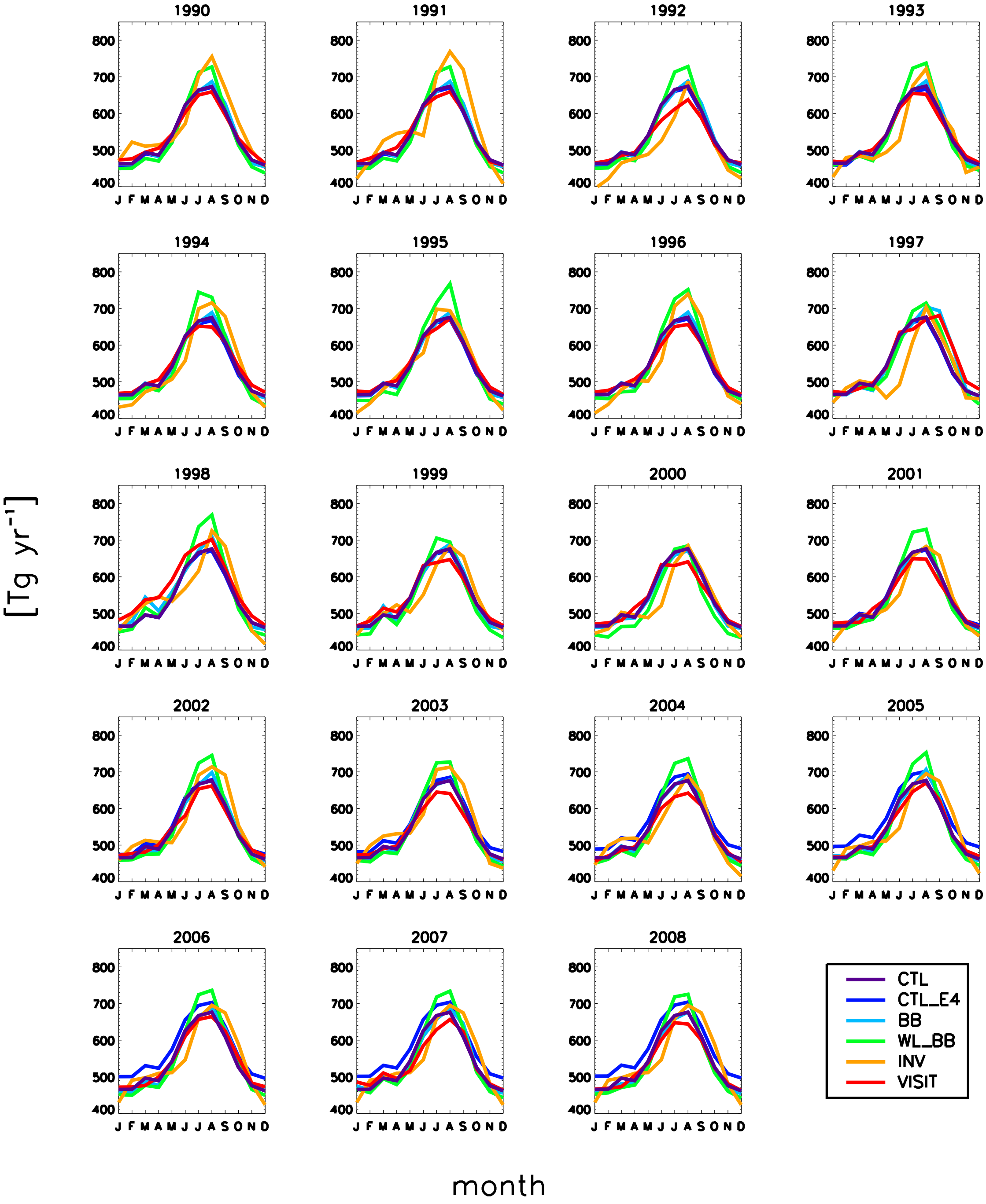


Figure 6.3 Monthly mean CH_4 emissions for each year in the period 1990 - 2008 for each of the six TransCom inventories in Tg yr^{-1}.



approximately 5% higher than that of CTL in the year 2005. The geographical distribution of emissions is very similar to that in CTL, as only the anthropogenic emissions have been altered.

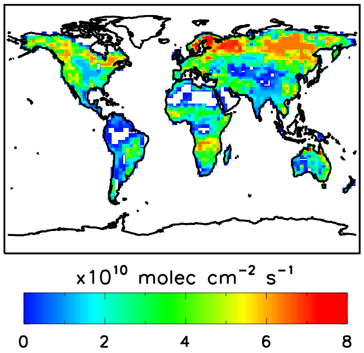
The BB inventory uses the same anthropogenic emissions as CTL, but introduces an interannual variability into the biomass burning emissions. This inventory is therefore similar to CTL in DJF, with only a redistribution of tropical African burning emissions and reduced emissions in South Asia. In summer, this inventory displays larger CH_4 emissions across the high-latitude NH and in sub-tropical Africa, with lower emission rates in South America. The seasonal cycle of the emissions is similar to that of CTL, but there is a greater level of inter-annual variation.

Inventory WL_BB introduces interannual wetland variability to the BB inventory. WL_BB emissions are increased in Northern Australia and South America during DJF compared to BB, while Indonesian emissions are reduced. Overall, WL_BB are lower in DJF than in BB. During JJA, WL_BB CH_4 emissions are up to 150 Tg yr^{-1} larger than BB, due to increased wetland emissions in North America and South Asia. Emissions are also redistributed across Siberia, Northern Europe and South America in JJA, while emissions in Alaska and North-East Canada are reduced. Aside from a large decrease in annual emissions during 2000 due to drought, annual WL_BB emissions are generally higher than those in CTL.

The top-down emission inventory, INV, displays the greatest interannual variation of emissions and a different seasonal cycle to the previous inventories. In 1990, for example, INV showed the highest emission rate of all the inventories in all months except May and June, while in 1992 it displayed the lowest emission rates throughout January to July. This inventory displays a dramatic increase in emission rate later in the year than other inventories, with emission rates in April, May and June regularly being lower than any other inventory. Total annual emissions are almost always lower than those of CTL, particularly throughout the period 1992 - 1997. Geographically, emissions are generally higher than those of CTL throughout the year in Europe, USA, East Africa and much of South Asia. However, South-East Asian emissions are lower in DJF and Canadian and Russian wetland emissions are reduced during JJA.

Finally, the VISIT dataset also shows large interannual variability compared to CTL. Emission rates are similar to those in CTL throughout the early- to mid-nineties apart from a summer decrease in 1992, which is probably due to the eruption of Mount Pinatubo a year earlier leading to a decrease in precipitation (Trenberth and Dai, 2007). Excluding a high emission rate in 1997/98 due to the strong El Niño that year, summertime emissions are generally lower than CTL from 1996 onwards. This inventory shows increased wetland emissions throughout the year in North America and central Africa, while Asian rice paddy emissions are lower in winter and higher in summer. Northern hemisphere Eurasian polar summer emissions are redistributed away from Scandinavia towards Siberia.

Figure 6.4 Annual mean distribution of CH_4 soil sink in molecules cm^{-2} s^{-1}.



6.2.2 CH_4 Soil Sink

The methane soil sink used in the simulations was supplied by Patra et al. (2011), and was formed by taking the mean seasonal cycle from the LMDZ atmospheric CH_4 inversion (Bousquet et al., 2006). Figure 6.4 shows the annual mean distribution of the soil sink. The soil sink has a small seasonal cycle which peaks in October due to the NH Arctic thaw, although the global sink remains at 27±0.1 Tg CH_4 yr^{-1}, approximately 5% of the total annual CH_4 sink. Methane uptake is greatest in the NH high latitudes across Scandinavia and Siberia, and lowest in desert areas.

6.3 Atmospheric CH_4 Chemistry

6.3.1 Tropospheric OH Chemistry and Model OH Fields

OH is produced in the troposphere as a product of the reaction of the excited oxygen atom, $O(^1D)$ with water vapour. $O(^1D)$ is produced through the photolysis of ozone (O_3) by ultraviolet (UV) radiation with a wavelength of 300 - 320 nm. The production of OH from O_3 is as follows;

$$O_3 + h\nu \longrightarrow O_2 + O(^1D) \tag{6.1}$$

$$O(^1D) + M \longrightarrow O + M \tag{6.2}$$

$$O(^1D) + H_2O \longrightarrow 2OH \tag{6.3}$$

Laboratory studies have indicated that reeaction 6.2 takes place at a much faster rate than reaction 6.3. This fact, together with the fact that much of the UV radiation within the narrow wavelength band necessary for reaction 6.3 to take place does not reach the troposphere due to high O_3 concentrations in the stratosphere, means that the probability of a tropospheric O_3 molecule producing

OH is fairly small. Since the tropospheric chemical lifetime of OH is very short, tropospheric OH concentrations are extremely low.

As discussed in Chapter 2, around 90% of CH_4 released into the atmosphere is destroyed by the OH, via Equation 2.4. Due to inherent difficulties in measuring atmospheric OH, three-dimensional OH fields for use in CTMs are not well constrained, with large uncertainties over the species' distribution and abundance in the troposphere (Heard and Pilling, 2003). In order to investigate the effect of the OH on atmospheric CH_4 concentration, two different OH fields were used in this study. The first was provided as part of the TransCom experiment (Patra et al., 2011), and was originally derived from Spivakovsky et al. (2000), while the second OH field was taken from the full chemistry version of the TOMCAT model (Arnold et al., 2005). The Spivakovsky and TOMCAT OH fields will henceforth be labelled OH_S and OH_T, respectively. Both of the fields were averaged in order to create monthly mean OH estimates, and had no interannual variation. Montzka et al. (2011a) inferred a small interannual variability of atmospheric OH concentrations using measurements of various atmospheric trace gases. In this work, the rate constant for the reaction of CH_4 with OH, k_{OH}, is specified as $k_{OH} = 2.45 \times 10^{-12} exp(-1775/T)$.

The OH_S distribution was first developed using observed distributions of O_3, H_2O, NO_X, CO, hydrocarbons and cloud optical depth together with a set of kinetic and chemical equations in a photochemical box model by Spivakovsky et al. (2000). This distribution was then reduced by 8% by Huijnen et al. (2010) in order to reproduce the observed decline of methyl chloroform (CH_3CCl_3, MCF) since 2000 in the TM5 transport model. The monthly mean, zonal mean distribution of OH is displayed in Figure 6.5. This field has low OH concentrations near the surface, and peak concentrations are at approximately 700 hPa. Annually, there is an equal abundance of OH in the NH and SH (Spivakovsky et al., 2000).

The TOMCAT full chemistry model simulates OH concentrations on-line through calculating its chemical reactions with other atmospheric species as already discussed in Chapter 3. The monthly mean OH fields used in this study represent the mean OH distribution for each month produced by the model over a 7-year period (2000-2006).The model grid had a resolution of 2.8° × 2.8°, and 31 vertical levels up to a pressure of 10 hPa. Emissions were taken from the 2001 IPCC report (Houghton et al., 2001), and the Global Fire Emissions Database (GFED) v.2. Chemistry was as described in Arnold et al. (2005). The resulting monthly mean, zonal mean OH_T distribution is shown in Figure 6.6. The annual global concentration of OH produced by this model is approximately 10% lower than OH_S, with OH concentrations peaking close to the surface, and a strong bias towards the NH.

Table 6.3 shows the tropospheric mass-weighted mean concentrations for both OH fields used in this study. Here, the tropopause is defined as per the standardised climatological definition suggested by Lawrence et al. (2001);

Figure 6.5 OH_S monthly zonal mean concentration in molecules cm^{-3}.

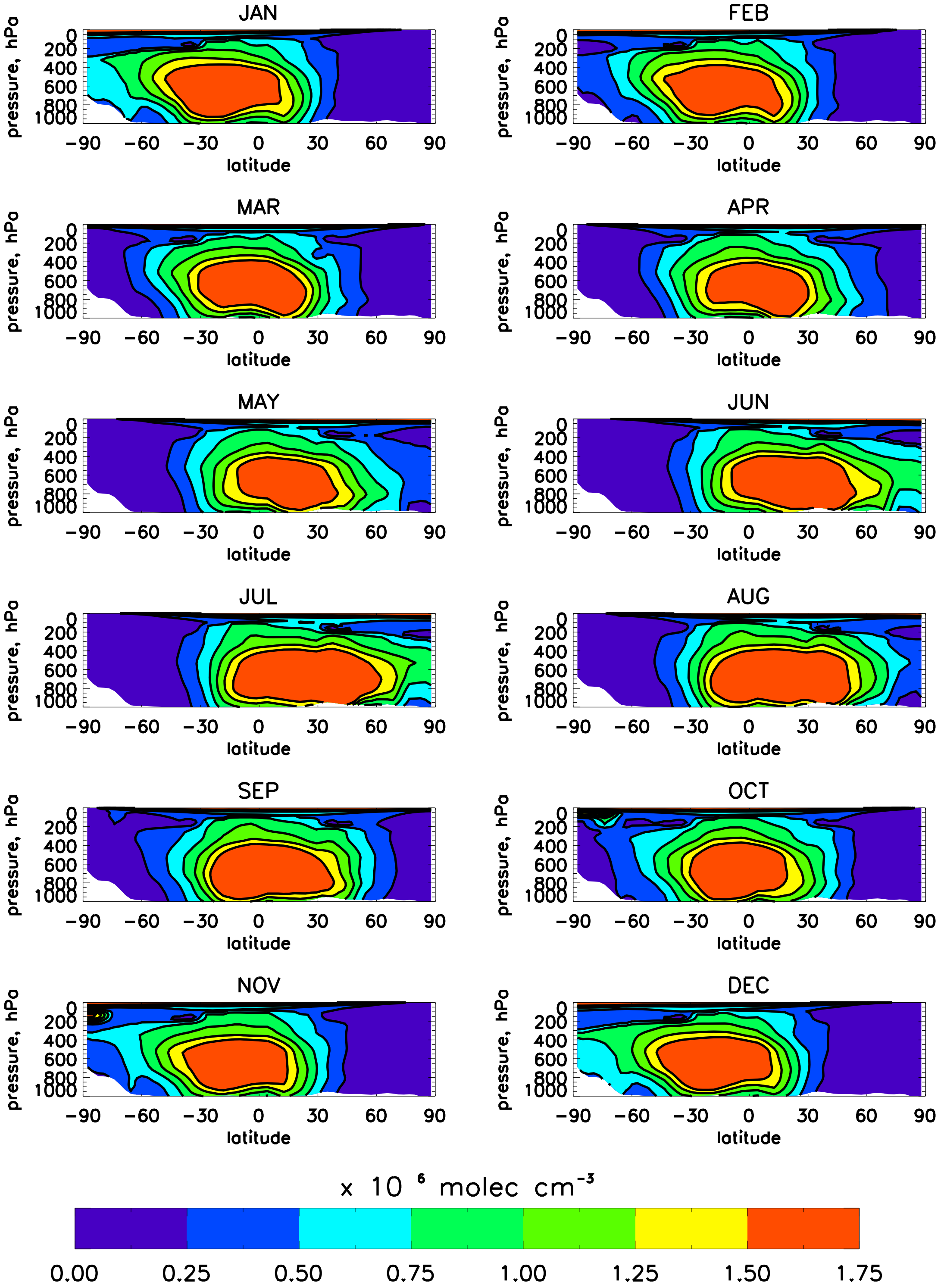


Figure 6.6 OH_T monthly zonal mean concentration in molecules cm^{-3}.

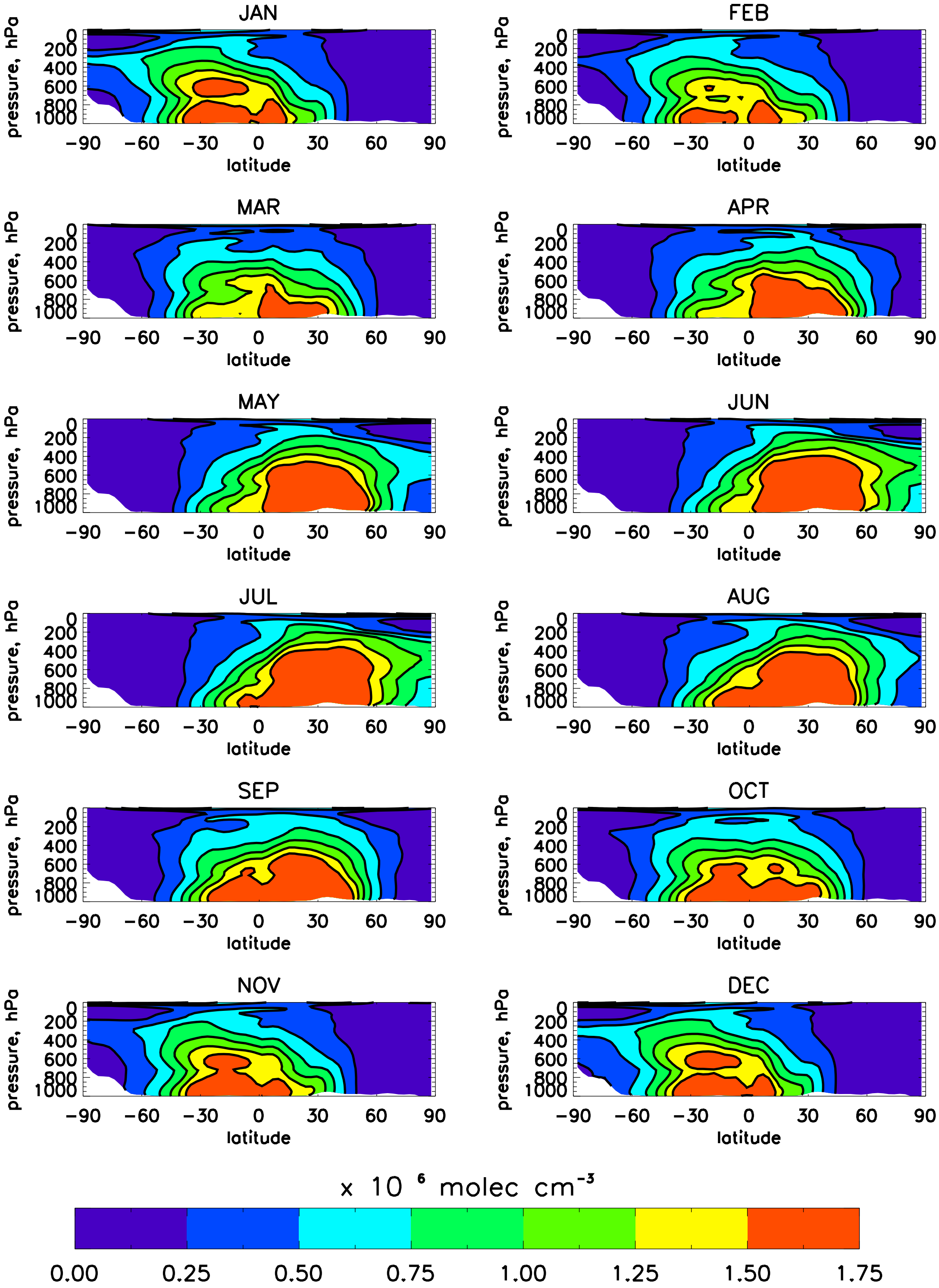
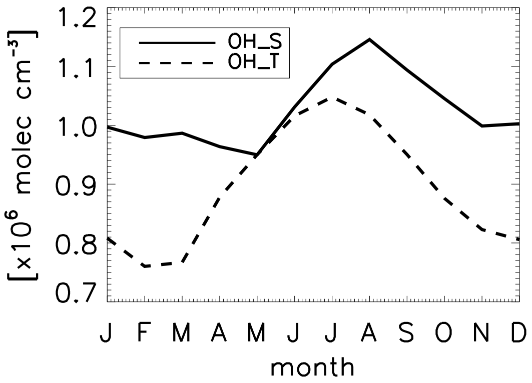


Table 6.3 Comparison of OH_S and OH_T mean concentration in molecules cm^{-3} for 1990. Concentrations are mass-weighted unless specified. % difference row represents difference of OH_T compared to OH_S.

	Global	**NH**	**SH**	**DJF**	**JJA**	**CH_4-weighted**	**MCF - weighted**
OH_S	1024800	1031230	1018840	992928	1093550	1208630	1158490
OH_T	891639	1024060	760513	791346	1027290	1158490	1087620
% difference	13.0	0.7	25.4	20.3	6.1	4.1	6.1

Figure 6.7 Global monthly mean tropospheric OH concentration for OH_S and OH_T in molecules cm^{-3}.



$$p_{tr} = 300 - 215(cos(\phi))^2 \qquad (6.4)$$

where p_{tr} is the pressure level of the tropopause in hPa and ϕ is the latitude. The annual global mean OH_T concentration is 13% smaller than OH_S each year, with the majority of the difference coming in the SH. Annual NH values are within 0.7% of each other, whereas the annual SH OH_T concentration is just 75% of the OH_S value. Figure 6.7 shows the monthly global mean concentration of OH in each field. While total OH in May and June is almost identical in both OH fields, the values can differ by up to 25% during November through to March. Possible reasons for the discrepancy between the two distributions is discussed in Section 6.4.

6.3.2 The PEM-tropics Aircraft Campaign

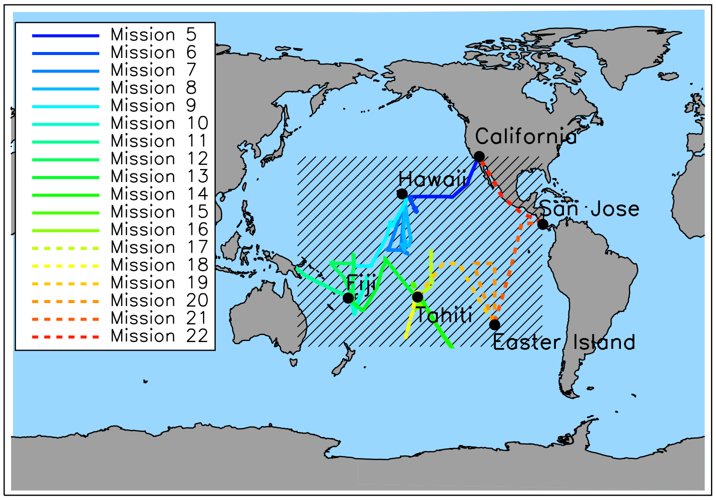
The hydroxyl radical is difficult to observe in the atmosphere due to its short lifetime - and therefore its low concentration - and there have been relatively few recent measurement campaigns to discover more about the atmospheric concentration of the species. Observations are generally obtained through spectroscopic methods, and efforts must be made to remove the effect of interference that the measurement itself can create (Heard and Pilling, 2003). These inherent difficulties mean that attempts to validate the two OH fields used in this study are fairly limited in scope.

However, tropospheric OH was measured as part of the NASA-sponsored Global Tropospheric Experiment (GTE) Pacific Exploratory Mission (PEM) Tropics A and B, during which large number of hydrocarbons, halocarbons and organic nitrates were quantified in the Pacific area. Carried out during March and April 1999, the PEM Tropics B mission consisted of around 40 flights on the P-3B aircraft and OH measurements were conducted using the Selected Ion Chemical Ionization mass spectrometry (SICIMS) technique. For more information on the aircraft campaign and measurement technique see Raper et al. (2001) and Mauldin III et al. (2001), respectively. For detailed results of the OH measurements see Tan et al. (2001). Measurements were taken day and night over a two month period in the areas surrounding California, Hawaii, Fiji,Tahiti, Easter Island and Costa Rica. Flight tracks from the mission are shown in Figure 6.8.

Since these measurements have a much higher temporal and spatial resolution than the monthly mean distributions used in this work, measurements were averaged into altitude bins of 100 hPa in size, and were also separated into those taken in March and those in April. OH is produced through the photolysis of O_3 by sunlight, and its extremely short lifetime, of the order of 0.1 - 1 seconds, means that OH concentrations are high during the day and practically zero at night (Bloss et al., 2005). In order to compare results from the aircraft campaign to the monthly mean OH fields used in the model, it is therefore necessary to sample both day-time and night-time measurements. Since the flights were mostly carried out during daylight hours, all night-time measurements were removed and the resulting OH concentrations were divided by two in order to compare with the monthly-mean model fields. As all flights were carried out in the tropics, day and night are approximately 12 hours each. Whilst this method does not produce accurate monthly mean OH concentrations for comparison with the model OH fields, it does provide some idea of the distribution and magnitude of tropospheric OH concentration during the campaign. OH_S and OH_T were distributed into the same altitude bins, and only the area surrounding the flight paths of the missions are considered. This area is shaded in Figure 6.8.

Figure 6.9 shows a comparison of the two model OH fields with the campaign measurements. The magnitude of the OH concentrations may not be directly comparable, due to the inexact nature of the data treatment, but the missions do give some information about the nature of the vertical distribution of the hydroxyl radical. During both March and April, the aircraft measurements show a vertical profile which peaks at approximately 500 - 600 hPa. This distribution is similar to that of OH_S, which also peaks at about 500 hPa. The magnitude of OH_S concentrations are also similar to the observations, with a background of approximately 1×10^6 molecules cm^{-3} at the surface

Figure 6.8 Flight paths of PEM-Tropics B campaign for Missions 5-22. Different flight paths are represented by different colours, with solid lines representing flights undertaken in March 1999, while dashed lines represent flights from April of that year. The shaded area represents the area over which the OH_S and OH_T fields are averaged.



and above 400 hPa, and a peak at approximately 1.5×10^6 molecules cm^{-3}. However, OH_T concentration has a different distribution, peaking at the surface and decreasing with altitude. Surface concentrations are approximately 1.6×10^6 molecules cm^{-3} and fall to 0.5×10^6 molecules cm^{-3} above 200 hPa. Away from the surface, OH_T concentration is consistently less than both the observations and OH_S, although in the mid troposphere it is often within one standard deviation of the observations. Although this is a limited comparison, it does imply that use of the OH_T may provide slow CH_4 loss in the upper troposphere and rapid loss at the surface, and this will affect the total methane budget.

6.3.3 The ARCTAS Mission

Carried out in 2008, the Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission was created with the aim of investigating the effect of long-range transport and boreal fires on pollution in the Arctic. The mission, which was carried out by the by the Global Tropospheric Chemistry Program and the Radiation Sciences Program of the US National Aeronautics and Space Administration (NASA), used the DC–8, P–3 and B–200 research aircraft in order to take measurements of a large number of atmospheric trace species to augment continuous satellite measurements of the Arctic atmosphere. For details of the mission's design, execution, and first results, see Jacob et al. (2010). Of particular interest here is the measurements of Arctic OH taken on-board the DC–8 aircraft in July 2008. These flights took place mostly in Western Canada, as shown in Figure 6.10. The results of these flights were compared to OH_S and

Figure 6.9 Monthly mean vertical distribution of OH for PEM Tropics B samples (black), OH_S (blue) and OH_T (red) in molecules per cm^{-3} for March 1999 *(Left)* and for April 1999 *(Right)*. All data is averaged into pressure bins of 100 hPa in size. Error bars represent 1σ of the observations, and the grey bars display the number of observations in each pressure bin.

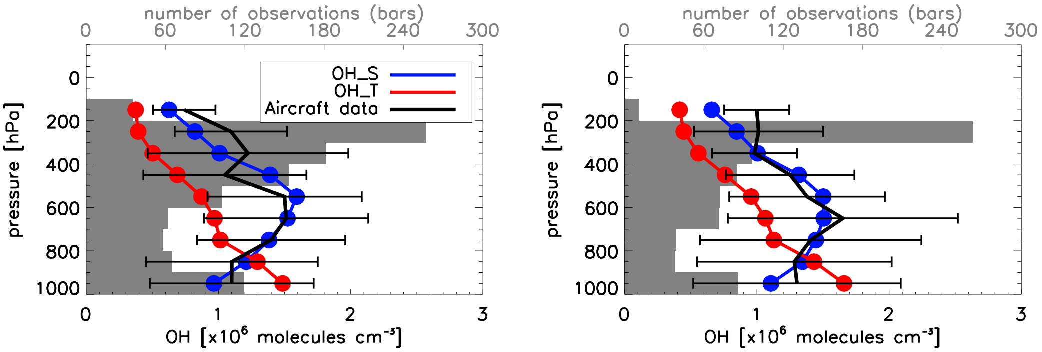
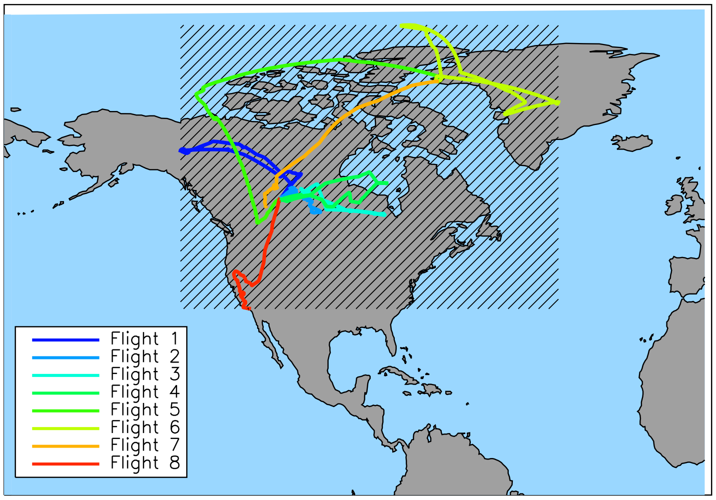


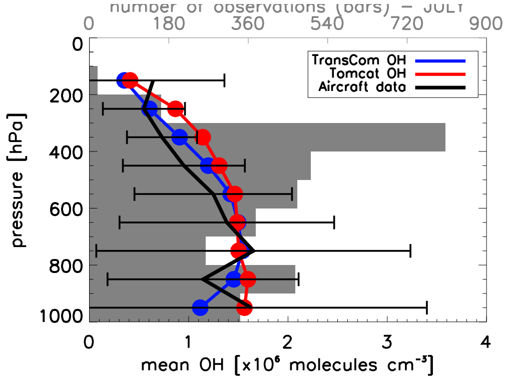
Figure 6.10 Flight paths of ARCTAS mission for flights 1-8. Different flight paths are represented by different coloured lines. The shaded area represents the area over which the OH_S and OH_T fields are averaged.



OH_T in the same way as the results of the PEM–Tropics campaign, described in Section 6.3.2. For each pressure bin, an average OH concentration is taken for the OH_S and OH_T fields over the shaded area in Figure 6.10.

Figure 6.11 shows a comparison between the results of the ARCTAS mission and the OH_S and OH_T fields. At this latitude, the two model fields show fairly similar distributions, unlike at tropical regions. The OH_T distribution again has increased OH near the surface, and also displays slightly higher concentrations than OH_S at pressure levels between 500 and 100 hPa. The

Figure 6.11 Monthly mean vertical distribution of OH for ARCTAS mission samples (black), OH_S (blue) and OH_T (red) in molecules per cm^{-3} for 2008. All data is averaged into pressure bins of 100 hPa in size. Error bars represent 1σ of the observations, and the grey bars display the number of observations in each pressure bin.



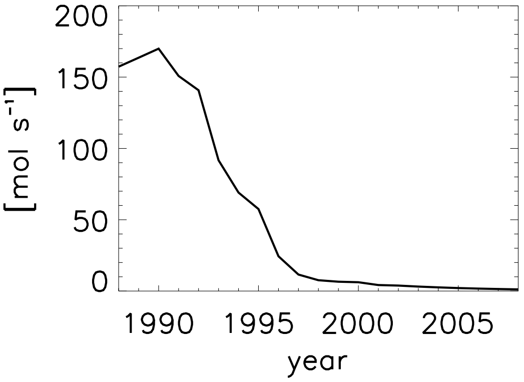
observed concentrations are lower than either of the model distributions in this pressure range, but observed concentrations rise to a similar level to those displayed by the model distributions towards the surface. However, the standard deviation of the observations towards the surface is large, and both OH_S and OH_T fall well within the error range.

6.3.4 Assessment of OH Distributions using Methyl Chloroform

Methyl chloroform (MCF) was widely used as an industrial solvent until 1992, when its role in the destruction of ozone in the stratosphere led to it being phased out as part of the Copenhagen Amendment to the Montreal Protocol, with a phase-out date of 1996 for developed countries. This meant that MCF emissions, which had grown substantially until 1992, rapidly shrank to almost zero. The annual global emissions of MCF for the period 1988 - 2008 are shown in Figure 6.12, which shows the sharp decline in MCF emissions after 1990. The major destruction process of MCF is due to reaction with the OH radical in the troposphere, with over 80% of the species being destroyed through this reaction (Prinn et al., 1995). Stratospheric photolysis and dissolution into the oceans account for around 13% and 6% of atmospheric destruction, respectively (Butler et al., 1991). The rapid decrease in MCF emissions in the early nineties, and the low uncertainty in their magnitude and distribution, mean that atmospheric MCF concentration is an excellent means for estimating tropospheric OH distribution (Prinn et al., 1995, 2001; Krol et al., 1998; Montzka et al., 2000; Spivakovsky et al., 2000).

In order to further assess the validity of each OH field, the TOMCAT model was used to simulate the atmospheric concentration of MCF. Monthly mean surface fluxes, stratospheric loss rates and

Figure 6.12 Globally integrated MCF emissions in mole s^{-1} for the period 1988-2008.



oceanic dissolution rates were provided as part of the TransCom project (Patra et al., 2011) and were represented in the model according to the following reactions:

$$CH_3CCl_3 + OH \xrightarrow{k_{OH}} Products \qquad [k_{OH} = 1.64 \times 10^{-12} exp(-1280/T)] \tag{6.5}$$

$$CH_3CCl_3 \xrightarrow{J_{CH_3CCl_3}} Products \tag{6.6}$$

$$CH_3CCl_3 \xrightarrow{DEP_{CH_3CCl_3}} Oceanic\ MCF \tag{6.7}$$

The oceanic MCF deposition rate, $DEP_{CH_3CCl_3}$, is estimated using the method of Kanakidou et al. (1995), and the loss rate is dependent on ocean temperature and ocean surface fraction. The global annual mean oceanic MCF deposition rate is approximately 2.7 × 10^{-6}m s^{-1} and the distribution of oceanic deposition velocities for January and July are displayed in Figure 6.13, indicating increased deposition rates in each hemisphere in winter. The annual mass-weighted mean of $J_{CH_3CCl_3}$, the stratospheric loss rate of MCF, is approximately 7.98 × 10^{-8} s^{-1}.

The two simulations were identical apart from the OH field. Each simulation was carried out using the same model grid, timestep, PBL scheme and advection scheme as the CH_4 simulations described in Section 6.2, and stratospheric destruction and oceanic deposition were identical. Both simulations were initialised with values provided as part of the TransCom CH_4 experiment on January 1, 1988. The simulation which used the OH_S field will henceforth be referred to as MCF_S, and the simulation using the OH_T field similarly referred to as MCF_T. Monthly mean 3D MCF concentration fields were created from model output and were compared to gas chromatograph data from the the NOAA/ESRL halocarbons in situ programme. Surface station details are displayed in Table 4.2. The global background MCF concentration, μ_{MCF}, is estimated each month as the mean of the MCF concentrations at BRW, MLO, SMO and SPO.

Figure 6.13 MCF oceanic deposition rates used in TOMCAT simulations in m s^{-1} for January *(Left)* and July *(Right)*.

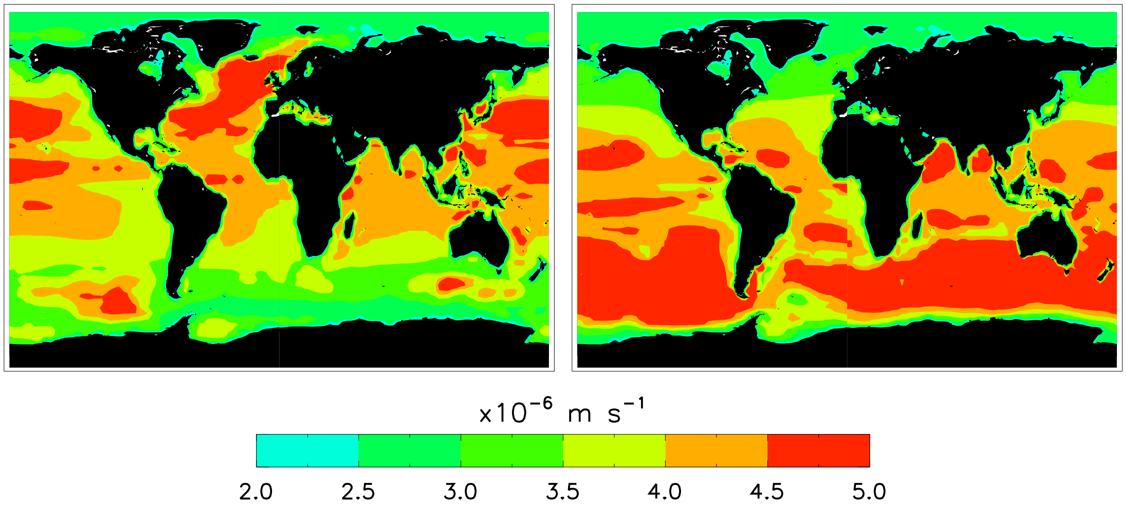
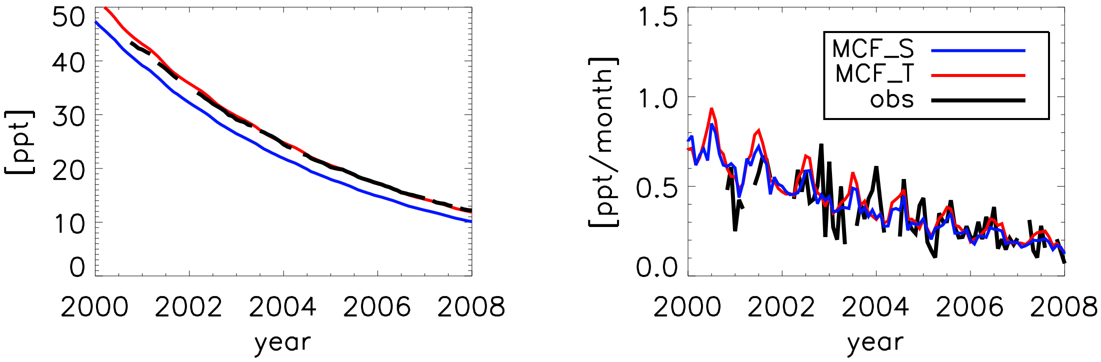


Figure 6.14 shows μ_{MCF} for the period 2000-2008 and $\Delta\mu_{MCF}$, the monthly change in μ_{MCF}. By 2000, global emissions of MCF are approximately zero, and the global MCF concentration is already decaying. μ_{MCF} falls from around 50 ppt in 2000 to around 15 ppt in 2008. Initially, μ_{MCF_S} is approximately 4 ppt lower than μ_{MCF_T}, before μ_{MCF_T} decays slightly faster so that the two quantities are within 1.5 ppt of each other by 2008. $\Delta\mu_{MCF}$ in both simulations is approximately equal to the decay rate estimated from the simulations, falling from 0.6 ppt month^{-1} to 0.2 ppt month^{-1}. μ_{MCF_T} is approximately equal to μ_{MCF_OBS}, which means that μ_{MCF_S} is 2 - 5 ppt lower. As shown in Table 6.3, the MCF-weighted mean OH_T concentration is only around 6% less than the OH_S concentration, which accounts for the small difference in the destruction rates. The low bias of μ_{MCF_S} does not necessarily indicate that there is too much atmospheric OH in that simulation, since other factors such as the stratospheric destruction or deposition could well be too strong, but it is clear that during this period, the loss rate of MCF due to OH in both simulations is accurate, and can be investigated further in the CH_4 simulations carried out later in this section.

6.3.5 Stratospheric Destruction

As discussed in Chapter 2, stratospheric destruction of methane occurs through reactions with the excited oxygen atom $O(^1D)$ and with atomic chlorine (Cl), along with reactions with stratospheric OH. This stratospheric sink of CH_4 accounts for approximately 5% of the total atmospheric methane loss. For this work, parameterised loss-rate fields were used in the stratosphere in order to remove the necessity for complex chemical schemes. The destruction rate used in the TOMCAT simulations was based upon the Cambridge/RIVM 2-D stratospheric chemistry model (Velders, 1995). The Cambridge 2-D model was originally developed by Harwood and Pyle (1980) and

Figure 6.14 *(Left)* Monthly mean modelled μ_{MCF} in ppt compared with μ_{MCF} from NOAA/ESRL flask observations. MCF_S is represented in blue, MCF_T is represented in red and observations are in black. *(Right)* $\Delta\mu_{MCF}$, the monthly change in μ_{MCF} in ppt.



further developments were made by Law and Pyle (1993). The model is Eulerian, and averages chemical species over longitude at a fixed latitude, altitude and time, up to an altitude of 90 km, although chemistry is only modelled up to 60 km. The model chemistry includes CH_4, OH, $O(^1D)$ and Cl, making it possible to derive monthly mean stratospheric CH_4 destruction rates from the model representation of chemical loss. The stratospheric destruction rates (SD) are shown in Figure 6.15. Loss rates are greatest in the NH in JJA and in the SH in DJF, since $O(^1D)$ and Cl are produced through the photolysis of O_3 and Cl_y photochemistry respectively. The loss rate given in SD is based upon the following reactions and reaction rates, which give a global mean loss rate of approximately 3.11×10^{-10} s^{-1}.

$$CH_4 + O(^1D) \longrightarrow CH_3 + OH \qquad [k_{O^1D} = 1.5 \times 10^{-10}] \qquad (6.8)$$

$$CH_4 + Cl \longrightarrow CH_3 + HCl \qquad [k_{Cl} = 7.3 \times 10^{-12} exp(-1280/T)] \qquad (6.9)$$

6.3.6 Assimilation of Satellite Data

A model simulation, CTL_A, was also carried out which was identical to the CTL simulation, but at each time step the model stratosphere was over-written with satellite data in order to provide a different stratospheric boundary condition, and to investigate the effect that changing the stratospheric concentration of CH_4 has at lower altitudes. Data was assimilated from the Halogen Occultation Experiment (HALOE) instrument from the period 1992-2003 and the ACE instrument afterwards. HALOE was launched in 1991 on board the Upper Atmosphere Research Satellite (UARS). The instrument uses solar occultation to measure vertical profiles of CH_4 in the middle

Figure 6.15 Zonal monthly mean CH_4 stratospheric destruction rates, SD, used in model simulations in Chapter 6 in s^{-1}.

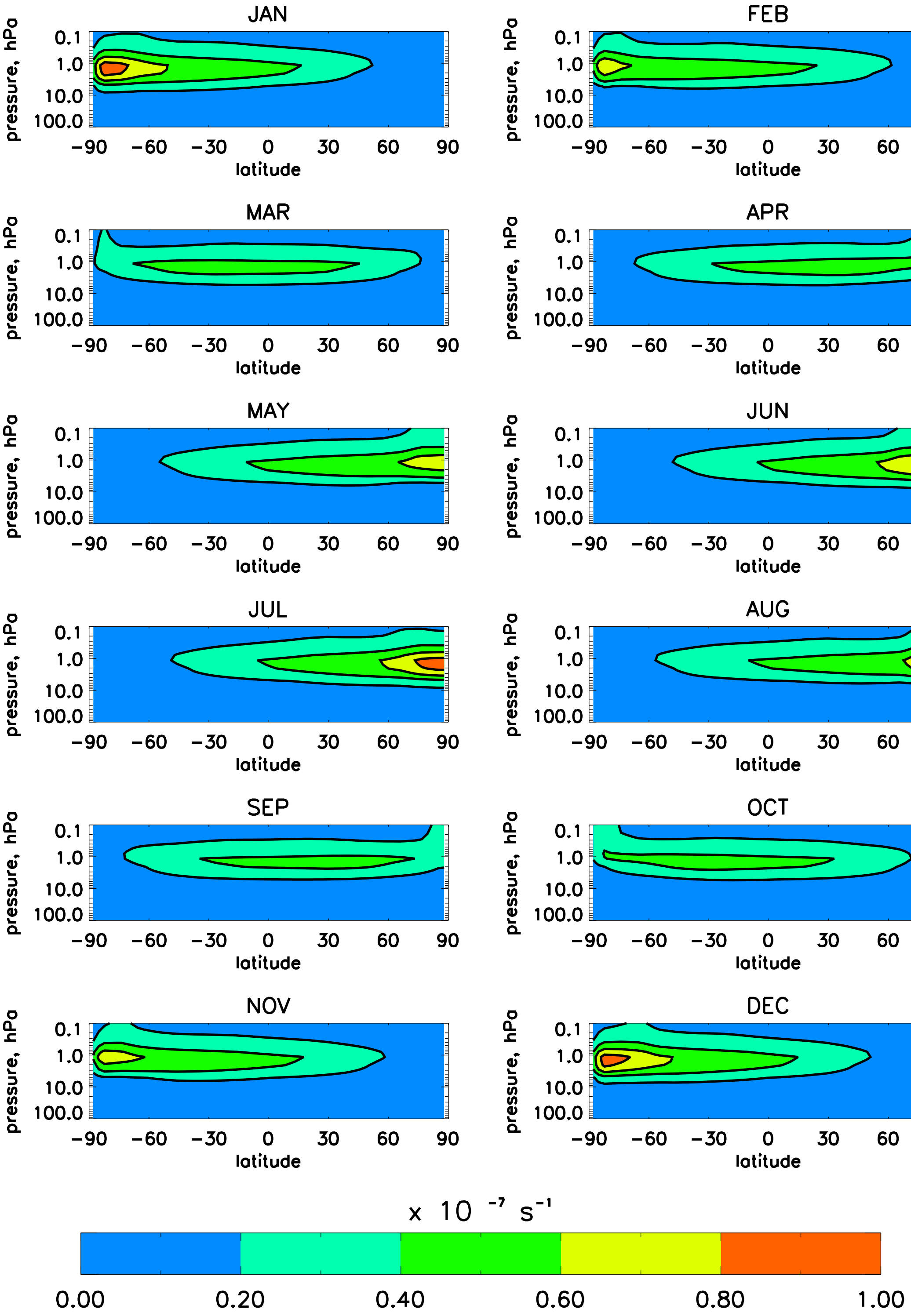
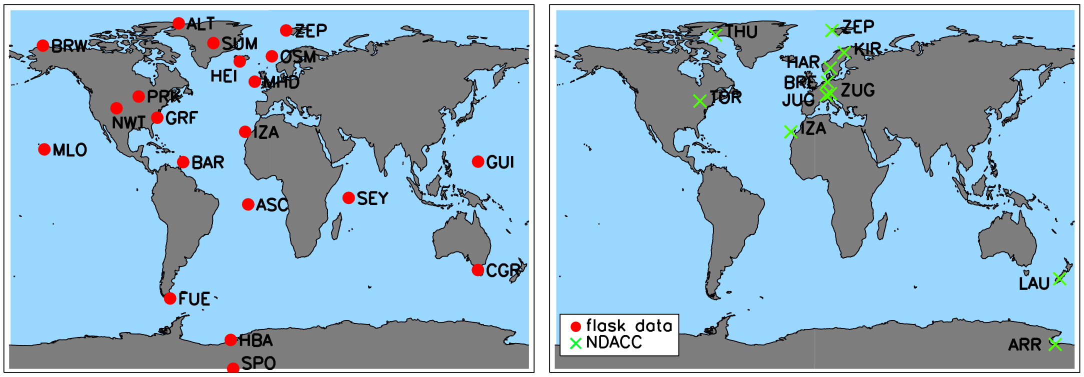


Figure 6.16 Locations of NOAA/ESRL flask sampling sites (*Left*) and NDACC FTIR sampling sites (*Right*). Flask sites are diaplayed as red dots and NDACC sites as green crosses.



and upper atmosphere with a 1.6 km instantaneous field of view at the Earth's limb. The ACE instrument is described in detail in Section 4.3. Although HALOE and ACE both give sparse coverage, the relatively long lifetime of CH_4 means that this should not affect the quality of the assimilation. The assimilation of the satellite observations into the TOMCAT model is described in Chipperfield et al. (2002), and was first used to assimilate long-term stratospheric observations of stratospheric trace gases from HALOE by Gunn (2008). This is the first time that the system has been used to assimilate data from ACE, however.

6.4 CH_4 Model Comparisons

Figure 6.17 shows a comparison between modelled and observed annual mean CH_4 mixing ratio at surface station sites for the period 1990-2008. Modelled data has been linearly interpolated to the surface sites from the model grid, except at MHD, which was treated as described in Section 4.3.3. Observed concentrations at each station show a period of growth from 1988 until 1999, after which the concentration remains approximately constant until 2005, when the concentration begins to increase again. This pattern was discussed in detail in Chapter 2. The period of levelling off is most clearly seen at SH stations such as CGR and SPO, where the mixing ratio is not dependent on local emissions. At these stations, each of the model simulations apart from CTL_E4 reproduces the pattern of CH_4 concentrations increasing throughout the 1990s before remaining stable in the early 2000s. The recent increase is not picked up by the model simulations due to the fact that most of the emission inventories do not vary beyond 2005. Only the VISIT wetland component varies annually until 2008, and emissions in this inventory produce only a small increase after 2005. The CTL_E4 inventory, which which has elevated emission rates after 2001 synchronised with Chinese economic growth, produces a rapid increase in CH_4 concentration in the 2000s, contrary to the observational record. This indicates that Asian emissions are unlikely to have

Table 6.4 Details of NOAA ESRL flask sampling sites and NDACC FTIR sites used for CH_4 comparisons in Section 6.4.

Station Code	**Station Location**	**Longitude (°)**	**Latitude (°)**	**Altitude (m)**	**Observation Type**
ALT	Alert, Canada	62.5W	82.5N	210	Flask
ZEP	Svalbard, Norway	11.9E	78.9N	475	Flask & NDACC
THU	Thule, Greenland	68.8W	76.5N	225	NDACC
SUM	Summit, Greenland	38.5W	72.6N	3238	Flask
BRW	Barrow, Alaska	156.6W	71.3N	11	Flask
KIR	Kiruna, Sweden	204.E	67.8N	419	NDACC
OSM	Ocean Station 'M'	2.0E	66.0N	5	Flask
HEI	Heimaey, Iceland	20.3W	63.4N	100	Flask
HAR	Harestua, Norway	10.8E	60.2N	596	NDACC
MHD	Mace Head, Ireland	9.9W	53.3N	8	Flask
BRE	Bremen, Germany	8.9E	53.1N	27	NDACC
ZUG	Zugspitze, Germany	11.0E	47.4N	2964	NDACC
JUG	Jungfraujoch, Switzerland	8.0E	46.6N	3580	NDACC
PRK	Park Falls WI, USA	90.3W	45.9N	868/470	Flask
TOR	Toronto, Canada	79.4W	43.7N	174	NDACC
NWT	Niwot Ridge, CO, USA	105.6W	40.1N	3523	Flask
GRF	Grifton, NC, USA	77.4W	35.4N	505	Flask
IZA	Izaña, Tenerife	16.5W	28.3N	2367	Flask & NDACC
MLO	Mauna Loa, HA, USA	155.6W	19.5N	3397	Flask
GUI	Guam	144.8E	13.4N	2	Flask
BAR	Rugged Point, Barbados	59.4W	13.2S	45	Flask
SEY	Seychelles	55.2E	4.7S	7	Flask
ASC	Ascension Island	14.4W	7.9S	54	Flask
CGR	Cape Grim, Australia	144.7E	40.7S	94	Flask
LAU	Lauder, New Zealand	169.7E	45.0S	370/390	NDACC
FUE	Tierra Del Fuego, Argentina	68.5W	54.9S	20	Flask
HBA	Halley Bay, Antarctica	26.5W	75.6S	33	Flask
ARR	Arrival Heights, Antarctica	166.7E	77.8S	200	NDACC
SPO	South Pole	24.8W	90.0S	2810	Flask

been increasing at the rate estimated in the EDGAR 4.0 database without a mitigating decrease in emissions elsewhere. The simulation which uses the VISIT inventory, which included a large wetland emissions increase in 1997 - 1998, produces a mixing ratio peak around the year 2000, before decreasing and finally levelling off in 2005. This modelled decrease is due to reduced JJA emissions in the VISIT inventory during these years.

In the NH, modelled concentrations are generally higher than observations, while the opposite is true in the SH. While this may be in part due to the emission distributions used in the model, Figure 4.8 indicates that model interhemispheric transport may be slower than observations suggest, which may also have an effect. Table 6.5 shows the root mean square difference (RMSD) between the model and the observations for each simulation for the periods 1990 - 1999 and 2000 - 2008. The RMSD is separated into the period covering the increase in concentration during the 1990s and the stable period during the 2000s, and is defined as in Equation 4.1. RMSD varies according to time and station, but model bias is generally largest in the NH during the 1990s and smallest in the SH during the same period.

In the NH during the 1990s, the model bias is large at the Arctic stations ALT and BRW, where the lowest RMSD is approximately 10 ppb and the highest is 19 ppb at ALT and 30 ppb at BRW. The BB simulation produces the largest RMSD at both of these stations, and Figure 6.2 shows that this inventory has increased emissions in the NH during JJA, which these Arctic stations will be sensitive to. RMSD at NIW and MLO during 1990 - 1999 is smaller, ranging between 4 and 15 ppb. At SH stations SEY, CGR and SPO, the range of RMSD is smaller still. While the CTL_E4 and INV simulations produce SH mixing ratios which are 8 - 13 ppb lower than observations, the BB inventory is within 5 ppb of observations at all SH stations. The fact that the BB inventory performs well in the SH and poorly in the NH, when combined with the fact that the TOMCAT model's IHD was found to be higher than inferred from observations in Chapter 4, indicates that the NH biomass burning estimates in JJA in BB are likely to be too high.

During the period 2000 - 2008, the CTL_E4 inventory produces large RMSDs at each station as it does not stabilise in the same way as the observations. Among the other simulations in the NH, RMSDs decrease at ALT, BRW and MLO as the modelled and observed concentrations level out within 15 ppb of the observations while at NIW the simulations underestimate the steady mixing ratio by around 11 - 17 ppb. At each of these stations, the WL_BB simulation produces the smallest RMSD, although since the inventory uses cyclic wetland emissions throughout this period, this is not due to any variation in the emissions that were not already captured in the BB inventory. The total emissions during this period, which are approximately 5 Tg/year larger than BB, may be more realistic, however. In the SH, the BB simulation again produces the smallest RMSD, stabilising with 3 - 6 ppb of the observations at each station, while other simulations are within 7 - 17 ppb of the observations.

Figure 6.17 Annual mean modelled and observed CH_4 mixing ratio in ppb at surface station sites for the period 1990 - 2008. Coloured lines represent the emission inventory used in the model simulation, while the black line represents NOAA/ESRL surface flask observations.

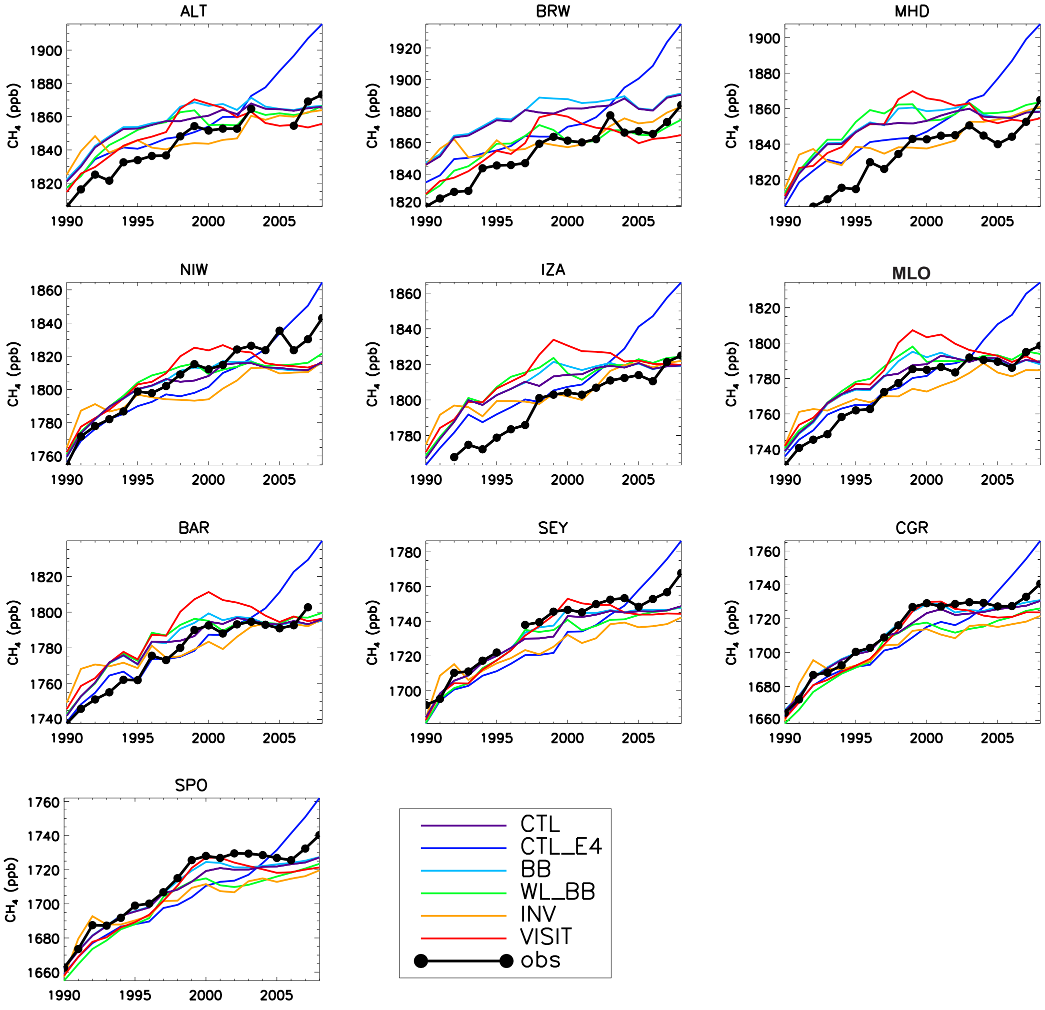


Figure 6.18 shows the annual mean surface level modelled mixing ratio of CH_4 for the year 2006 for each emission inventory, and the difference between the modelled and observed concentrations. 2006 was chosen to represent the last year in which the global CH_4 mixing ratio remained stable before increasing from 2007 onwards. By this point, the CTL_E4 simulated mixing ratios were up to 50 ppb higher than observations in the NH and around 10 ppb higher in the SH. Each of the other simulations displays similar global distributions with high mixing ratios close to emission regions with a NH - SH background concentration difference of approximately 150 ppb. Examining the model - observation differences reveals the influence of the variations in the emission inventories. Every simulation produces high concentrations at MHD, which must be due to either high European anthropogenic emissions in the EDGAR databases or over-estimated model transport to this station. Modelled concentrations are too high at the Arctic stations ALT, ZEP, OSM and HEI in each simulation. This may be due to high local emissions from Canada and Scandinavia. VISIT

produces mixing ratios closest to the observed data, which may be due to decreased JJA emissions in these two regions in the VISIT inventory. SUM, which is in the same region as the other Arctic stations, is underestimated in the model simulations. This is likely to be due to its high altitude of approximately 3200m, since it was posited in Chapter 4 that vertical transport in this region is too slow during SON. BRW mixing ratios are overestimated in simulations CTL, BB and INV and underestimated in WL_BB and VISIT, which both produce lower JJA emission totals local to BRW. In the SH, CTL and BB show the smallest model - observation difference, and both display increased mixing ratios in the South Atlantic and South-western Pacific Oceans due to increased emissions south of the inter-tropical convergence zone (ITCZ, described in Chapter 4) in South America and Indonesia.

Figure 6.19 shows the average monthly mean CH_4 at station sites for the period 2000-2006 for each emissions scenario, while Table 6.6 shows the Pearson correlation coefficients, r, between the modelled and observed seasonal cycle at selected stations. The seasonal cycle at a station is dependent upon the transport, atmospheric chemistry and local emissions. The model reproduces the observed seasonal cycle well at remote southern hemisphere (SH) sites. At SPO, all six emission scenarios reproduce the observed seasonal cycle, in which concentration is lowest in February and highest in September, with correlations to the observations of 0.995 or larger. However, the range of the seasonal variation is much higher in each of the model simulations than in the observations. The seasonal variation is due to the oscillation of tropospheric OH, and therefore of peak CH_4 destruction, between the SH in DJF and the NH in JJA. Results are similar at CGR, with model correlations with observations between 0.98 and 1.0, although the seasonal range is again larger in the model than in the observations. This increased model seasonal range in the SH may be due to OH concentrations being too high in the SH summer or too low in the SH winter. At SEY, the seasonal cycle is mainly due to the island's position near the ITCZ. As with the SF_6 measurements at TUT discussed in Section 4.4, the SEY station samples NH air with a high CH_4 concentration during DJF, and lower concentration SH air for the rest of the year. The data shows a second, smaller concentration peak in August and September which is due to the OH oscillation. The model correlation with observations is 0.99 or above in each simulation.

Stations in the NH display the opposite seasonal cycle to those in the SH, since CH_4 destruction due to OH is lowest during DJF and highest during JJA. NH concentrations of CH_4 therefore peak in the NH winter and are lowest in summer. At these NH stations the influence of atmospheric transport and emissions on the seasonal CH_4 cycle becomes greater. At NIW, IZA, MLO and BAR, the smooth cycle due to OH oscillation is disrupted by the influence of atmospheric transport, altering the seasonal cycle in different ways depending on the station's location. For example, comparing the seasonal cycle at MLO to that displayed for SF_6 in Figure 4.7, which is dependent only on transport from anthropogenic source regions, reveals that increased CH_4 concentrations

Figure 6.18 Annual mean modelled surface CH_4 mixing ratio in ppb for the year 2006 for six emission inventories. Triangles represent difference between model and NOAA/ESRL surface site flask observations. Upright triangles represent stations at which the model estimates are higher than observed, while inverted triangle represent stations where the model estimate is lower than observations. The size of the triangle indicates the magnitude of the model/observation difference. For station names, see Figure 6.16.

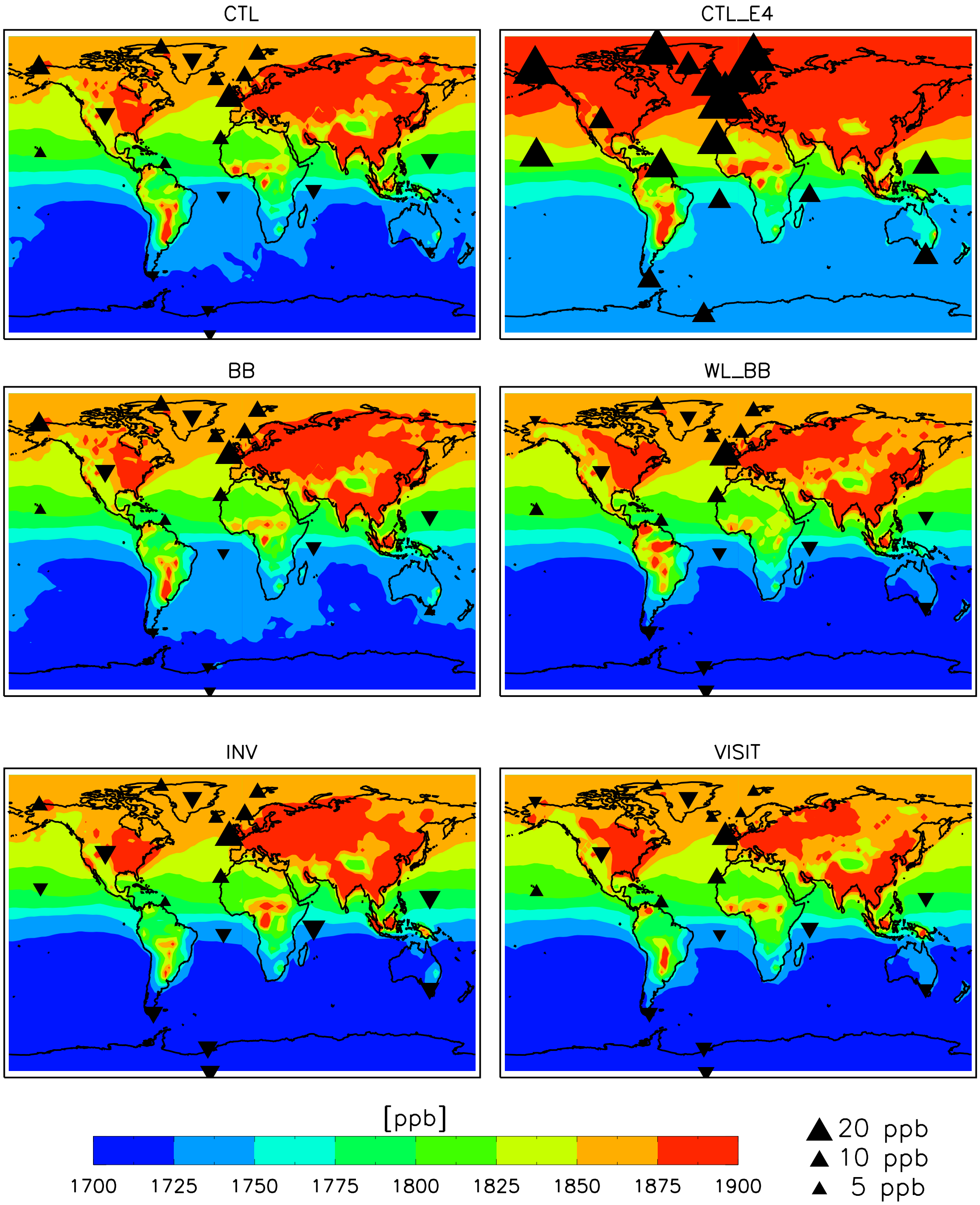


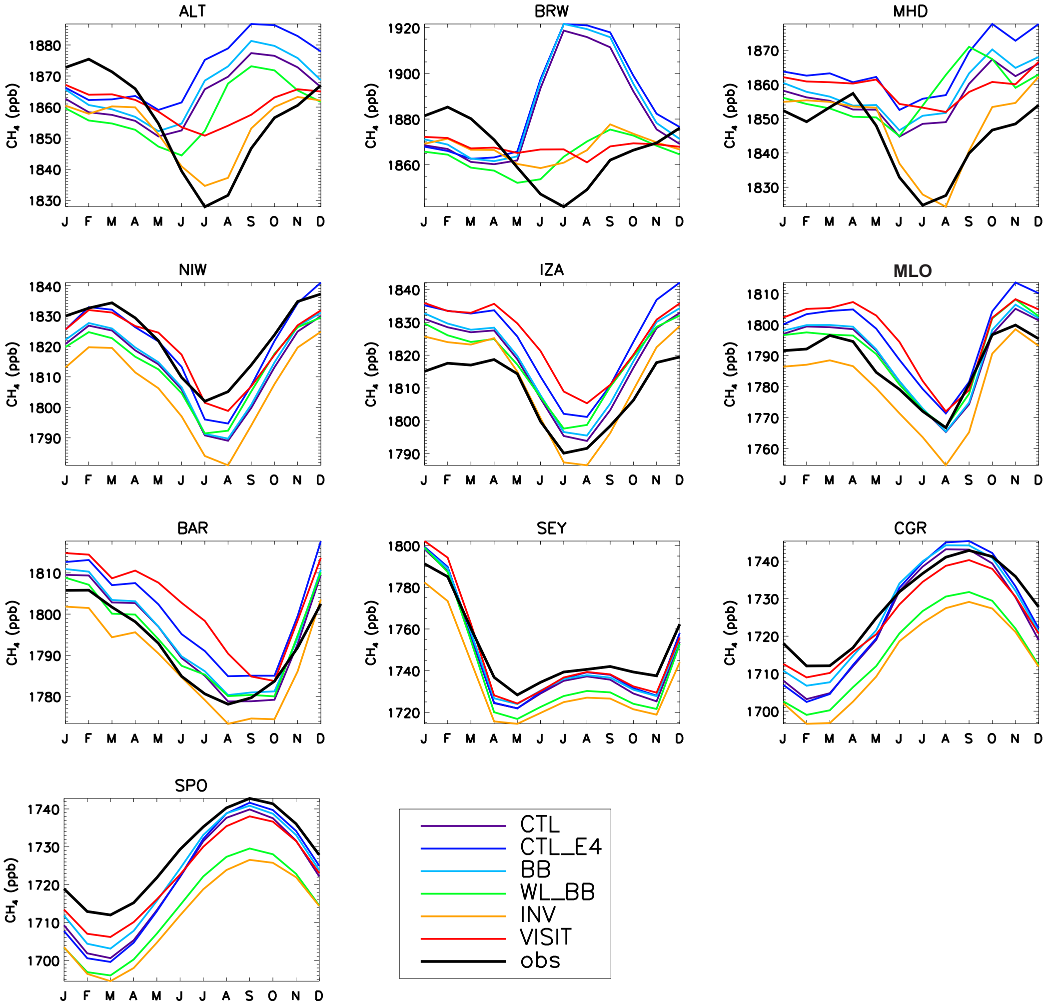
Table 6.5 RMSD for annual mean observed and modelled CH_4 concentration at selected surface sites for six emission inventories. Observations are NOAA/ESRL surface flask estimates. RMSD is given separately for the periods 1990 - 1999 and 2000 - 2008, and the smallest RMSD at each station is highlighted in bold.

STATION	LAT	LON	Inventory	RMSD 1990 - 1999	RMSD 2000-2008
ALT	82.4N	62.5W	CTL	17.309	8.110
			CTL_E4	**9.201**	23.702
			BB	19.137	10.887
			WL_BB	15.329	**4.257**
			INV	14.553	6.206
			VISIT	12.266	10.865
BRW	71.3N	156.6W	CTL	27.970	17.562
			CTL_E4	14.000	29.172
			BB	30.204	19.987
			WL_BB	11.989	**3.737**
			INV	19.380	5.776
			VISIT	**11.084**	9.133
NIW	40.1N	105.6W	CTL	5.706	12.671
			CTL_E4	7.846	11.496
			BB	**4.392**	12.693
			WL_BB	6.783	**11.262**
			INV	11.583	17.346
			VISIT	7.445	12.074
MLO	19.5N	155.6W	CTL	10.557	4.298
			CTL_E4	**5.175**	18.009
			BB	11.107	5.091
			WL_BB	14.125	**3.826**
			INV	11.436	8.130
			VISIT	15.695	11.170
SEY	4.7S	55.2E	CTL	7.037	6.477
			CTL_E4	13.709	12.128
			BB	**4.516**	**5.849**
			WL_BB	6.862	9.821
			INV	12.120	15.920
			VISIT	5.342	7.022
CGR	40.7S	144.7E	CTL	3.713	4.791
			CTL_E4	8.835	13.392
			BB	**2.154**	**3.235**
			WL_BB	7.250	12.330
			INV	8.183	14.840
			VISIT	4.717	5.777
SPO	90.0S	24.8W	CTL	5.146	7.378
			CTL_E4	11.090	13.890
			BB	**3.071**	**5.719**
			WL_BB	9.163	14.501
			INV	8.372	16.550
			VISIT	6.460	7.319

during March and April and the sharp minimum during August are both due to atmospheric transport rather than emissions or chemical destruction. Similar high concentrations at IZA in March and April are likely to be due to increased transport to the station from source regions. Model-observation correlations at these stations are 0.9 or above using every emission inventory, although a larger range is displayed by the six simulations as local emissions start to become an influence upon the seasonal cycle. The VISIT simulation is an outlier at these stations, producing concentrations that are too high in the NH spring at each of the four stations, possibly due to increased NH winter emissions in this inventory.

At the NH stations MHD, BRW and ALT the seasonal cycle is strongly dependent on all three of the contributing factors. As well as OH and transport dictating the seasonal variations, the influence of the timing and magnitude of local NH emissions is also important. The concentration at MHD peaks during April, although there are also high concentrations in December and January which are likely to be due to transport from anthropogenic source regions, as they also appear in the seasonal cycle of SF_6 (see Figure 4.7). The model generally does a poor job of reproducing the seasonal cycle at this station, although this is likely to be due to the distribution of CH_4 emissions used in each simulation due to the wide range of results produced. The INV simulation reproduces the correct seasonal cycle with a range similar to that of the observations, especially during the NH summer, while the CTL, CTL_E4, BB and VISIT simulations produce a weak cycle that does not have a minimum concentration during JJA. The modelled range is much smaller (∼15 ppb) in these simulations than the observed range (∼35 ppb). Finally the WL_BB simulation's seasonal cycle peaks in September and has a minimum in May and June, putting it almost exactly out of phase with the observations. This is likely to be due to the influence of increased wetland emissions in the USA during JJA in the WL_BB inventory. At both BRW and ALT, CH_4 concentrations peak in February and reach their minimum during July with smooth transitions between the two, aside from a sharp concentration increase in September. Model simulations are again largely dependent upon the emission distribution used. In Section 4.4 it was suggested that model vertical transport in the Arctic was too strong in DJF and not strong enough during September and October, so we would expect concentrations to be low in NH winter and high in autumn. The INV simulation performs well, producing correlations of 0.908 at ALT and 0.505 at BRW, although the seasonal variations at BRW are not large enough, and concentrations decrease throughout November and December, when the observed concentrations increase. At BRW, the VISIT simulation produces a seasonal cycle with very little variation apart from a slight decrease in August, while the WL_BB inventory is out of phase with the observations by around three months, as it peaks in September and reaches its minimum in May. At ALT, the VISIT inventory reproduces the correct seasonal cycle with a correlation of 0.935, although the magnitude of the seasonal variations are far too small. However, the CTL, CTL_E4 and BB inventories perform extremely poorly at both stations, producing a seasonal cycle which is completely out of phase with the observations. Correlations

Figure 6.19 Monthly mean modelled and observed CH$_4$ concentration in ppb, averaged over the period 2000 - 2006, at surface station sites for six emission inventories and NOAA/ESRL flask observations. Lines are coloured as in Figure 6.17.



range from -0.214 to -0.320 at ALT and from -0.760 to -0.773 at BRW, indicating inaccuracies in the emission distributions in the NH. Examining the seasonal cycle at these NH stations indicates the importance of having the correct geographical and temporal distribution of emissions in the model, and the influence of emissions on Arctic methane concentrations will be studied in more detail in Section 7.2.

In order to examine the distribution of CH$_4$ at higher altitudes, modelled CH$_4$ was compared to data from the Infrared Working Group (IRWG), part of the Network for the Detection of Atmospheric Composition Change (NDACC (http://www.ndsc.ncep.noaa.gov/)) which provides information on total column CH$_4$ at a number of station sites using Fourier Transform Spectrometers recording direct solar spectra in the near-infrared spectral region to produce total column CH$_4$. Figure 6.16 shows the location of NDACC station sites.

Table 6.6 Correlations for monthly mean observed and modelled CH_4 concentration at selected remote surface sites averaged over the period 2000 - 2006 for each of the six emission inventories.

		Station						
		ALT	BRW	NIW	MLO	SEY	CGR	SPO
Inventory	CTL	-0.214	-0.768	0.976	0.966	0.990	0.984	**0.999**
	CTL_E4	-0.320	-0.773	0.972	0.966	0.993	0.989	0.997
	BB	-0.230	-0.760	0.978	0.969	0.992	0.982	**0.999**
	WL_BB	-0.008	0.082	**0.982**	**0.978**	**0.995**	0.998	**0.999**
	INV	0.908	0.505	**0.982**	0.976	0.993	**0.998**	0.995
	VISIT	**0.935**	**0.689**	0.931	0.909	0.992	0.989	**0.999**

Figure 6.20 Monthly mean observed and modelled total column CH_4 in molecules cm^{-2} at NDACC FTIR sites. Lines are coloured as in Figure 6.17. Shaded area represents observational error range supplied with data.

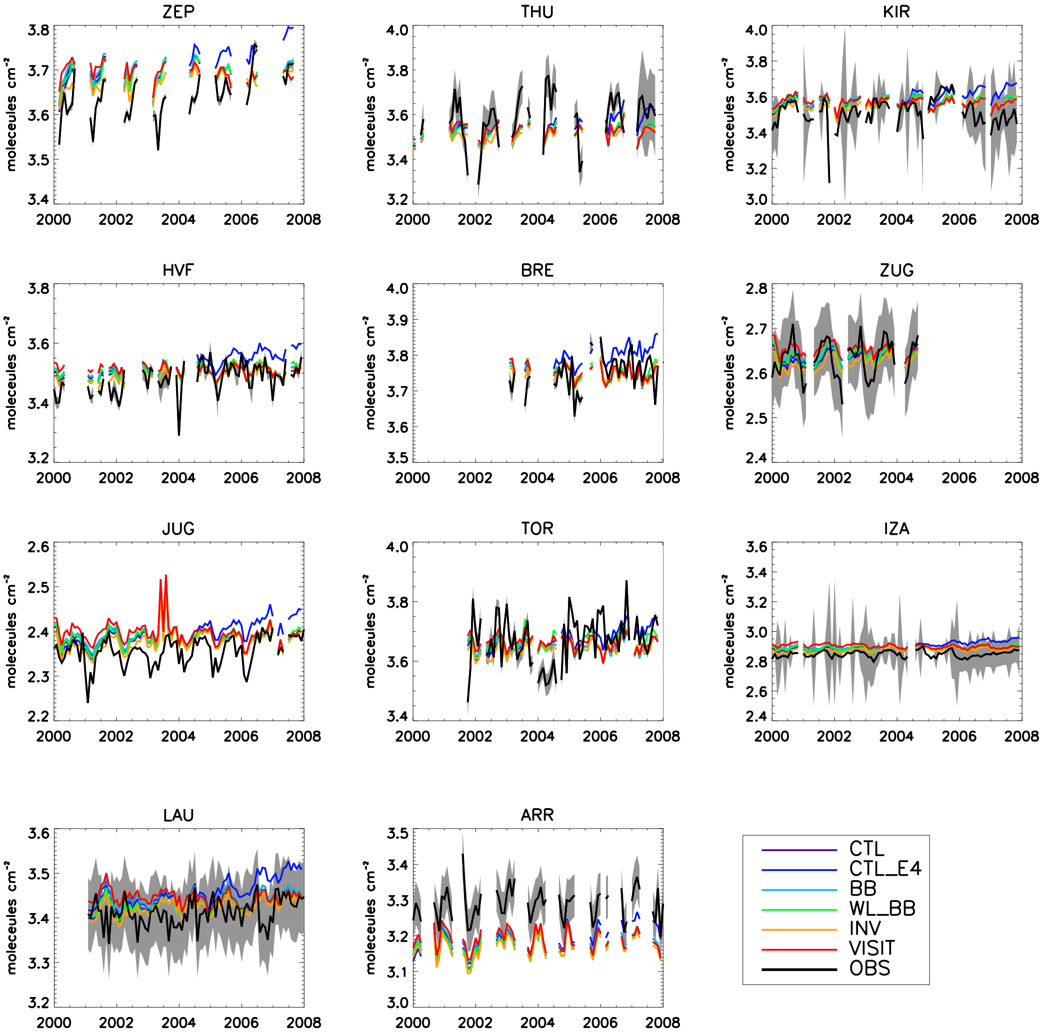


Figure 6.20 shows comparisons of TOMCAT total column CH_4 for the six emission inventories and observations from NDACC stations for the period 2000-2008. Since FTIR spectrometers require daylight in order to take measurements, only summer-time measurements are available for stations at polar latiudes (ZEP, THU, ARR). Many of the NDACC stations (KIR, HAR, BRE, ZUG and JUG) are in Western Europe, with sparse coverage elsewhere. There is little difference between the total column CH_4 predicted by the model using most of the emission inventories, indicating that the geographical distribution of emissions has little effect on the seasonal cycle of the total column. The CTL_E4 inventory does produce a significantly larger column mixing ratio than the other inventories in the late 2000s. At each station the observed total column CH_4 is well matched by the model, which is consistently within the observational error range, although modelled values are high at the Arctic station ZEP and low at the Antarctic station ARR, which is likely to be due to the slow interhemispheric transport of the model. Modelled values are within 0.1×10^{19} molecules cm^{-2} of observations at each of the stations in the tropics and mid-latitudes, and produce excellent matches at each of the European stations, TOR, IZA and LAU. However, the timing and, especially, the magnitude of seasonal variations at many of the stations is not reproduced well in the model, indicating that modelled high-altitude transport is too homogeneous.

Figure 6.21 shows a comparison of stratospheric CH_4 in CTL, the model simulation using the CTL emission inventory and the Spivakovsky OH field, and CTL_A, which assimilates stratospheric mixing ratios from the HALOE and ACE instruments into the simulation. CTL annual mean CH_4 mixing ratio is comapred to CTL_A for the years 1998, when HALOE data is assimilated into CTL_A, and 2004, when ACE data has been assimilated. In both years, the concentration of CH_4 in the upper stratosphere (1.0 - 0.1 hPa) in CTL is lower by up to 400 ppb than in CTL_A, indicating that the stratospheric rate of vertical mixing in the TOMCAT model simulation is too slow, or that the stratospheric destruction rate due to OH, $O(^1D)$ or Cl is too high. Monge-Sanz et al. (2007) found that the use of ERA-Interim ECMWF winds in the TOMCAT model, which are used in all of these simulations, performed well in comparison with age-of-air measurements. This indicates that the Brewer-Dobson circulation, which is the large-scale motion responsible for stratospheric transport of air, is modelled well. It is likely, therefore, that the rate of stratospheric loss of CH_4 is too high in the upper stratosphere. There are few model levels at pressures less than 10 hPa, and this may account for the discrepancies. However, a key point is that both the HALOE and the ACE assimilations produce low CH_4 values at the poles at 100 hPa compared with the CTL simulation. There is large-scale subsidence of air into the troposphere at mid-high latitudes (Holton et al., 1995), and decreased CH_4 mixing ratios above the polar tropopause, as seen in CTL_A, will decrease tropospheric CH_4. Figure 3 in Patra et al. (2011) shows a similar result for a range of CTMs which took part in the TransCom CH_4 intercomparison. Each model used the same OH_S and SD fields that are used in this study, and the majority of them produce high CH_4 concentrations at 50 - 100 hPa in the NH compared with observations from the ACE satellite.

Figure 6.21 Annual mean stratospheric methane mixing ratio in ppb for CTL (*Top Left*) and CTL_A (*Top Right*) for 1998 (with assimilated HALOE data) and for CTL (*bottom left*) and CTL_A (*bottom right*) for 2004 (with assimilated ACE data).

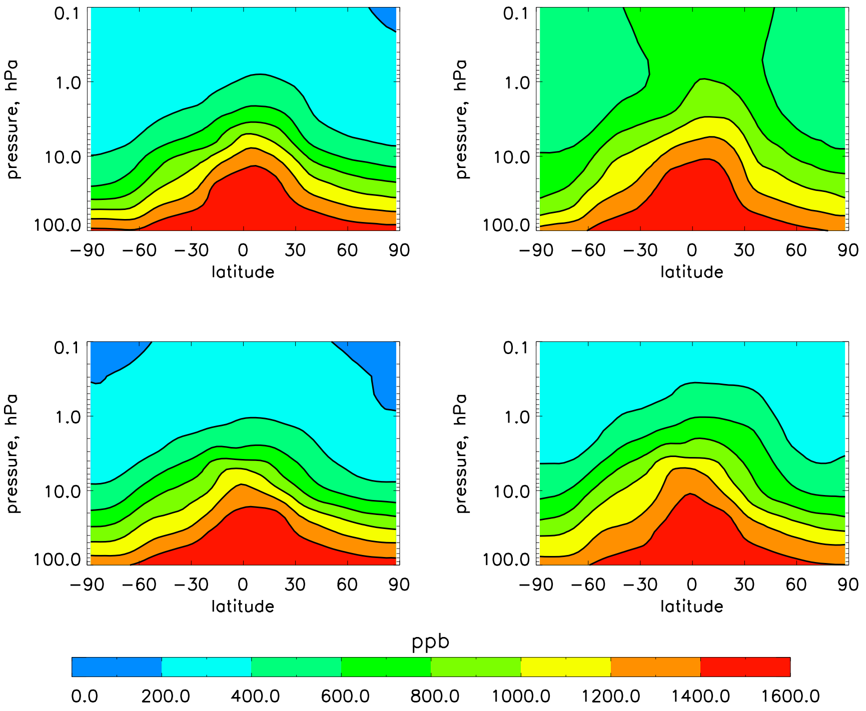


Figure 6.22 shows the annual mean CH_4 concentration at surface sites from NOAA flask measurements and for model simulations CTL, CTL_T and CTL_A. CTL_T, which has approximately 10% less tropospheric OH than CTL, does not produce the correct CH_4 mixing ratio at the surface, and by the end of the simulation period over-estimates surface CH_4 by approximately 100 ppb, having not yet reached the steady state in the 2000s, unlike the observations and the CTL simulation. This indicates that the tropospheric sink in CTL_T does not balance the emissions in the second half of the studied period in the same way that the OH_S field does. Since the OH_T field is produced from the TOMCAT model with the full chemistry scheme, this implies that there some sources or sinks of OH in that model are incorrect. Comparison with the OH_S field suggests that the discrepancy may be in the SH in DJF. Discussing the reasons for this inconsistency is beyond the scope of this study, but it seems that if the total CH_4 emission estimate of 500 - 580 Tg year^{-1} is correct, then the OH field produced by the full chemistry TOMCAT model is inaccurate.

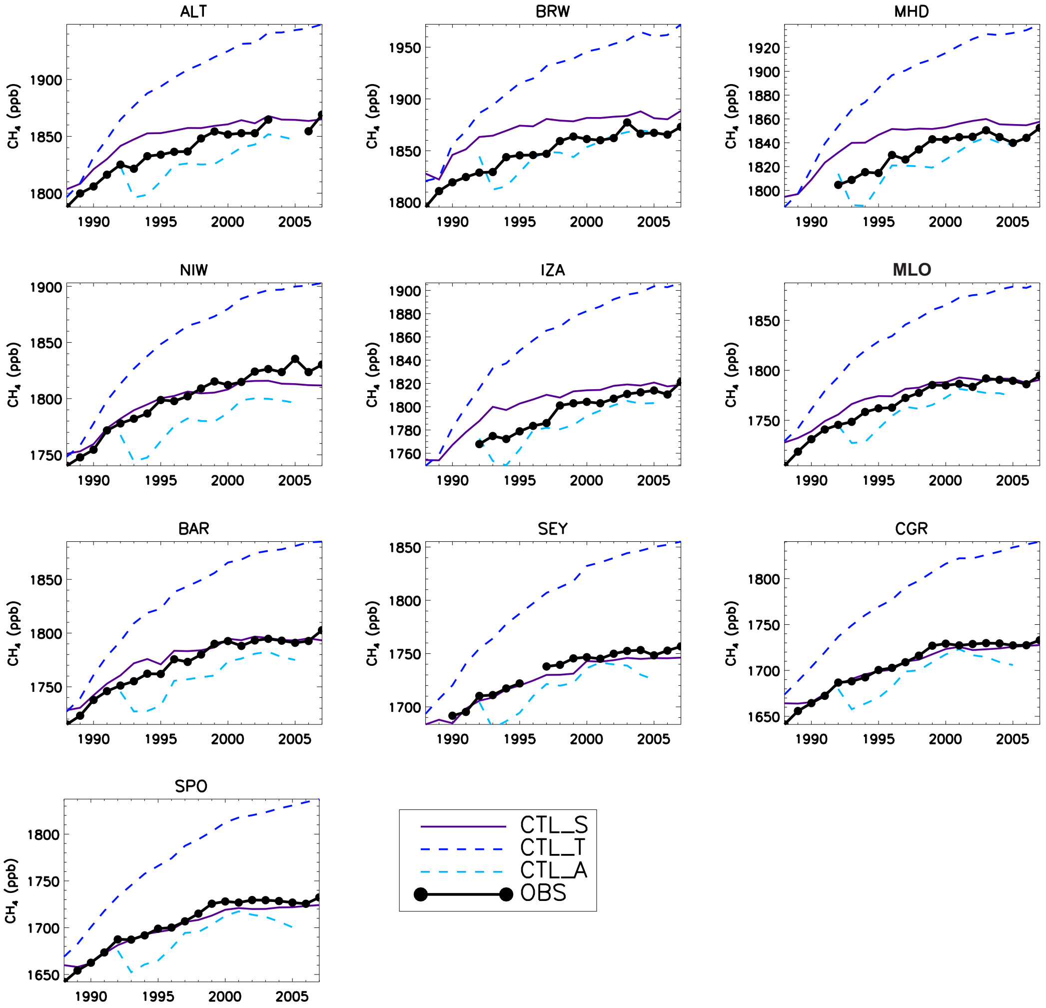
There are a number of possible causes for the discrepancies between the OH_T and OH_S distributions. As discussed in Section 6.3.1, OH concentrations are driven by the photolysis of ozone (O_3) and the reaction of the excited $O(^1D)$ molecule with water vapour. OH is lost through reactions with other trace species such as CH_4 and CO. The fact that the OH_T distribution is much more weighted towards the surface and the NH than that of OH_S implies that there is some discrepancy between the two models which were used to produce the two distributions. In TOMCAT, the atmospheric H_2O distribution is read in from ECMWF re-analyses, which are likely to be accurate. However, there may be inaccuracies in the photolysis scheme or the O_3 distribution. The O_3 distribution is in turn dependent upon the model's ability to simulate tropospheric NO_x. The distribution of OH in the TOMCAT model is therefore sensitive to the accuracy of the anthropogenic and soil emissions of NO_x and also to the model representation of atmospheric NO_x emissions due to lightning. Furthermore, errors in the model's vertical transport may also be responsible for the high OH concentrations towards the surface. As discussed in Chapter 3, vertical convection in the TOMCAT model is known to occur at too slow a rate, and does not extend up to a high enough altitude, which may mean that the controlling chemistry for OH occurs too close to the surface. Further investigation is necessary in order to assess the accuracy of the complex processes which control the model's OH distribution, although it is beyond the scope of this thesis.

The CTL_A simulation produces mixing ratios around 10-20 ppb lower than CTL, illustrating the effect that stratospheric air can have at the surface as it subsides at the poles. Surface CH_4 concentrations in the CTL_A simulation rise after 1993 until 2001, when they begin to decrease slightly, while the observations stabilise. The CTL_A simulation concentrations are more consistent with observations than CTL at BRW and MHD, but are generally 10 - 20 ppb lower elsewhere, especially in the SH. Assuming that the CTL_A simulation stratosphere is more consistent with observations than that of the CTL simulation, this implies that some tropospheric sink of CH_4 may in fact be too large or that emissions may be too small, although it is not possible to say which.

Figure 6.23 shows the mean seasonal cycle at surface sites for CTL, CTL_T and CTL_A compared with NOAA flask data over the period 2000-2006. Leaving the various biases of the three simulations aside, the seasonal variation of each modelled concentration is extremely similar, indicating that neither the distribution of the model OH nor the influence of stratospheric air has a significant influence on the seasonal cycle of surface methane, which is thus largely dependent on transport and emissions. Although the OH_T field has lower OH concentrations throughout much of the year, the NH-SH oscillation of OH is captured, and so the seasonal cycle at those stations at which OH is the contributing factor is unchanged, although the annual range is slightly altered at some locations.

Figure 6.24 displays statistics on the rate of CH_4 destruction and the balance between destruction and emissions in the model simulations CTL and CTL_T. Figure 6.24(a) shows the annual atmospheric burden increase for the two simulations, confirming that the destruction and emission of CH_4 fluctuates around a steady state from 2001 onwards in the OH_S simulation, while the atmospheric burden is still increasing by around 10 Tg year^{-1} in 2008 when the OH_T field is used. The burden increase of the CTL simulation compares favourably with the annual growth rate shown

Figure 6.22 As Figure 6.17, but for runs CTL, CTL_T and CTL_A.



in Figure 2.6(b), which also fluctuates around zero after 2001 (note the different units of the two plots). Figure 6.24(b) shows that in each simulation, the amount of CH$_4$ destroyed as a percentage of the total atmospheric burden does not change over the two decades simulated, remaining at approximately 11.1% of the total burden for CTL and 10.4% for CTL_T. Variations in this value will be due to changes in temperature and emissions. Figure 6.24(c) shows the annual global mean CH$_4$ lifetime, τ in each simulation, which is estimated as

$$\tau = \frac{\sum M_{CH_4}}{\sum k_{OH}[OH](M_{CH_4})} \tag{6.10}$$

where M_{CH_4} is the mass of CH$_4$, [OH] is the OH concentration in molecules cm^{-3} and k_{OH} is reaction rate of CH$_4$ with OH used in the model. τ has a mean value of 9.9 years for the CTL simulation and 10.8 years for the CTL_T simulation. This is close to the model-mean CH$_4$ lifetime of 9.99 years estimated by Patra et al. (2011), in which eleven CTMs used the OH_S field and CTL

Figure 6.23 As Figure 6.19, but for CTL, CTL_T and CTL_A.

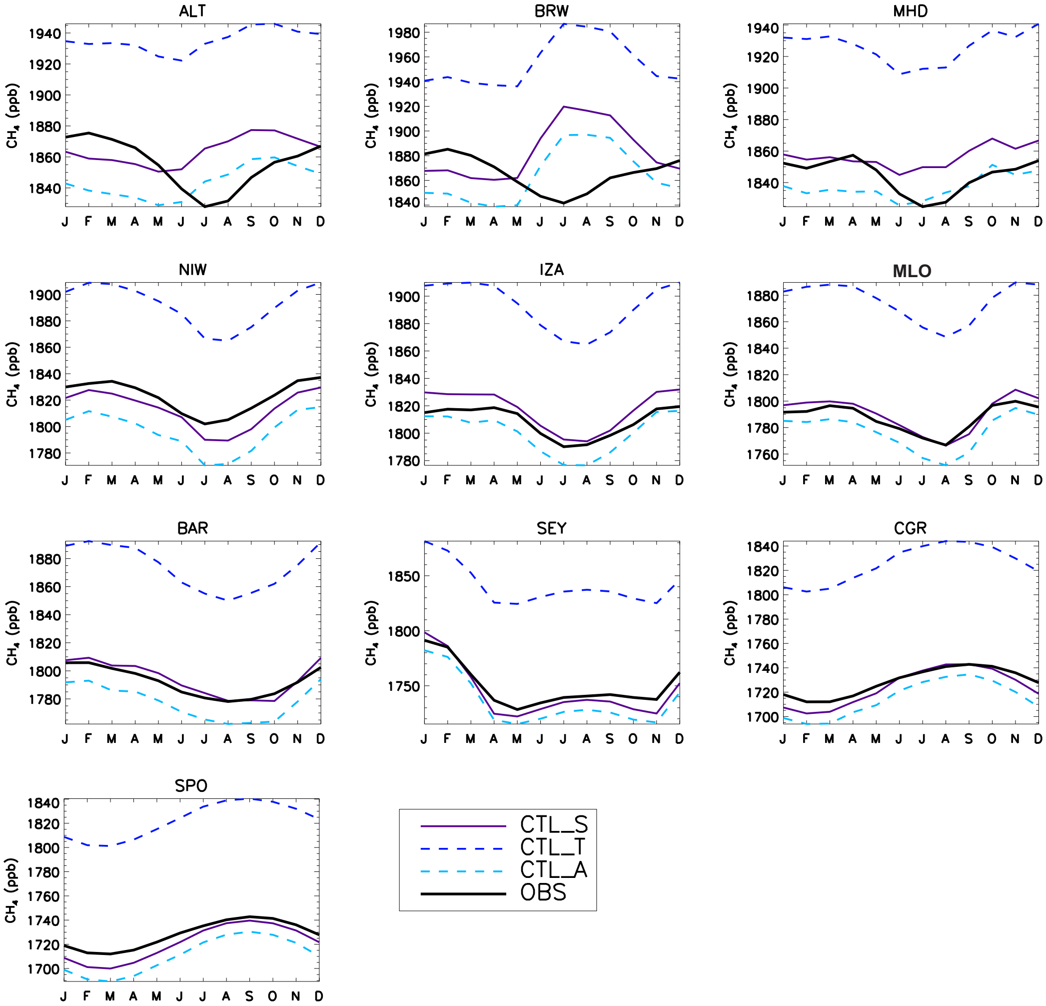
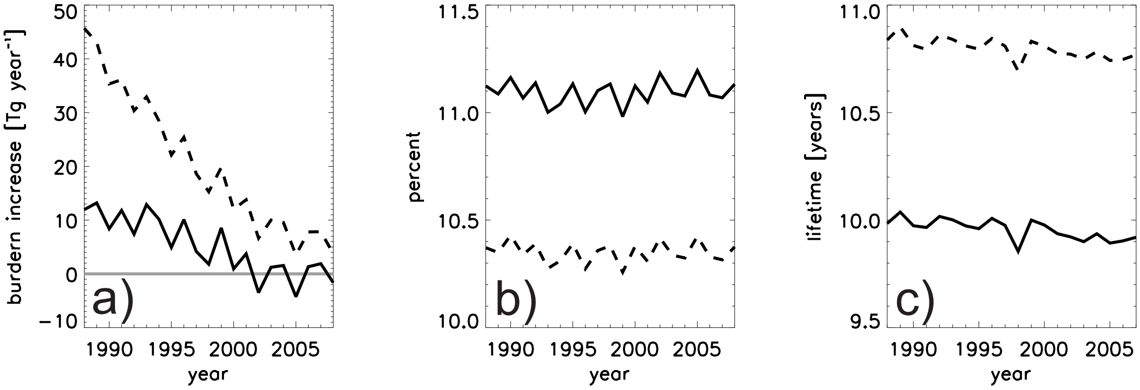


Figure 6.24 Annual CH_4 burden increase in Tg year^{-1} *(Left)*, annual CH_4 destruction as a percentage of total atmospheric CH_4 burden *(Centre)* and CH_4 life time in years *(Right)* for the period 1988 - 2008 for the model simulations CTL (solid line) and CTL_T (dashed line).



emission inventory to simulate tropospheric CH_4 concentrations. It is also within the range of 7.79 - 10.25 years found by Lawrence et al. (2001) in a study of the effect of different OH fields upon CH_4 lifetime, and it is just within the estimate of 8.7 ± 1.3 years years supplied in the 2007 IPCC report (Solomon et al., 2007) Note that the CH_4 lifetime decreases by approximately 0.1 years during 1998 due to the strong El Niño that year, and that, comparing with Figure 6.24(a), this small decrease in the lifetime leads to around a 9 Tg increase in the atmospheric burden compared with the surrounding years 1997 and 1999.

6.5 Summary

This chapter examined recent variations in the atmospheric CH_4 budget through the comparison of the results of model simulations to observed data. It was found that it was possible for the atmospheric growth rate of the species to decrease during this time due to a decrease in the growth of surface emissions of the species, regardless of whether there were changes in the loss rate of the species.

Although the two OH fields used in this comparison performed similarly in producing the correct loss rate of MCF, the OH_S field was more consistent with observations and produced CH_4 concentrations similar to those seen at surface station sites and in total column observations. The OH_T field was found to have the incorrect distribution in comparison with observations, and produced model results which overestimated CH_4 concentrations by up to 100 ppb.

It was found that assimilation of satellite observations of stratospheric CH_4 produced significant effects on surface CH_4 concentrations. Assimilation from the HALOE and ACE satellites reduced surface concentrations by around 10 ppb, due to subsidence of stratospheric air through the tropopause at the poles. indicating the importance of accurate estimation of the stratospheric CH_4 sink despite its relatively small size compared to the OH sink.

A number of emission inventories were used in order to examine the effect of emissions on the atmospheric CH_4 budget. Many of the emission inventories captured the decline the atmospheric growth rate of CH_4 well. However, the EDGAR V4.0 inventory of anthropogenic emissions, which estimated that CH_4 emissions rose throughout the 2000s in line with Asian economic growth, significantly overestimates the growth rate of CH_4 throughout the last decade. This implies that for the estimated growth of anthropogenic CH_4 emissions to be correct, there must have been a corresponding decrease in some other emission process to mitigate the growth. Bousquet et al. (2006) suggested that wetland emissions may have decreased during the early years of the 2000s, which may be the necessary mitigating factor. The VISIT model, however, produced CH_4 concentrations which decreased throughout the 2000s, indicating that CH_4 emissions were larger than 530 Tg year^{-1} during the last decade.

The TOMCAT model captured the seasonal cycle of CH_4 at stations south of 50°N well, with correlations greater than 0.9 with all inventories. However, at stations north of 50°N, some of the model simulations perform poorly. This is likely due to discrepancies in the emission budgets, since some simulations do capture the seasonal variation well, and this is further investigated in Chapter 7.

Chapter 7

Revised Estimates of Northern Hemisphere High Latitude CH_4 Emissions

7.1 Introduction

This chapter investigates the seasonal cycle of CH_4 in the Arctic produced by the TOMCAT model using the emission inventories described in Chapter 6. Some of these inventories produced elevated atmospheric concentrations during the NH summer at Arctic stations. Since the simulated transport in TOMCAT has been validated using SF_6 observations, and changes to the chemical sinks had little effect on the observed seasonal cycle, this indicates that inconsistencies in the emissions are likely to be responsible for the discrepancy. It should be noted, however, that the two OH fields examined in Chapter 6 are fairly similar at high latitudes, as shown in Figure 6.11. The possibility that the seasonal variability of Arctic OH may not be modelled correctly therefore does exist. This chapter attempts to locate the source of the error using various forward and adjoint modelling techniques, before using the TOMCAT 4D-Var inverse model to produce an updated emission inventory which is consistent with observations. Section 7.2 examines the seasonal cycle of CH_4 at a number of Arctic stations, and examines the likelihood that emissions are responsible for the discrepancies between the model output and the observations, while Section 7.3 uses the adjoint version of the TOMCAT model, which was described in Chapter 5, to examine the sensitivity of the Arctic to transport of emissions from other regions. Section 7.4 uses the TOMCAT forward model to investigate the influence of emissions from large NH wetland areas on the Arctic CH_4 concentration, while Section 7.5 describes the results of an inverse simulation carried out with the aim of constraining surface CH_4 emissions and improving the consistency between the model simulations and observations in the Arctic. Section 7.6 summarises the findings of this chapter.

7.2 Seasonal cycle at Arctic Stations

Chapter 6 found that although the TOMCAT model can accurately reproduce the inter-annual variation of CH_4 in the troposphere, there are substantial discrepancies between observations and simulations which point to substantial misattributions of surface emissions. Section 6.4 showed that different emission inventories produced large variations in the seasonal cycle of CH_4 mixing ratio at a number of NH surface sites, and that some inventories produced modelled concentrations which were completely out-of-phase with a site's observational record. In this chapter model simulations will be carried out with the aim of closely examining the relationship between the seasonal cycle of the mixing ratio of CH_4 at NH polar station sites and the magnitude and timing of surface CH_4 fluxes. Investigating how differences in the emission inventories lead to discrepancies in the atmospheric concentration can help us to improve our surface flux estimates.

The major contributing factors to producing the distinctive CH_4 seasonal cycle at a specific station are the atmospheric transport, the seasonal oscillation of the OH field, and the magnitude and timing surface fluxes of the species, all of which must be accurately simulated in the model in order to reproduce the observed concentrations. Examination of the modelled seasonal cycle of SF_6 in Section 4.4 indicated that model transport is unlikely to be the cause of the poor representation of CH_4 seasonal cycle at these stations, as modelled SF_6 does not display discrepancies to the same extent as those produced in the CH_4 simulations. Changes in the distribution and magnitude of CH_4 destruction due to OH and changes to the stratospheric loss in Section 6.4 produced no alteration in the seasonal cycle, indicating that the problem is also unlikely to be dependent on the chemistry used in the model. However, the two OH fields examined in the previous chapter displayed similar distributions at Arctic latitudes, meaning that further investigation into the effect of OH on the seasonal cycle of CH_4 may be necessary before chemistry can be discounted as a contributing factor. It is likely, however, that the distribution, timing or magnitude of the CH_4 emission inventories may be the root of the problem. This hypothesis is backed up by the fact some scenarios perform much better than others, with the INV inventory, which was developed using assimilated data from these stations, producing output which was much more consistent with observations. The VISIT inventory also produced the correct seasonal cycle at the NH stations, although the range of the annual variation was usually smaller than in the observations.

In order to confirm that the problem was indeed due to inconsistencies in the emission inventories, two TOMCAT CH_4 simulations were carried out which removed the seasonal cycle of the emissions globally (CTL_MN) and north of 60°N only (CTL_MN60). The total annual CH_4 emissions in each of these simulations was the same as in CTL. These simulations were otherwise identical to CTL. Figure 7.1 shows the seasonal cycle produced by these simulations at surface sites along with CTL and NOAA flask data. At remote SH stations such as CGR and SPO, and as far north at SEY, there was no change to the monthly mean distribution of CH_4, confirming that the seasonal cycle at these stations relies almost entirely upon the quality of the long-range model transport and chemistry, with little contribution from emissions, which are mixed out and have no local impact. At stations in tropical and mid-latitudinal locations such as NIW, IZA, MLO and BAR, differences appear in CTL_MN. These stations are not highly sensitive to changes in emissions north of 60°, and so CTL_MN60 displays little deviation from CTL at these stations. CTL_MN does display differences, however, indicating that these stations are dependent on the accuracy of the emission inventory. At IZA, for example, there is a large increase in CH_4 during MAM in the CTL_MN inventory, signifying that the station usually samples air from a location with decreased CH_4 emissions during these months. MHD, at a latitude of 53.3°N, is sensitive to emissions north of 60°, as the seasonal cycle here in the CTL_MN60 simulation is different to that of CTL, displaying a minimum during August and a maximum during May. This is much closer to the seasonal cycle displayed by the observations at this station, which have an April maximum and a July minimum. This suggests that inaccurate emissions in the inventory are degrading the CTL simulation.

Figure 7.1 also displays results for six Arctic stations, each of which show a similar observed seasonal cycle. The observed seasonal cycle, and that produced by CTL, are similar at each of ALT, ZEP, OSM and HEI, with the observations peaking in January and reaching a minimum during July while the model produces maxima in September or October and minima around June. Removing the seasonality of the model emissions above 60° produces a cycle which is much closer in distribution to that of the observations, confirming that polar NH emissions are responsible for the incorrect seasonality in CTL. At BRW the same effect is seen, as a large peak in CH_4 concentration in July in CTL becomes a minimum in CTL_MN60. SUM is a special case due its high altitude of 3200m. This means that the station is less directly reliant on the seasonality of emissions than the other polar stations, with vertical mixing in the boundary layer being a more important factor. The seasonality is still shifted slightly at this station, however. Table 7.1 shows the correlation between the observations and the CTL, CTL_MN, CTL_MN60, INV_S and VISIT_S simulations. At each station apart from SUM, the CTL simulation produces either a very small or a negative correlation with the observations. At these stations, however, both the CTL_MN and CTL_MN60 simulations produce a correlation greater than 0.7, indicating both these stations are highly sensitive to changes in local emissions and that, in reality, local NH summertime emissions may be too large, creating discrepancies in the seasonal cycle of CH_4 mixing ratio at sites above 60°N. At SUM, removing the seasonality of the emissions reduces the model correlation with the emissions, which was already close to 0.6. The INV_S and VISIT_S may provide better estimates of these Arctic emissions, as they produce high correlations with the

Table 7.1 Correlations for seasonal cycle of observed and modelled CH$_4$ concentration at selected remote surface sites for CTL, CTL_MN and CTL_MN60, INV_S and VISIT_S. The highest correlation for each station is shown in bold.

		Station					
		ALT	ZEP	SUM	BRW	OSM	HEI
Inventory	CTL	-0.214	0.058	0.581	-0.768	0.078	0.126
	CTL_MN	0.842	0.981	0.300	**0.852**	0.935	0.969
	CTL_MN60	0.761	**0.985**	0.129	0.732	0.922	0.827
	INV_S	0.908	0.983	0.932	0.505	**0.937**	**0.970**
	VISIT_S	**0.935**	0.973	**0.659**	0.689	0.931	0.756

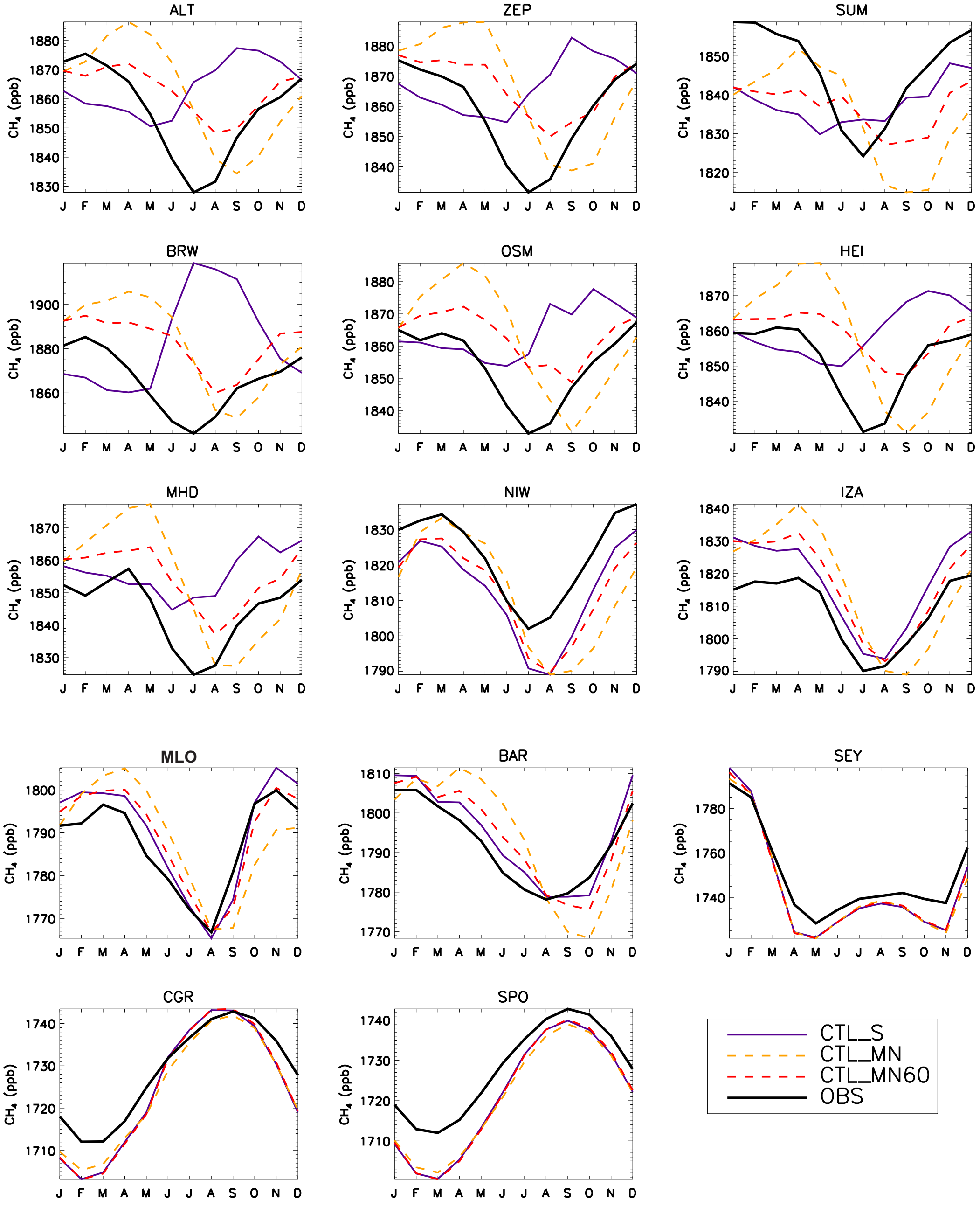
observed data at each station. The INV inventory in particular produces correlations greater than 0.9 at each station apart from BRW, at which the correlation is almost 0.7.

7.3 Adjoint Modelling of Arctic CH$_4$

In order to investigate the differences of the emission estimates, the CTL and INV inventories were compared in more detail. The adjoint version of the TOMCAT model, described in detail in Chapter 5 can be used as tool to provide $\chi = \left(\frac{\partial c}{\partial f}\right)^T$, the sensitivity of a concentration to a change in a surface flux elsewhere. This means that if the adjoint model is initialised with a concentration field which is zero everywhere except for the grid box containing the surface station in which we are interested, a simulation will indicate the sensitivity of that station to the surface flux at each of the surface grid boxes. This sensitivity field may be multiplied by a 2D flux field to obtain χf, the sensitivity of the station to actual fluxes. This process was carried out for the two Arctic stations ALT and BRW, investigating each station's sensitivity in July and August to both the CTL and INV surface flux fields over the preceding two months. Results from ALT are shown in Figure 7.2 and for BRW are shown in Figure 7.3.

Figure 7.2 shows the mean sensitivity field χ for the ALT station site for July and August 2008, at points 10 days (t_{10}), 30 days (t_{30}) and 60 days (t_{60}) previous to the starting point. This shows that most air sampled at ALT is from local regions such as Greenland and North-east Canada which has arrived at ALT within ten days. The majority of air which has arrived at ALT within one month is also from these regions, although there is also a greater influence from Eastern Siberia and North America. At t_{60}, the influence of Siberian air becomes comparable to that of North America, and air arrives at ALT from much of the NH. Also shown in Figure 7.2 is χf at ALT for the CTL and INV surface flux fields. This shows that ALT is highly sensitive to recent local emissions from Canada, Alaska and Eastern Siberia in the CTL inventory, which are not as large in INV, and that there is also a smaller influence from Scandinavian emissions. At t_{30}, emissions from North

Figure 7.1 As Figure 6.19, but for runs CTL, CTL_MN (with the seasonal cycle of emissions removed globally) and CTL_MN60 (with the seasonal cycle of emissions removed only above 60°N). Additional panels showing Arctic sites are displayed.



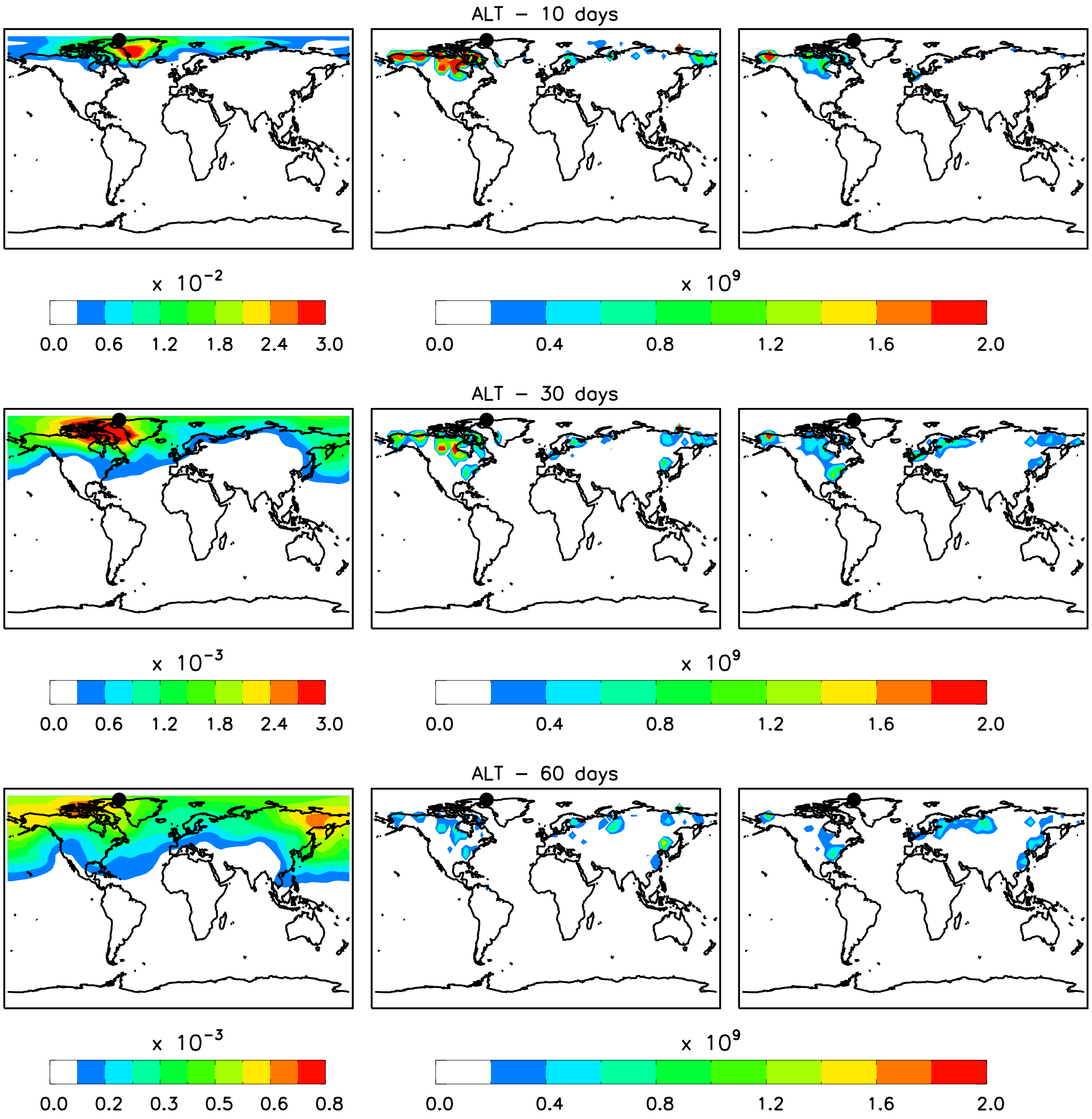
America still have a large influence when the CTL inventory is used, which are again much lower in INV. The influence of anthropogenic emissions from the USA and Europe also begin to have an influence at this time scale, more so in the INV inventory than in CTL. At t_{60}, the sensitivity of the station to local emissions is much smaller, with emissions from the Ural region of Russia, South East Asia and the USA being the most important regions to ALT's CH_4 concentration in both emission inventories. In summary, the modelled JJA concentration increase at ALT when using the CTL inventory is likely to be due to local emission increases in North America and Eastern Siberia, which do not occur to the same extent in the INV inventory.

Figure 7.3 displays the sensitivity of the BRW station to emissions in the CTL and INV inventories. Although the distribution of χ at this station is similar to that of ALT, BRW is much more sensitive to changes in surface fluxes in the NH. The majority of air sampled at BRW at t_{10} and t_{30} has come from North America and Siberia, and from throughout much of the NH at t_{60}. BRW is much more sensitive than ALT to emissions from the USA, Canada and Siberia than ALT, which may explain why the large increase in CH_4 mixing ratio during JJA produced using the CTL inventory is more pronounced at BRW than at other stations. These regions still important at the t_{30} time scale, but the influence is more dispersed at t_{60}. χf indicates that BRW is sensitive to high emissions in North America and Eastern Siberia, which are larger in the CTL inventory than in INV. At t_{30}, the CH_4 concentration at the station is influenced strongly by emissions throughout North America, Russia, Scandinavia and South-East Asia, the majority of which are larger in the CTL inventory than in INV. Anthropogenic emissions from the East Coast of the USA and from Europe are larger in INV, however. At the t_{60} time scale, BRW is still sensitive to the majority of NH emissions, and emissions from the large wetland area in the Ural region of Russia are increasingly important, while the influence of Eastern Siberian and Canadian emissions decreases. The dramatic increase in CH_4 concentration at BRW during June - September is therefore likely to be due to increased emissions from North America and East Siberia which are influencing BRW air samples on a time scale of around one month.

7.4 Sensitivity of Arctic CH_4 to Regional Emissions

As discussed in Chapter 2, the majority of sub-annual variations in CH_4 emissions above 60°N are due to wetlands created by large-scale winter freeze and summer thaw of polar water. Summertime emissions increase greatly in Siberia, Scandinavia, Canada and Alaska due to the northern retreat of ice, leaving behind vast boggy areas which release large amounts of CH_4. However, difficulties arise when trying to measure the extent and distribution of this summer flux increase, and our emission inventories are therefore most inaccurate at these wetland areas. Figures 7.2 and 7.3 showed the influence that emissions from these areas can have on Arctic CH_4 concentrations. In order to quantify the effect that the NH polar emissions are having in the model, two simulations

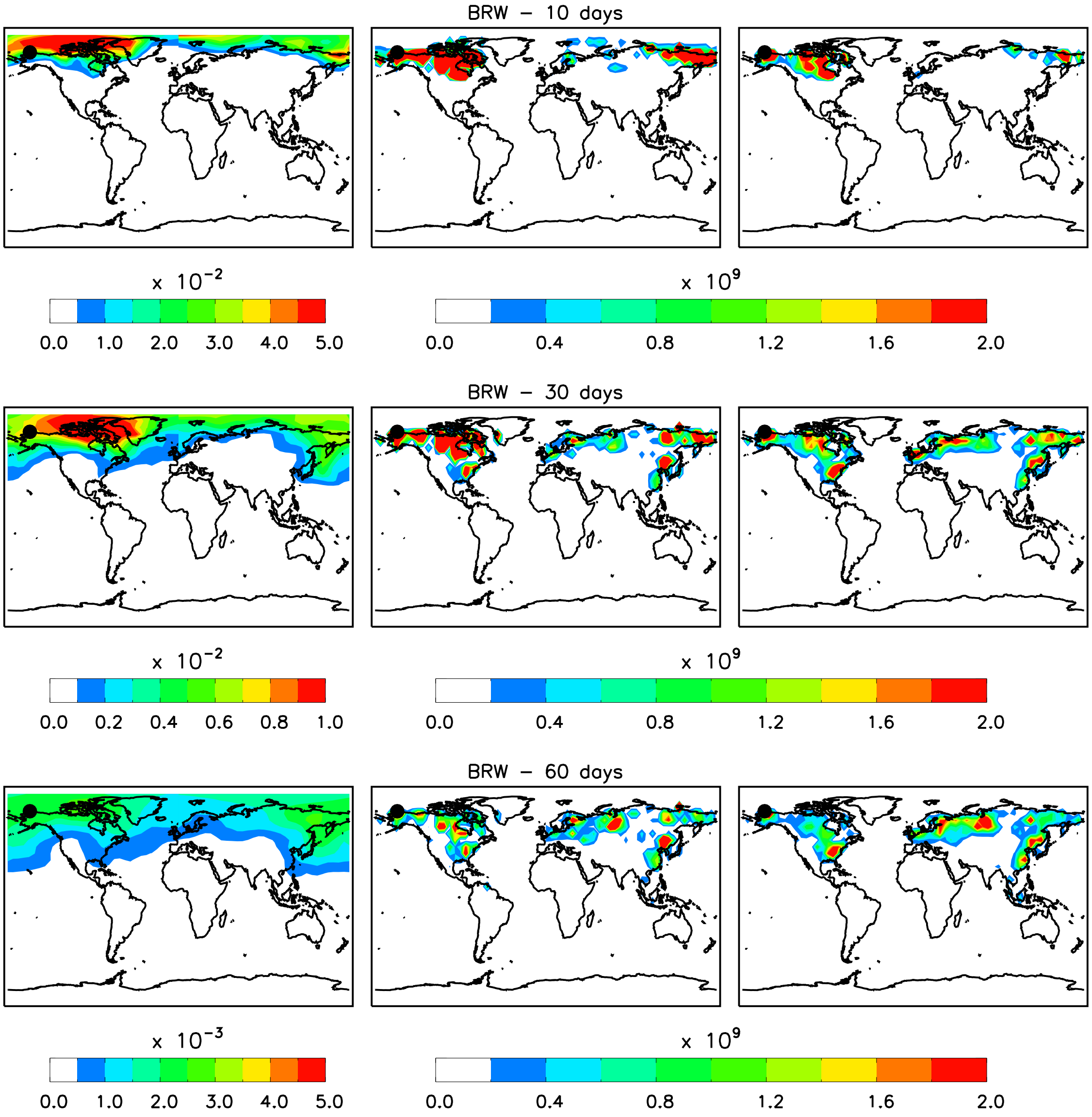
Figure 7.2 *(Left column)* Mean sensitivity field to surface fluxes, χ, for July and August at ALT surface station 10 days *(Top)*, 30 days *(Middle)* and 60 days *(Bottom)* previously. *(Centre column)* shows χf for CTL emission inventory and *(Right column)* shows χf for INV emission inventory.



were carried out which traced emissions from the largest NH emission regions separately. In these simulations, CTL_WL and INV_WL, emissions from the Western Siberian Plain (WSP), to the east of the Ural mountains in Russia, from the North East of Siberia (E_SIB) and from Alaska and Canada (AL_CAN) were treated separately from other regions (ROW), and were also split by season. The distribution of the emission regions are displayed in Figure 7.4.

Figure 7.5 shows the total monthly emissions in 2008 from each of the regions defined in Figure 7.4 for the CTL and INV inventories. At E_SIB, the CTL inventory produces a high emission rate of approximately 1.3 Tg month^{-1} in June - September. Emissions are negligible throughout

Figure 7.3 As Figure 7.2 but for BRW station site.



the rest of the year. The INV inventory produces emissions which are slightly larger than CTL during February to April but which remain low until June, when they increase to a peak rate of approximately 1.0 Tg month^{-1}. The emission rate then decreases , reaching a similar rate to that of CTL in October. Not only is the peak emission rate lower in INV than in CTL, therefore, but the the emission period also begins one month later. At AL_CAN, the CTL emission rate shows a similar timing and magnitude to that of E_SIB, with high emission rates of approximately 1.5 Tg month^{-1} during June - September. The INV inventory, however, displays a gradual increase in emission rate throughout May - August, peaking at just 0.8 Tg month^{-1}, before decreasing to a negligible rate in October. At WSP, total annual emissions are much higher than at the other regions for both emission inventories, peaking at approximately 3.0 Tg month^{-1}. In CTL the emission rate is

Figure 7.4 Emission regions separated in CTL_WL and INV_WL. E_SIB is shaded blue, WSP is shaded green, AL_CAN is shaded red and ROW is shaded grey.

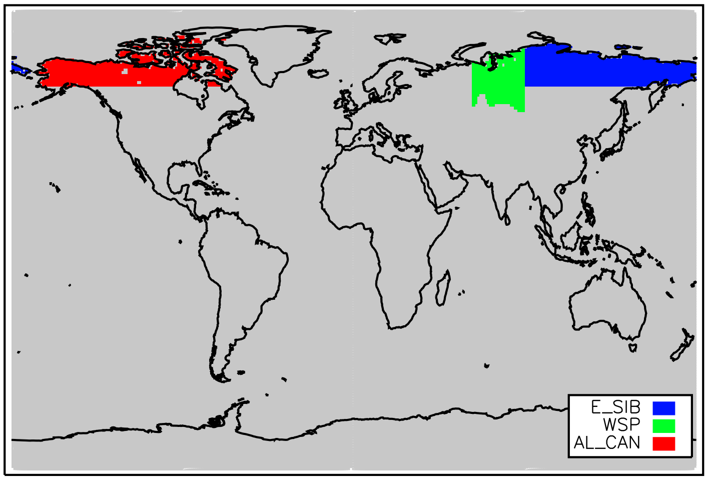
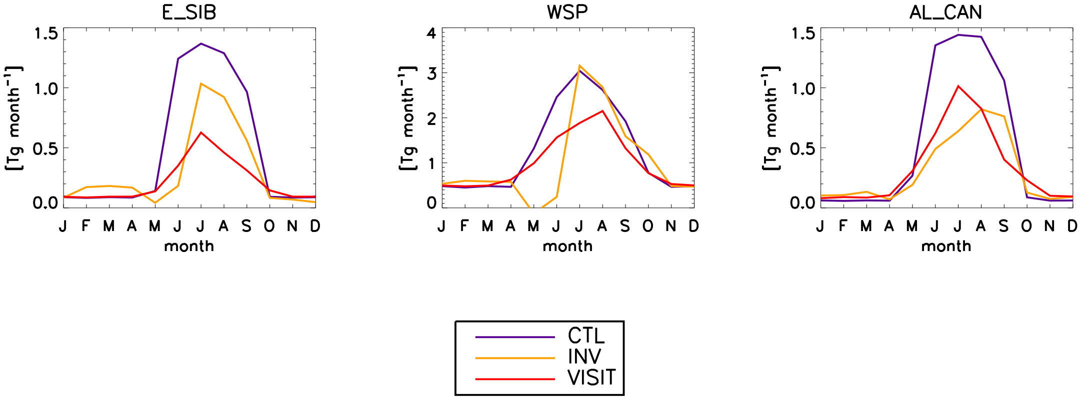


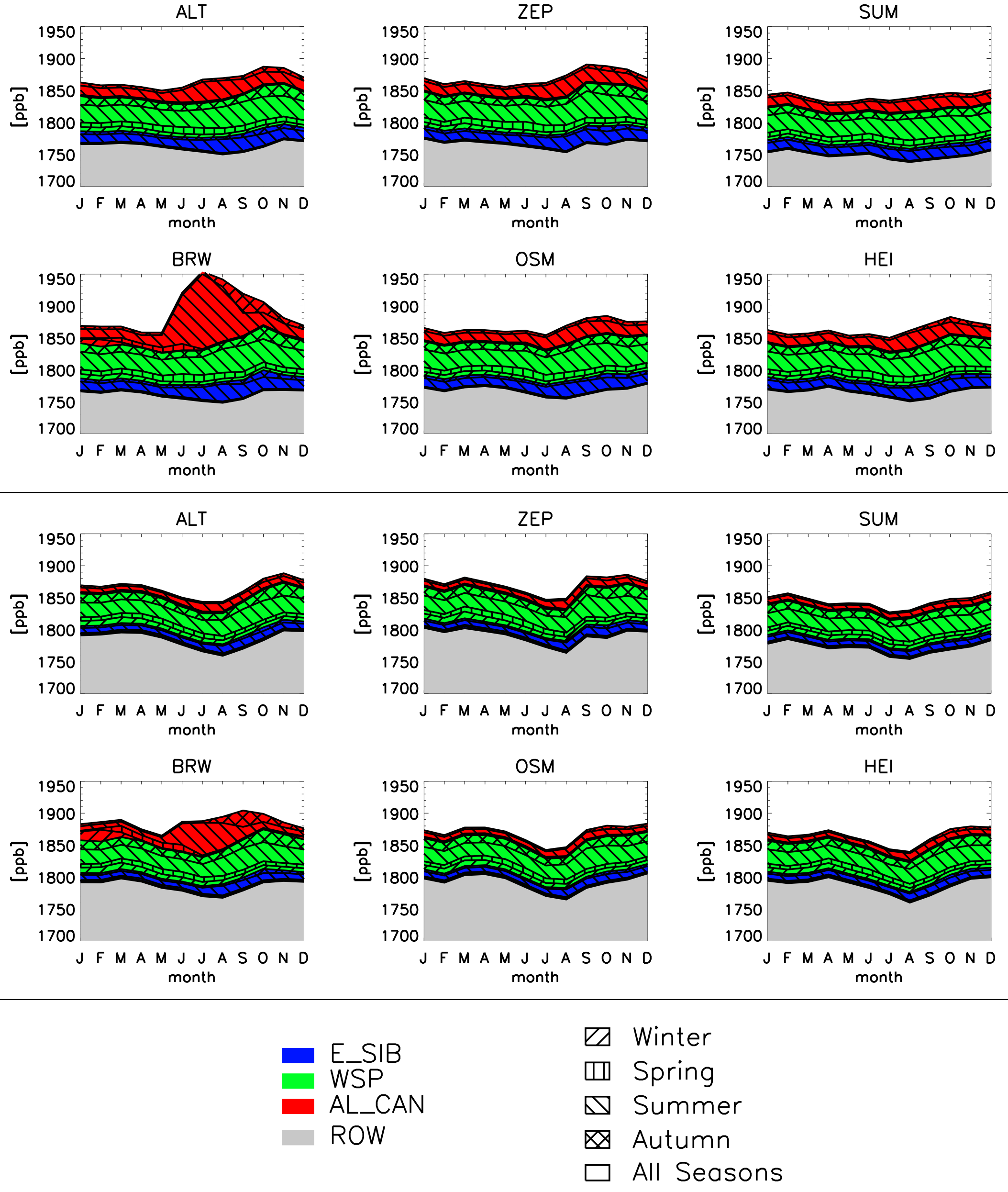
Figure 7.5 Total emissions in Tg month^{-1} for CTL_WL (purple lines) and INV_WL (orange lines), and also for the VISIT inventory at the same locations (red lines) at each of the regions defined in Figure 7.4. *(Left)* E_SIB, *(Centre)* WSP, *(Right)* AL_CAN.



steady at approximately 0.5 Tg month^{-1} until it steadily rises from May - July before decreasing back to the background emission rate in November. The INV inventory is similar until May, when the region becomes a small CH_4 sink before increasing up to the peak emission rate in July a remaining similar to CTL for the rest of the year. This distribution may be slightly flawed due to inaccuracies in the inverse modelling procedure used to create the inventory, but the emission season clearly begins later in INV than in CTL. Total emissions in each of these wetland regions is smaller in INV than in CTL, and this can affect both the global CH_4 total and the seasonal cycle of CH_4 concentration at Arctic stations sites.

Figure 7.6 displays the contribution of each region's emissions to the total CH_4 mixing ratio at

Figure 7.6 Monthly mean CH_4 mixing ratio in ppb at each of the Arctic station sites for CTL_WL *(Top six plots)* and for INV_WL *(Bottom six plots)*, displaying the contribution from each region in a different colour, as in Figure 7.4, while the contribution from each season is cross-hatched differently.



each of the Arctic station sites shown in Table 7.1. Since the total annual emissions of each region are lower in the INV inventory than in the CTL inventory, one would expect that the total contribution of emissions from each of these regions would be lower when the INV inventory is used, and this is indeed true. At ALT, for example, emissions from these three regions when using the CTL inventory contribute approximately 100 ppb, or 5%, of the total CH_4 concentration in January and 120 ppb, or 7%, of the total during July. When using the INV inventory, however, the totals are 80 ppb (4%) and 100 ppb (5%), respectively. The contribution at each of these stations from the rest of the world is similar for both inventories, having the same seasonal cycle as observed data. At each station apart from BRW, contributions from E_SIB, WSP and AL_CAN then remain approximately constant throughout the year in INV_WL, producing the observed seasonal cycle. However, CTL_WL indicates that the concentration increases at these stations from July onwards due to the contribution of emissions during the NH summer and autumn from E_SIB, WSP and AL_CAN. The huge increase in mxing ratio at BRW during June - October when using the CTL inventory is due to local emissions from the AL_CAN region. Although NH summer-time emissions are the main contributing factor to the increase, autumnal emissions also have an effect later in the year. This implies that these emissions from this region are far too high in the CTL inventory. At the same station, the INV inventory shows a similar sensitivity to these local emissions. However increased DJF emissions from the region increase the winter mixing ratio, and the decreased JJA emissions brings down the concentration during the summer. However, Figure 6.19 shows that the concentration of CH_4 at BRW during JJA in the INV_S simulation is also too high, which suggests that even the decreased AL_CAN emissions in this inventory shown in Figure 7.5 may be excessive.

7.5 Inverse Modelling of Arctic CH_4 Emissions

7.5.1 Model Set-up

The TOMCAT 4D-Var system was used to produce a new CH_4 emission inventory for the year 2008. Due to time and computational limitations, it was not possible to include any spin-up for the inversion which runs from January 1, 2008 to January 3, 2009, and produces updated monthly-mean emission estimates for each month of 2008. In order to constrain NH emissions, and to keep to a minimum the number of data to be assimilated, observations were taken only from six station sites in the NOAA/ESRL measurement network. These stations are shown in Figure 7.7, and further details are given in Table 7.2.

Since this is the first inversion to be completed using the TOMCAT 4D-Var system, many of the parameters of the inversion were chosen for their simplicity, with a coarse temporal resolution for the observations and uncorrelated error covariance matrices. However, the results from this

Figure 7.7 Locations of NOAA/ESRL station sites with flask observations available for assimilation into the TOMCAT 4D-Var inversion system in Chapter 7.

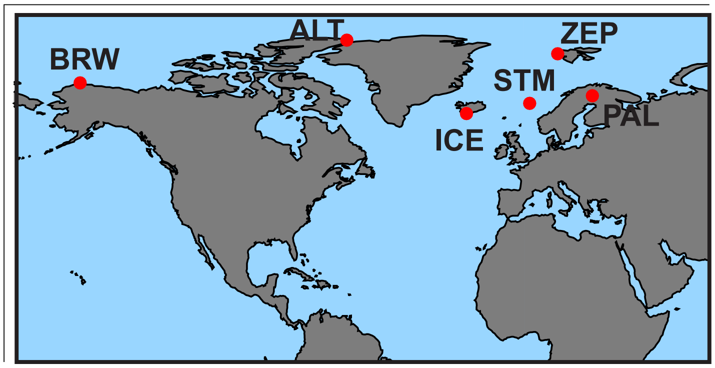


Table 7.2 Details of NOAA ESRL flask sampling sites used in inversions performed in Chapter 7.

Station Code	Station Location	Longitude (°)	Latitude (°)	Altitude (m)	Observation Type
ALT	Alert, Canada	62.5W	82.5N	210	Flask
ZEP	Svalbard, Norway	11.9E	78.9N	475	Flask
BRW	Barrow, Alaska	156.6W	71.3N	11	Flask
PAL	Pallas, Finland	24.1E	68.0N	565	Flask
STM	Ocean Station ‘M’	2.0E	66.0N	5	Flask
ICE	Heimaey, Iceland	20.3W	63.4N	100	Flask

initial inversion will indicate the system’s potential for producing accurate high resolution flux inventories in the future. The inversion was carried out using the T21_L model set-up, described in Chapter 4. Observations were averaged temporally over 8 days, which gives 46 assimilated observations at each station over the course of the one-year inversion. Due to the fact that there was missing data at some stations, however, there were in fact only 244 assimilated data points, rather than 276. During assimilation into the adjoint model, if there were multiple stations inside the same model grid box, a mean value was used. For simplicity, the errors between observations were assumed to be uncorrelated and equal to 3 ppb for every measurement, as this is the value suggested as the accuracy of CH_4 flask observations by the NOAA/ESRL measurement network.

For this study, since there is a large disparity in scale between the model-observation differences ($\approx 10^2$ ppb) and the changes made to the emission rates in order to correct these differences ($\approx 10^6$ kg grid box^{-1} hr^{-1}), a scaling factor γ was introduced to the cost function as follows;

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2\gamma}(\mathbf{y}-\mathbf{H}(\mathbf{T}[\mathbf{x}]))^T\mathbf{R}^{-1}(\mathbf{y}-\mathbf{H}(\mathbf{T}[\mathbf{x}])) \tag{7.1}$$

where $\gamma = 1\times10^{-4}$. A similar ‘scaling’ method was used in Henze et al. (2007). This means that model-observation errors and changes to the emission rate both have a similar ‘cost’. The gradient of the cost function is then evaluated as;

$$\nabla_{\mathbf{x}} J(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{\gamma}\mathbf{T}^T \left[\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}(\mathbf{T}[\mathbf{x}]))\right] = 0 \tag{7.2}$$

This choice of constraint allows the inversion to make fairly large alterations to the emission rate in order to match the observations, without the possibility that a poorly defined error covariance for the a priori may mean that the system will not converge (F. Chevallier, personal communication).

Negative emission rates were not allowed in the inversion, and any negative values were set to zero before evaluation of the cost function. The CTL emission inventory for 2008 was used as the a priori flux estimate, and the inversion was initialised using output from the CTL simulation described in Chapter 6. The OH_S and SD chemical fields were used to represent atmospheric chemistry in the model, and the 2-D soil sink used in the forward CH_4 simulations was also included. The optimisation procedure was repeated until four consecutive iterations had passed without minimisation. The inversion code has not yet been parallelised so the local serial machines were used for the simulations. This meant that individual iterations required a significant amount of computational time due the large amount of reading and writing to files necessary to perform the adjoint transport.

7.5.2 Inversion Results

Figure 7.8 shows the results of the inversion as updates to the monthly-mean CTL emission inventory for 2008. Changes are small in the NH winter. In January – April and November – December, updates are generally limited to relatively small increases in emission rates in oceanic regions around North America, with extremely small decreases over land-based NH regions. There are more substantial updates in May – October, however, with significant decreases in high-latitude NH land-based regions throughout these summer months. The largest decreases were in Alaska, Canada, Eastern Siberia and Scandinavia in JJA, with reductions of up to 100% in some grid boxes in these regions. The influence of SH emissions on the Arctic stations is insignificant on the one-year time-scale of the inversion, although this may not be true for a longer inversion. In total, the annual emissions have fallen from 543.4 Tg yr^{-1} in the CTL inventory to 521.6 Tg yr^{-1}, a decrease of 21.8 Tg yr^{-1}.

Figure 7.9 shows the changing values of the cost function and the norm of the gradient of the cost function, $||\mathbf{g}||$, throughout the iterations of the inversion Only three iterations were carried out, and the value of J was reduced by approximately 70% at the end of the third iteration, while the value of $||\mathbf{g}||$ has been decreased by approximately 85%. There were relatively few minimisations required for a significant reduction of the cost function due to the fact that a large step size along

Figure 7.8 CH_4 emissions for the CTL inventory *(Left)* and the difference between the updated TOM inventory and CTL *(Right)* (kg grid box^{-1} hr^{-1}) for January, February and March *(Top-Bottom)*.

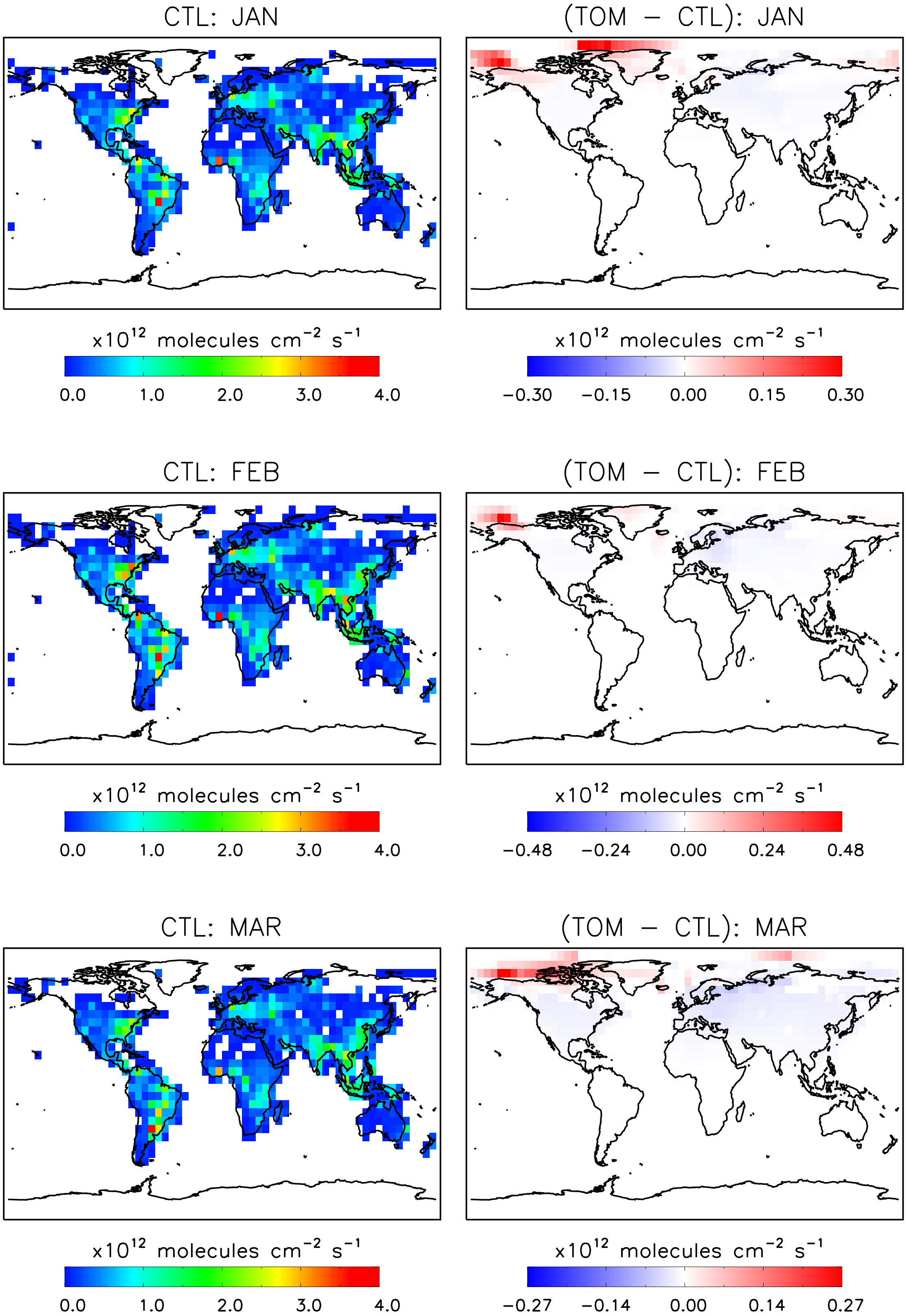


Figure 7.8 As previous page, but for April, May and June.

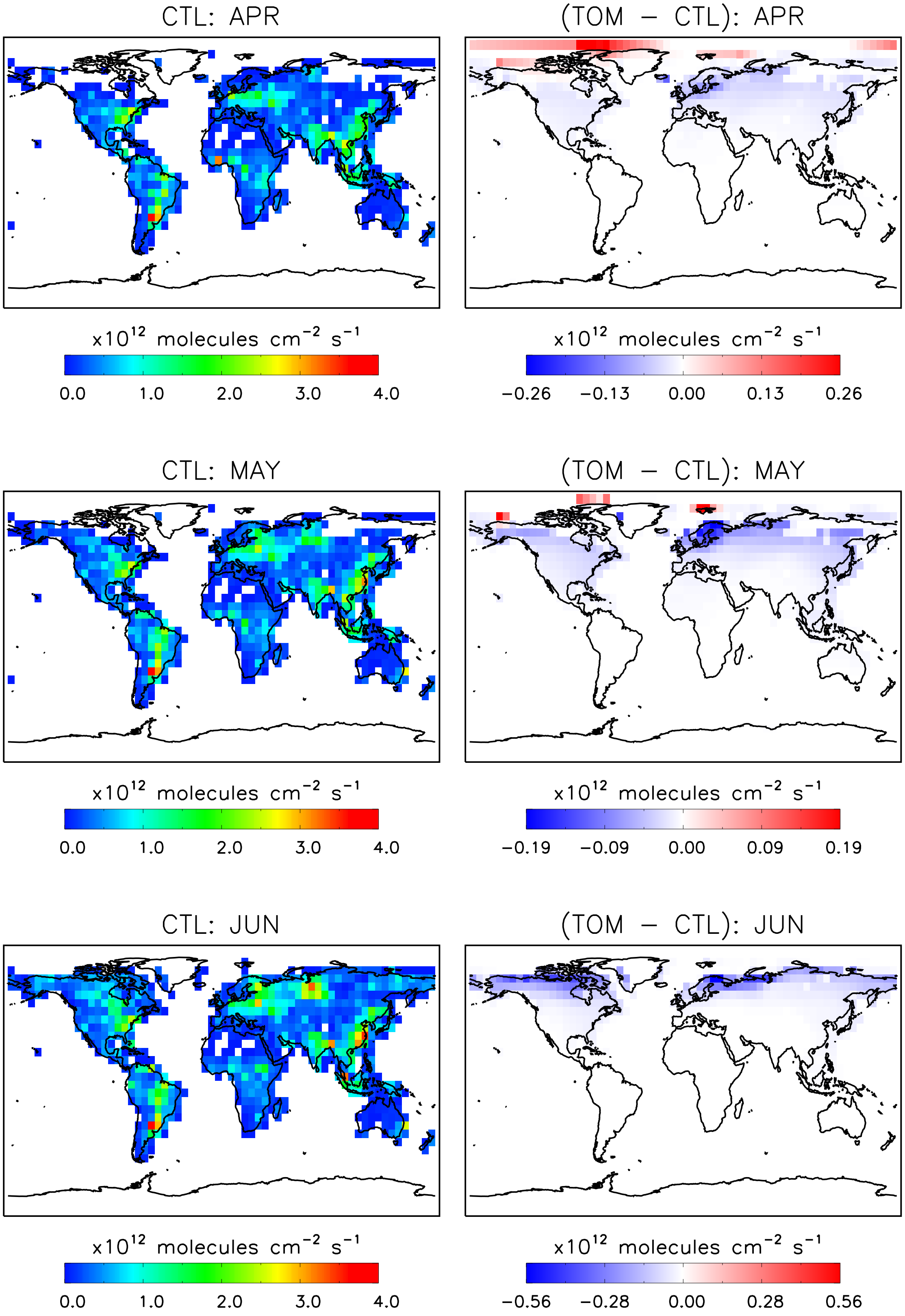


Figure 7.8 As previous page, but for July, August and September.

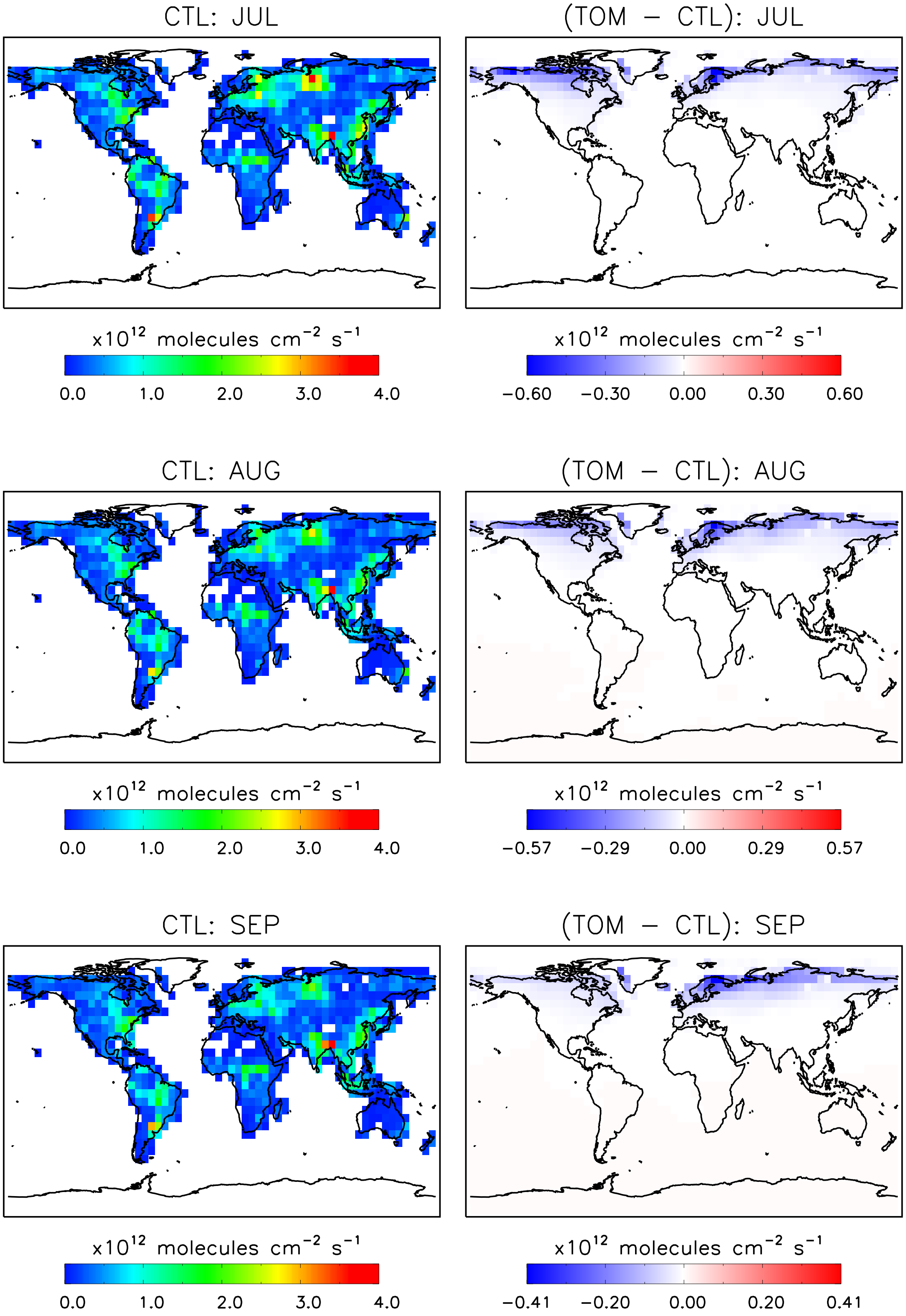


Figure 7.8 As previous page, but for October, November and December.

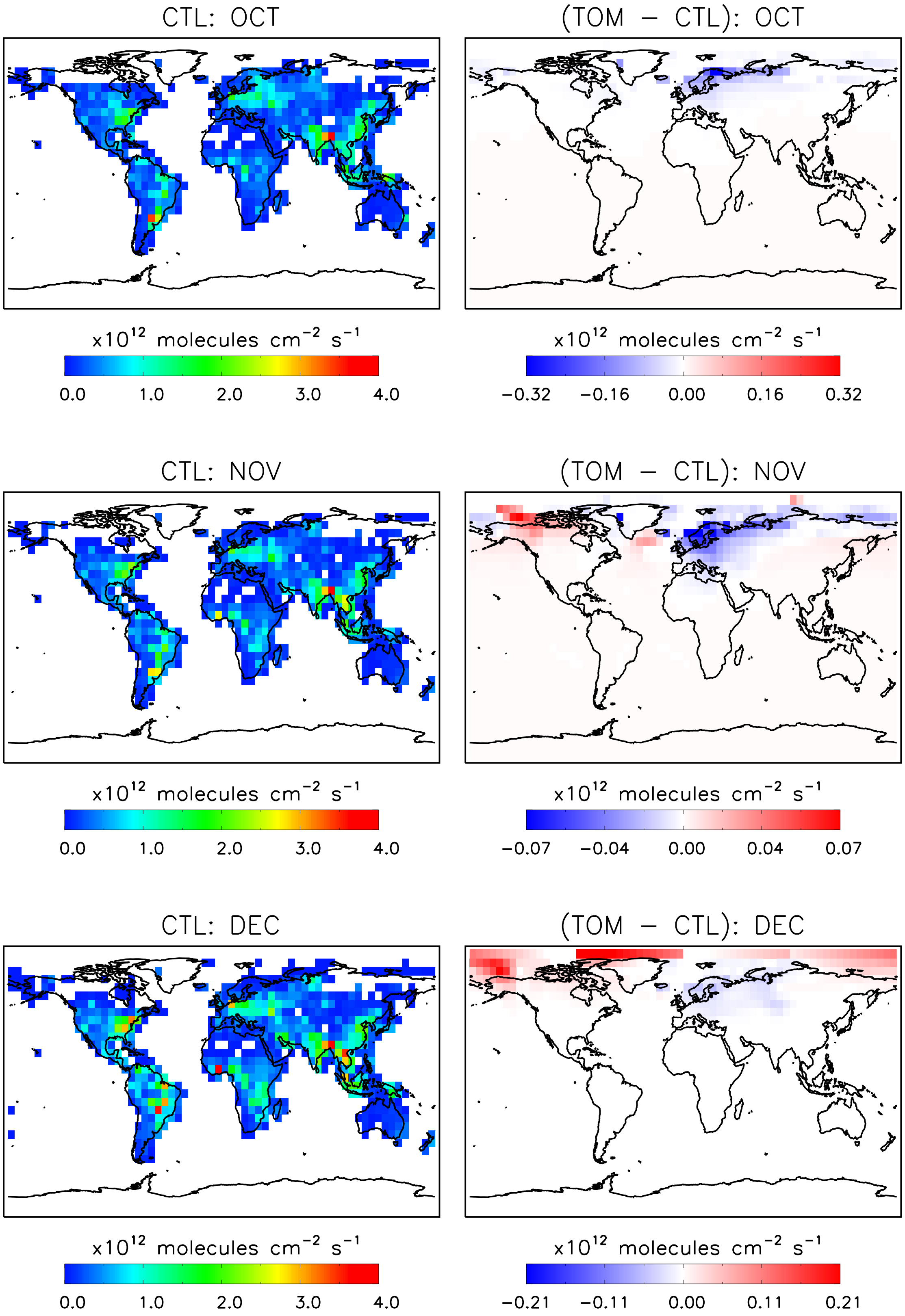
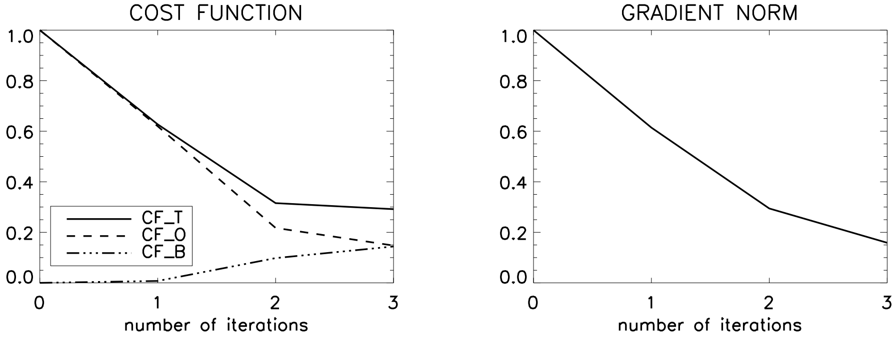


Figure 7.9 *(Left)*Value of the cost function J at each iteration of the 2008 inversion and *(Right)* the norm of the gradient of J.



each minimisation vector was chosen. Also shown are the contribution of the background term, CF_B, and the observation term, CF_O, to the total value of the cost function, where;

$$CF\_B = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) \tag{7.3}$$

and;

$$CF\_O = \frac{1}{\gamma} (\mathbf{y} - H(T[\mathbf{x}]))^T \mathbf{R}^{-1} (\mathbf{y} - H(T[\mathbf{x}])) \tag{7.4}$$

where all terms are defined as in Equation 7.1. By the end of the minimisation the observational cost term, CF_O, which is a measure of the differences between the model output and the observations, has in fact been reduced to approximately 15% of the initial cost function value. However, the background term is becoming significant by the final iteration, and also contributes a value approximately 15% the size of the initial cost function, or 50% of the final value of J. It is interesting to note that the cost function can be reduced by up to 40% with minimal background contribution to the cost function.

In order to compare the updated emission inventory produced by the inversion, which is on the model's 5.6° × 5.6° grid, to the CTL, INV and VISIT inventories, it was necessary to map the 1° × 1° inventories onto the model grid. Figure 7.10 shows the AL_CAN, WSP and E_SIB regions, previously described in Section 7.4 on the coarser grid. The emission inventory produced by the inversion is henceforth referred to as TOM.

Figure 7.11 shows the total monthly emissions from the AL_CAN, WSP and E_SIB regions in the emission inventories on the 5.6° emission grid. The TOM inventory produces wetland emissions which have been reduced by up to 50% in the NH summer in the AL_CAN and E_SIB regions,

Figure 7.10 As Figure 7.4, but for the 5.6° emission grid.

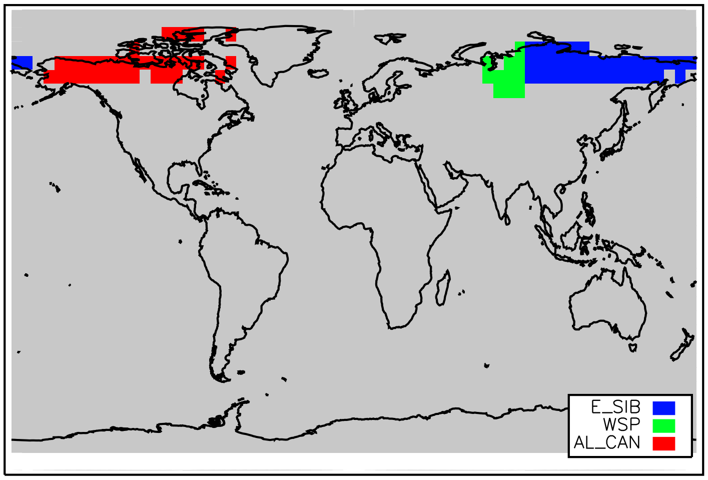
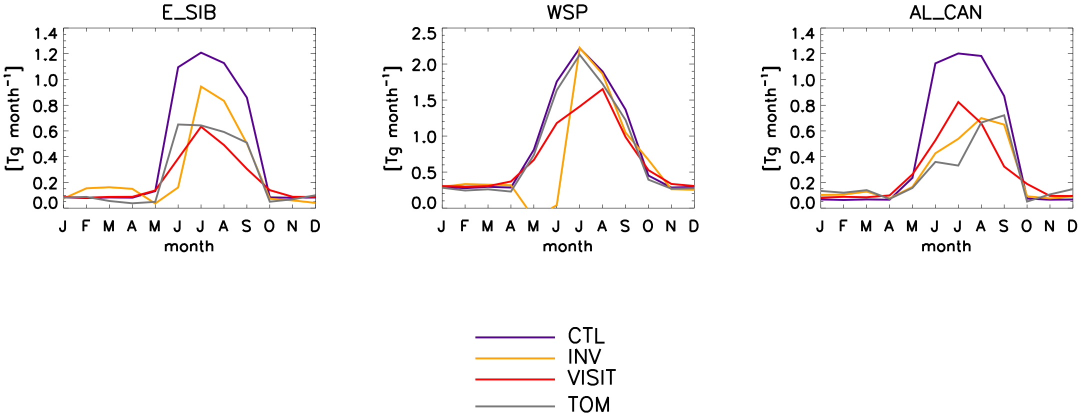


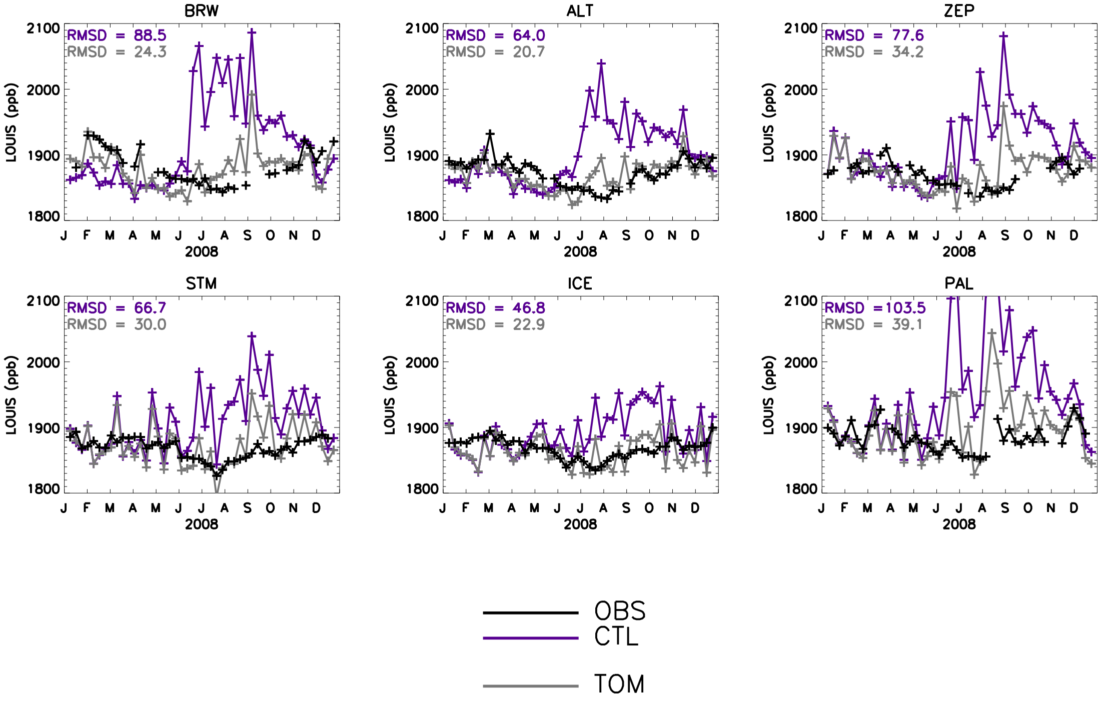
Figure 7.11 Total emissions (Tg month^{-1}) for the E_SIB, WSP and AL_CAN regions in 2008 for the CTL, INV, VISIT and TOM inventories on the 5.6° emission grid shown in Figure 7.10.



when compared to the a priori CTL inventory. These changes are in good agreement with the INV and VISIT inventories. In particular, the INV and TOM inventories have both favoured a late summer peak in wetland emissions in the AL_CAN region, with relatively low emissions in May and June. The emissions in the WSP region are relatively unchanged, meanwhile.

Figure 7.12 shows the results of a forward simulationcarried out using the TOM emission inventory. The simulation was initialised at January 1, 2008 using output from the CTL simulation described in Chapter 6. The model grid was identical to the T21_L29 set-up described in Chapter 4, which was also used for the inversion which produced TOM. This set-up uses a grid resolution of 5.6° × 5.6° and 29 vertical levels. The simulation used the FOM advection scheme and the Louis boundary layer scheme, so that the results of this simulation are the same as those prodused when the forward model is run during the inversion. At each station, the TOM inventory gives

Figure 7.12 CH_4 concentration (ppb) at Arctic surface station sites in 2008 from flask measurements (black lines) and a TOMCAT forward model simulation using the Louis PBL scheme and the CTL emission inventory (purple lines) and the TOM emission inventory produced using the TOMCAT 4D-Var inversion system (grey lines).

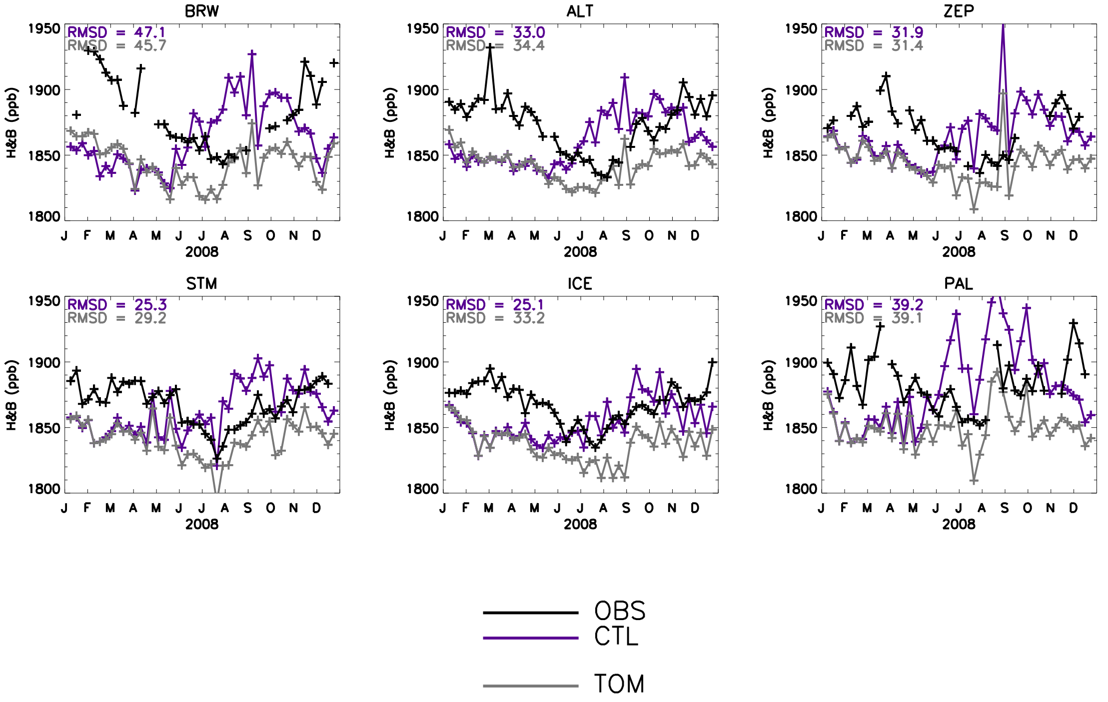


a smaller RMSD (as defined in Section 6.4) than that produced when the CTL inventory is used, largely due to the decreased atmospheric concentrations produced by TOM in June–November. The RMSD has been decreased by 23.9–67.4 ppb by the use of the TOM inventory, which corresponds to an improvement of 51.1%–72.5%. However, some observations are still over-estimated by up to 150 ppb by the model, indicating that changes to the emission inventory may still lead to improved model results if more flexibility around the a priori were to be allowed in an inversion.

7.5.3 Limitations of the Inversions

Although the TOM emission inventory produced by the inversion described in this chapter has led to improved results when a forward model simulation carried out with the new inventory uses the same set-up as the inversion system, this clearly does not guarantee improvements when the model set-up is altered. To further investigate the accuracy of the TOM inventory, a one-year simulation identical to that which produced the results shown in Figure 7.12 was performed, but with the SOM advection scheme and the Holtslag & Boville boundary layer scheme used instead of their simpler counterparts. The results of this simulation are shown in Figure 7.13. In this case, the TOM inventory has provided no improvement upon the CTL inventory in terms of the accuracy of the modelled CH_4 concentration at the Arctic stations. The RMSD has in fact increased at

Figure 7.13 As Figure 7.12, but for a TOMCAT forward model simulation using the Holtslag & Boville PBL scheme rather than the Louis scheme.



three of the stations due to the use of TOM, while there is no significant improvement at the remaining three station sites. It seems that the inversion has underestimated the necessary increase in emissions during the NH winter months, while the decrease in summertime emissions is larger than necessary to bring modelled concentrations in line with observations.

The limitations of the Louis boundary layer scheme have already been discussed in Chapters 3 and 4, and the fact that the scheme 'traps' emissions close to the surface has led to inaccurate results in the inversion, assuming that the Holtslag & Boville scheme is the more accurate of the two available schemes. The inverse system is attempting to decrease artificially high summertime concentrations of CH_4 produced by the use of the Louis scheme, which leads to over-compensation when the Holtslag & Boville scheme is used in the forward model. This implies that the decreased emissions in the E_SIB and AL_CAN regions shown in Figure 7.11 may be exaggerated. In order to accurately find an updated emission inventory which is consistent with the use of the Holtslag & Boville scheme and adjoint version of this boundary layer scheme will need to be developed and used in future inversions.

7.5.4 Differences in the Process-based Models

Clearly, the estimation of the magnitude of NH wetland fluxes differ greatly in the CTL, INV and VISIT inventories used in this study. As already discussed, the INV inventory is from a synthesis inversion, and so contains no diagnosis of the underlying processes of CH_4 emission. However, the natural emissions in the CTL inventory and the VISIT inventory both come from process-based bottom-up models, and so the large differences in emission estimates at NH wetlands indicates that the two may use very different methods to produce their estimates.

The natural wetland emissions from the CTL inventory are provided by the GISS natural emissions inventory using a wetland model described in Matthews and Fung (1987), while the VISIT inventory estimates wetland emissions using the Vegetation Integrative Simulator for Trace gases (VISIT) (Ito, 2010). The GISS inventory estimated NH wetland emissions, classified as forested and non-forested bog, using the fairly crude assumptions, based on the small amount of measurements available at the time, that the CH_4 emission rate is constant everywhere in this type of wetland and that the length of the production emission season is 100 days. The VISIT inventory, meanwhile, used a one-dimensional multi-layer scheme for CH_4 emission from wetlands developed by Walter and Heimann (2000), in which CH_4 emission was dependent upon diffusive flux, ebullition (bubbling), plant mediated flux and CH_4 production and oxidation in soils. In turn CH_4 production was calculated as a function of biome type, substrate availability, temperature and inundation. This meant that although the wetland map used by Matthews and Fung (1987) estimated a smaller wetland area in the AL_CAN region that that of Ito (2010), emissions were far greater in the GISS inventory due to the lack of variation in the emission rate and thaw season length. The VISIT inventory produces atmospheric concentrations which are more consistent with observations in the NH summer in the Arctic.

7.6 Summary

This chapter examined the sensitivity of Arctic concentrations of CH_4 to surface emissions of the species, particularly to those in the high latitude NH, and investigated the accuracy of our current CH_4 emission inventories. It was found that surface concentrations of CH_4 in the Arctic are highly sensitive to local emissions such as those from wetlands and thaw lakes in Siberia, Alaska, Canada and Scandinavia, and that inaccuracies in our estimates of emissions from these sources can greatly impair our ability to model atmospheric CH_4. Using the adjoint TOMCAT model it was found that in the NH summer, the stations at BRW and ALT are highly sensitive to CH_4 emissions from across the NH on a time scale of at least two months, with large contributions from wetland regions. The forward TOMCAT model was then used to find the extent of the influence of large seasonal wetland emissions from Eastern Siberia, the West Siberian Plain and Alaska and Canada in the CTL and INV inventories. It was found that the CTL inventory, which uses natural emissions from the GISS inventory, must overestimate the scale of these emissions by up to a factor of 2. In particular, the emissions in the AL_CAN and E_SIB regions appear to be responsible for large increases in the CH_4 concentration at BRW and ALT during the NH summer.

The TOMCAT 4D-Var system was used in order to assimilate Arctic surface measurements of CH_4 so that the emission inventory could be constrained to let the model results match the observations. A one-year inversion was performed for 2008, and it was found that, compared with the CTL inventory, the inversion reduced emissions in the AL_CAN and E_SIB regions by up to 0.6 Tg CH_4 month^{-1} during the wetland emission season, a result which is is good agreement with the INV and VISIT inventories. These reductions corresponded to a decrease in modelled atmospheric concentrationof CH_4 of up to 150 ppb when using the Louis boundary layer scheme and 50 ppb when using the Holtslag & Boville scheme. There is little change in emissions from the WSP region.

The 4D-Var inversion system displayed its potential for producing revised estimates of emissions through the assimilation of observations, although the use of the Louis scheme may be responsible for inaccuracies in the results of the inversion. The creation and use of an adjoint version of the Holtslag & Boville boundary layer scheme may lead to more accurate updates to the CH_4 emission inventory. Also, an inversion with a long time period of multiple years, allowing for a spin-up period either side of the inversion period, would produce temporally consistent and accurate results.

Chapter 8

Summary

8.1 Completion of Aims

The objective of this thesis was to investigate recent variations in the global atmospheric methane budget using the TOMCAT CTM. As discussed in Chapter 2, the atmospheric CH_4 budget is controlled by the balance of a number of sources and sinks, and our understanding of changes to the global atmospheric growth rate of methane is therefore highly dependent upon a full awareness of the natures, locations and magnitudes of these fluxes. This study aimed to improve our understanding of this balance through a combination of forward model simulations and inverse modelling. Through use of these two complementary techniques, the sensitivity of the atmospheric methane budget to various estimates of emission and destruction processes was revealed, with a specific focus on the nature of high-latitude northern hemisphere emissions of methane from wetlands.

An important first step in the process of modelling the atmospheric concentration of CH_4 is to examine the accuracy of the simulated model transport. This was achieved through comparison of simulated atmospheric concentrations of SF_6 to surface measurements and satellite observations. The standard model set-up used a 2.8° model grid, an advection scheme which conserved up to second-order moments (SOM), and a non-local boundary layer mixing scheme. It was found that with this set-up, the model performed well in comparison to the observations, capturing both the annual increase and the seasonal cycle at a number of surface stations. This indicated that both the simulated transport and SF_6 emissions in the model are accurate. There is a possibility that the interhemispheric transport in the model may be slow, however, and there is too little temporal variation throughout the southern hemisphere. Although reducing the resolution of the model grid to 5.6° and changing to an advection scheme which conserved only first-order moments (FOM) did not significantly reduce the accuracy of the model transport, use of a local boundary layer mixing scheme reduced the accuracy of the model in regions close to SF_6 emissions. The scheme in question has been found in the past to produce slow vertical mixing out of the PBL, leading to a build up of emissions near the surface (Wang et al., 1999), and this was confirmed in this thesis. Increasing the depth of the surface model layer improved the issue to some extent, although surface concentrations were still too high in some locations. The standard model set-up also compared favourably with measurements in the lower troposphere taken on aircraft flights and also with satellite observations from the ACE satellite in the upper troposphere and lower stratosphere.

Once the accuracy of the transport model had been investigated, a number of forward simulations of CH_4 were carried out. These simulations investigated the effect of changes to the emission inventory, the atmospheric OH sink and the stratospheric concentration on the tropospheric concentration of methane. Two OH fields were examined, provided by Spivakovsky et al. (2000) and the TOMCAT full chemistry model. The total tropospheric OH provided in the Spivakovsky field was approximately 13% higher than that from TOMCAT, with a distribution weighted towards the mid-troposphere rather than the surface. Due to the lack of extensive tropospheric OH observations the two available model OH fields were compared to measurements taken on the PEM-Tropics campaign. It was found that the OH field produced by Spivakovsky et al. (2000) was more consistent than the TOMCAT full chemistry field with the limited observations provided by the aircraft campaign. The two OH fields performed similarly to one another in simulations of the destruction of atmospheric methyl chloroform, producing similar destruction rates for the period 2000-2008. However, in forward CH_4 simulations, the total TOMCAT OH appeared to be too low, producing CH_4 concentrations that were approximately 100 ppb too high at the surface.

For the period 1995-2005, a simulation was carried out in which stratospheric satellite observations of CH_4 were assimilated into the forward model. Observations were assimilated from the HALOE satellite until 2003, after which ACE observations were assimilated. This was the first time that ACE data had been assimilated into the TOMCAT model. It was found that although assimilated stratospheric CH_4 was overall higher than in the standard simulations, assimilation produced decreased CH_4 concentration at the surface, most likely due to subsidence of polar CH_4 near the tropopause, which had a lower concentration in the assimilated simulation. This emphasises the importance of representing a realistic stratospheric sink for the accurate simulation of surface CH_4.

Forward CH_4 simulations were carried out using six emission inventories, each with varying contributions from anthropogenic, biomass burning and wetland emissions. This was the first time that long-term simulations of CH_4 with full emission inventories had been carried out using the TOMCAT model, so that the CH_4 distribution was fully determined by the model emissions, transport and chemistry. Each of the emission inventories captured the decrease in the atmospheric growth rate of CH_4 throughout the 1990s to some extent, and five of the six inventories captures the period of stability in the early 2000s. The CTL_E4 inventory, in which anthropogenic CH_4

emissions were provided by the EDGAR V4.0 emission inventory estimated that anthropogenic emissions had greatly increased since 2002 in line with Asian economic growth. However, it was found that this meant that the model overestimated the global CH_4 concentration in the 2000s. This implies that for anthropogenic emissions to have risen to the extent estimated in the CTL_E4 inventory, they must have been mitigated by a corresponding decrease in some other emission process. The VISIT inventory, which used natural CH_4 emissions provided from the Vegetation Integrative Simulator for Trace gases, produced CH_4 concentrations which decrease throughout the 2000s, indicating that emissions throughout this decade was higher than the $\sim$530 Tg yr^{-1} estimated in this inventory.

The seasonal cycle of atmospheric CH_4 at a number of surface stations was investigated. At stations south of 50°N, each emission inventory produces similar results, in very good agreement with observations (correlations with observations greater than 0.9). This indicates that the seasonal cycle in the tropics and SH is dependent more upon transport and chemical destruction than on the distribution of emissions. The results at high latitude NH stations vary between inventories, however, with those using natural emissions provided by the GISS process-based emission model performing poorly. Since the model transport and OH chemistry had been previously validated, it was deduced that it was likely that discrepancies in the emission inventories were responsible for the poor representation of the seasonal cycle. Using separate simulations for different regional emissions, it was found that wetland emissions in Alaska, Canada and Siberia were likely to be overestimated in the GISS inventory by up to 100%. The overestimation is likely due to simplified estimates of emission rates and thaw periods, and emphasises the importance of reducing the potential for 'scaling-up' errors in bottom-up models.

In order to produce an updated CH_4 emission estimate, a 4D-Var inverse version of the TOMCAT model was developed as part of this study. The inversion system was produced with the aim of assimilating atmospheric CH_4 measurements in order to constrain surface emissions. First, the adjoint version of the TOMCAT model was coded explicitly, and thoroughly and extensively tested in order to validate the accuracy of the adjoint transport. The TOMCAT adjoint model was tested numerically using the adjoint identity test, before the conservation and reciprocity of the adjoint transport were also tested. The adjoint model was then incorporated into a 4D-Var optimisation code, which could be used in order to assimilate atmospheric observations. The optimisation code was tested using synthetic observations produced by the TOMCAT model, and it was found that model error had only a small effect on optimisation of the surface fluxes.

The 4D-Var system was used in order to to assimilate Arctic CH_4 observations from a small sample of surface sites, producing updated monthly-mean surface flux estimates for the year 2008. The inverse model performed well, reducing the value of the cost function by over 60%, producing an updated 'TOM' CH_4 emission inventory which improved upon the CTL inventory when used in the TOMCAT model with the Louis boundary layer scheme applied. The inversion confirmed that wetland fluxes in Alaska, Canada and Eastern Siberia in the GISS inventory needed to be reduced in order to produce model results consistent with observations, although the extent of the reductions is unconfirmed due to the inconsistencies produced by the use of the Louis scheme in the inverse model.

Recent variations in the growth rate of CH_4 have highlighted our lack of understanding of the source and sink processes which contribute to the atmospheric CH_4 budget. While a number of theories have been put forward as to the nature of the steady state reached in the 2000s, and the subsequent increase in the growth rate in recent years, there is currently no consensus as to the reasons for the variations. This study has illustrated the difficulties of modelling atmospheric CH_4 and highlighted the importance of a thorough understanding of each process contributing to variations in the atmospheric CH_4 budget. Accurate representation of emissions, destruction processes and transport are all necessary in order to produce model estimates that are consistent with observations. The machinery developed within this thesis will provide insight into spatial and temporal variations in the methane emission cycle, and together with complementary bottom-up models, will improve our understanding of the processes which control the atmospheric CH_4 budget.

8.2 Future Work

The 4D-Var system developed as part of this work lends itself to a range of future applications. A necessary first step, however, is to optimise the speed and memory requirements of the inverse model, which currently makes long simulations prohibitive. As the system has been explicitly coded, it will be possible to parallelise it to allow it run on non-local, high-speed computers. The system can then be applied to a well constrained inversion with a long simulation period. This requires a well-defined background covariance matrix and a large, well filtered observation set. The resulting inversion should be for one year, after a year of spin-up in order that model-observation bias does not influence the updated emissions. Repeating this process for multiple years will provide an idea of how CH_4 emissions are varying interannually, and how this influences the atmospheric budget. The results of such inversions may be compared to those of bottom-up models in order to provide an indication of the reason for the recent reduction and subsequent increase in the atmospheric CH_4 growth rate. Pre-conditioning of the variable which is minimised in the cost function, discussed in Courtier et al. (1994), can lead to faster convergence to an optimal solution, and should be added to the TOMCAT 4D-Var system. The system allows for the assimilation of large amounts of satellite data, and this provides a high resolution constraint for the assimilation. Satellites such as SCIAMACHY, ACE, HALOE and GOSAT all produce CH_4 measurements which may be assimilated into the TOMCAT 4D-Var model.

It is likely that the boundary layer scheme in the inverse model will need to be upgraded to the Holtslag & Boville scheme, as the Louis scheme generally produces poor results. In the inverse model, slow mixing through the boundary layer towards the surface may lead to an underestimation of surface fluxes. Although the Holtslag & Boville PBL scheme is a more complicated, non-linear piece of code, it is possible to produce its adjoint and include it in the inverse model.

As mentioned in Chapter 3, the isotopic composition of atmospheric CH_4 provides information on its emission, and including this in the inverse and forward models will allow partitioning of surface fluxes into emission types. This can also provide information on sinks of CH_4, such as the tropospheric OH distribution.

Finally, the inverse model may be extended to other atmospheric species. 4D-Var inversions have been carried out in the past for CO, CO_2, NO_X and H_2 (e.g. Hooghiemstra et al. (2011); Chevallier et al. (2005); Yver et al. (2011); Chai et al. (2009)), and now that the adjoint transport subroutines are in place, creating the necessary adjoint chemistry routines is a small step. Palmer et al. (2006) examined the possibility of inverse modelling of CO_2 by coupling its emissions with those of CO, and this would be a good first step for the new inverse model.

Bibliography

Andreae, M. O. and P. Merlet, 2001: Emission of trace gases and aerosols from biomass burning. *Global Biogeochemical Cycles*, **15**, 955–966, doi:10.1029/2000GB001382.

Archer, D., 2007: Methane hydrate stability and anthropogenic climate change. *Biogeosciences*, **4**, 521–544.

Arnold, S. R., M. P. Chipperfield, and M. A. Blitz, 2005: A three-dimensional model study of the effect of new temperature-dependent quantum yields for acetone photolysis. *Journal of Geophysical Research*, **110**, D22305, doi:10.1029/2005JD005998.

Aydin, M., K. R. Verhulst, E. S. Saltzman, M. O. Battle, S. A. Montzka, D. R. Blake, Q. Tang, and M. J. Prather, 2011: Recent decreases in fossil-fuel emissions of ethane and methane derived from firn air. *Nature*, **476**, 198–201, doi:10.1038/nature10352.

Bergamaschi, P., C. Frankenberg, J. F. Meirink, M. Krol, F. Dentener, T. Wagner, U. Platt, J. O. Kaplan, S. Koerner, M. Heimann, E. J. Dlugokencky, and A. Goede, 2007: Satellite chartography of atmospheric methane from SCIAMACHY on board ENVISAT: 2. Evaluation based on inverse model simulations. *Journal of Geophysical Research*, **112**, doi:10.1029/2006JD007268.

Bergamaschi, P., C. Frankenberg, J. F. Meirink, M. Krol, M. G. Villani, S. Houweling, F. Dentener, E. J. Dlugokencky, J. B. Miller, L. V. Gatti, A. Engel, and I. Levin, 2009: Inverse modeling of global and regional CH_4 emissions using SCIAMACHY satellite retrievals. *Journal of Geophysical Research*, **114**, doi:10.1029/2009JD012287.

Bergamaschi, P., R. Hein, M. Heimann, and P. J. Crutzen, 2000: Inverse modeling of the global CO cycle 1. Inversion of CO mixing ratios. *Journal of Geophysical Research*, **105**, 1909–1927, doi:10.1029/1999JD900818.

Bergamaschi, P., M. Krol, F. Dentener, A. Vermeulen, F. Meinhardt, R. Graul, M. Ramonet, W. Peters, and E. J. Dlugokencky, 2005: Inverse modelling of national and european CH_4 emissions using the atmospheric zoom model TM5. *Atmospheric Chemistry and Physics*, **5**, 2431–2460.

Bergamaschi, P., M. Krol, J. F. Meirink, F. Dentener, A. Segers, J. van Aardenne, S. Monni, A. T. Vermeulen, M. Schmidt, M. Ramonet, C. Yver, F. Meinhardt, E. G. Nisbet, R. E. Fisher,

S. O'Doherty, and E. J. Dlugokencky, 2010: Inverse modeling of european CH_4 emissions 2001-2006. *Journal of Geophysical Research*, **115**, doi:10.1029/2010JD014180.

Bernard, B., J. Brooks, and W. Sackett, 1978: Light-hydrocarbons in recent Texas Continental-Shelf and slope sediments. *Journal of Geophysical Research*, **83**, 4053–4061, doi:10.1029/JC083iC08p04053.

Bingemer, H. and P. Crutzen, 1987: The production of methane from solid-wastes. *Journal of Geophysical Research*, **92**, 2181–2187, doi:10.1029/JD092iD02p02181.

Blaxter, K. and J. Czerkaws, 1966: Modifications of methane production of sheep by supplementation of its diet. *Journal of the Science of Food and Agriculture*, **17**, 417–&, doi:10.1002/jsfa.2740170907.

Bloom, A. A., P. I. Palmer, A. Fraser, D. S. Reay, and C. Frankenberg, 2010: Large-Scale controls of methanogenesis inferred from methane and gravity spaceborne data. *Science*, **327**, 322–325, doi:10.1126/science.1175176.

Bloss, W. J., M. J. Evans, J. D. Lee, R. Sommariva, D. E. Heard, and M. J. Pilling, 2005: The oxidative capacity of the troposphere: Coupling of field measurements of OH and a global chemistry transport model. *Faraday Discussions*, **130**, 425–436, doi:10.1039/b419090d.

Bogner, J., K. Spokas, E. Burton, R. Sweeney, and V. Corona, 1995: Landfills as atmospheric methane sources and sinks. *Chemosphere*, **31**, 4119–4130.

Bohn, T. J., D. P. Lettenmaier, K. Sathulur, L. C. Bowling, E. Podest, K. C. McDonald, and T. Friborg, 2007: Methane emissions from western siberian wetlands: heterogeneity and sensitivity to climate change. *Environmental Research Letters*, **2**, doi:10.1088/1748-9326/2/4/045015.

Bousquet, P., P. Ciais, J. B. Miller, E. J. Dlugokencky, D. A. Hauglustaine, C. Prigent, G. R. Van der Werf, P. Peylin, E. Brunke, C. Carouge, R. L. Langenfelds, J. Lathiere, F. Papa, M. Ramonet, M. Schmidt, L. P. Steele, S. C. Tyler, and J. White, 2006: Contribution of anthropogenic and natural sources to atmospheric methane variability. *Nature*, **443**, 439–443, doi:10.1038/nature05132.

Bousquet, P., P. Ciais, P. Peylin, M. Ramonet, and P. Monfray, 1999: Inverse modeling of annual atmospheric CO_2 sources and sinks 1. Method and control inversion. *Journal of Geophysical Research*, **104**, 26161–26178.

Bousquet, P., B. Ringeval, I. Pison, E. J. Dlugokencky, E. G. Brunke, C. Carouge, F. Chevallier, A. Fortems-Cheiney, C. Frankenberg, D. A. Hauglustaine, P. B. Krummel, R. L. Langenfelds, M. Ramonet, M. Schmidt, L. P. Steele, S. Szopa, C. Yver, N. Viovy, and P. Ciais, 2011: Source attribution of the changes in atmospheric methane for 2006-2008. *Atmospheric Chemistry and Physics*, **11**, 3689–3700, doi:10.5194/acp-11-3689-2011.

Breider, T. J., M. P. Chipperfield, N. A. D. Richards, K. S. Carslaw, G. W. Mann, and D. V. Spracklen, 2010: Impact of BrO on dimethylsulfide in the remote marine boundary layer. *Geophysical Research Letters*, **37**, 6 PP., doi:201010.1029/2009GL040868.

Brooks, J., D. Reid, and B. Bernard, 1981: Methane in the upper water column of the northwestern Gulf-of-Mexico. *Journal of Geophysical Research*, **86**, 1029–1040, doi:10.1029/JC086iC11p11029.

Brown, A. T., M. P. Chipperfield, C. Boone, C. Wilson, K. A. Walker, and P. Bernath, 2011: Trends in atmospheric halogen containing gases since 2004. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **112**, 2552–2566.

Bruhwiler, L. M. P., A. M. Michalak, W. Peters, D. F. Baker, and P. Tans, 2005: An improved kalman smoother for atmospheric inversions. *Atmospheric Chemistry and Physics*, **5**, 2691–2702.

Butenhoff, C. L. and M. A. K. Khalil, 2007: Global methane emissions from terrestrial plants. *Environmental Science & Technology*, **41**, 4032–4037, doi:10.1021/es062404i.

Butler, J. H., J. W. Elkins, T. M. Thompson, B. D. Hall, T. H. Swanson, and V. Koropalov, 1991: Oceanic consumption of CH_3CCl_3: Implications for tropospheric OH. *Journal of Geophysical Research*, **96**, PP. 22,347–22,355, doi:199110.1029/91JD02126.

Cao, M. K., S. Marshall, and K. Gregson, 1996: Global carbon exchange and methane emissions from natural wetlands: Application of a process-based model. *Journal of Geophysical Research*, **101**, 14399–14414, doi:10.1029/96JD00219.

Carver, G. D., P. D. Brown, and O. Wild, 1997: The ASAD atmospheric chemistry integration package and chemical reaction database. *Computer Physics Communications*, **105**, 197–215, doi:10.1016/S0010-4655(97)00056-8.

Chai, T., G. R. Carmichael, Y. Tang, A. Sandu, A. Heckel, A. Richter, and J. P. Burrows, 2009: Regional NO_x emission inversion through a four-dimensional variational approach using SCIAMACHY tropospheric NO_2 column observations. *Atmospheric Environment*, **43**, 5046–5055.

Chen, Y. H. and R. G. Prinn, 2005: Atmospheric modeling of high- and low-frequency methane observations: Importance of interannually varying transport. *Journal of Geophysical Research*, **110**, doi:10.1029/2004JD005542.

Chevallier, F., P. Bauer, J. F. Mahfouf, and J. J. Morcrette, 2002: Variational retrieval of cloud profile from ATOVS observations. *Quarterly Journal of the Royal Meteorological Society*, **128**, 2511–2525, doi:10.1256/qj.01.153.

Chevallier, F., F. Breon, and P. J. Rayner, 2007: Contribution of the Orbiting Carbon Observatory to the estimation of CO_2 sources and sinks: Theoretical study in a variational data assimilation framework. *Journal of Geophysical Research*, **112**, D09307.

Chevallier, F., M. Fisher, P. Peylin, S. Serrar, P. Bousquet, F. M. Breon, A. Chedin, and P. Ciais, 2005: Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data. *Journal of Geophysical Research - Atomspheres*, **110**, doi:10.1029/2005JD006390.

Chipperfield, M., 2006a: The TOMCAT/SLIMCAT Off-Line 3D CTM: A description of the model and a guide to using it. http://homepages.see.leeds.ac.uk/˜lecmc/papers/uniman0.8.pdf.

Chipperfield, M., D. Cariolle, P. Simon, R. Ramaroson, and D. Lary, 1993: A 3-Dimensional modeling study of trace species in the arctic lower stratosphere during winter 1989-1990. *Journal of Geophysical Research*, **98**, 7199–7218, doi:10.1029/92JD02977.

Chipperfield, M. P., 2006b: New version of the TOMCAT/SLIMCAT offline chemical transport model: Intercomparison of stratospheric tracer experiments. *Quarterly Journal of the Royal Meteorological Society*, **132**, 1179–1203, doi:10.1256/qj.05.51.

Chipperfield, M. P., B. V. Khattatov, and D. J. Lary, 2002: Sequential assimilation of stratospheric chemical observations in a three-dimensional model. *Journal of Geophysical Research*, **107**, 4585.

Christensen, T. R., I. C. Prentice, J. Kaplan, A. Haxeltine, and S. Sitch, 1996: Methane flux from northern wetlands and tundra - An ecosystem source modelling approach. *Tellus*, **48**, 652–661, doi:10.1034/j.1600-0889.1996.t01-4-00004.x.

Cofala, J., M. Amann, Z. Klimont, K. Kupiainen, and L. Hoeglund-Isaksson, 2007: Scenarios of global anthropogenic emissions of air pollutants and methane until 2030. *Atmospheric Environment*, **41**, 8486–8499, doi:10.1016/j.atmosenv.2007.07.010.

Conrad, R., 1989: Control of methane production in terrestrial ecosystems. *Exchange of Trace Gas between Terrestrial Ecosystems and the Atmosphere*, M. Andreae and D. Schimel, eds., John Wiley & Sons LTD, volume 47 of *Life Sciences Research Report*, 39–58, Berlin, Germany, Feb 19-24, 1989.

Courtier, P., J. Thepaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, **120**, 1367–1387, doi:10.1256/smsqj.51911.

Daniel, J. W., 1967: The conjugate gradient method for linear and nonlinear operator equations. *SIAM Journal on Numerical Analysis*, **4**, 10, doi:10.1137/0704002.

Dee, D., 1995: Testing the perfect-model assumption in variational data assimilation. WMO Symposium on Assimilation of Observations in Meteorology and Oceanography, Tokyo.

Denman, K., G. Brasseur, A. Chidthaisong, P. Ciais, P. Cox, R. Dickinson, D. Hauglustaine, C. Heinze, E. Holland, D. Jacob, U. Lohmann, S. Ramachandran, P. da Silva Dias, S. Wofsy, and X. Zhang, 2007: Couplings between changes in the climate system and biogeochemistry. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]]*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Denning, A. S., M. Holzer, K. R. Gurney, M. Heimann, R. M. Law, P. J. Rayner, I. Y. Fung, S. Fan, S. Taguchi, P. Friedlingstein, Y. Balkanski, J. Taylor, M. Maiss, and I. Levin, 1999: Threedimensional transport and concentration of SF_6. *Tellus*, **51**, 266–297, doi:10.1034/j.1600-0889.1999.00012.x.

Dentener, F., W. Peters, M. Krol, M. v. Weele, P. Bergamaschi, and J. Lelieveld, 2003: Interannual variability and trend of CH_4 lifetime as a measure for OH changes in the 1979-1993 time period. *Journal of Geophysical Research*, **108**, 12 PP., doi:200310.1029/2002JD002916.

Dickens, G. R., 2003: Rethinking the global carbon cycle with a large, dynamic and microbially mediated gas hydrate capacitor. *Earth and Planetary Science Letters*, **213**, 169–183, doi:10.1016/S0012-821X(03)00325-X.

Dlugokencky, E., K. Masaire, P. Lang, P. Tans, L. Steele, and E. Nisbet, 1994a: A dramatic decrease in the growth-rate of atmospheric methane in the northern hemisphere during 1992. *Geophysical Research Letters*, **21**, 45–48, doi:10.1029/93GL03070.

Dlugokencky, E., L. Steele, P. Lang, and K. Masarie, 1994b: The growth-rate and distribution of atmospheric methane. *Journal of Geophysical Research*, **99**, 17021–17043, doi:10.1029/94JD01245, 1st Scientific Conference of the International-Global-Atmospheric-Chemistry (IGAC)-Project, Israel Inst. Biol. Res., Elat, Israel, Apr 18-22, 1993.

Dlugokencky, E. J., L. Bruhwiler, J. W. C. White, L. K. Emmons, P. C. Novelli, S. A. Montzka, K. A. Masarie, P. M. Lang, A. M. Crotwell, J. B. Miller, and L. V. Gatti, 2009: Observational constraints on recent increases in the atmospheric CH_4 burden. *Geophysical Research Letters*, **36**, doi:10.1029/2009GL039780.

Dlugokencky, E. J., E. G. Dutton, P. C. Novelli, P. P. Tans, K. A. Masarie, K. O. Lantz, and S. Madronich, 1996: Changes in CH_4 and CO growth rates after the eruption of Mt Pinatubo and their link with changes in tropical tropospheric UV flux. *Geophysical Research Letters*, **23**, 2761–2764, doi:10.1029/96GL02638.

Dlugokencky, E. J., S. Houweling, L. Bruhwiler, K. A. Masarie, P. M. Lang, J. B. Miller, and P. P. Tans, 2003: Atmospheric methane levels off: Temporary pause or a new steady-state? *Geophysical Research Letters*, **30**, doi:10.1029/2003GL018126.

Dlugokencky, E. J., K. A. Masarie, P. M. Lang, and P. P. Tans, 1998: Continuing decline in the growth rate of the atmospheric methane burden. *Nature*, **393**, 447–450, doi:10.1038/30934.

Dlugokencky, E. J., E. G. Nisbet, R. Fisher, and D. Lowry, 2011: Global atmospheric methane: budget, changes and dangers. *Philosophical Transactions of the Royal Society*, **369**, 2058–2072, doi:10.1098/rsta.2010.0341.

Dlugokencky, E. J., B. P. Walter, K. A. Masarie, P. M. Lang, and E. S. Kasischke, 2001: Measurements of an anomalous global methane increase during 1998. *Geophysical Research Letters*, **28**, 499–502, doi:10.1029/2000GL012119.

Duncan, B. N., R. V. Martin, A. C. Staudt, R. Yevich, and J. A. Logan, 2003: Interannual and seasonal variability of biomass burning emissions constrained by satellite observations. *Journal of Geophysical Research*, **108**, doi:10.1029/2002JD002378.

Dutton, E. and J. Christy, 1992: Solar radiative forcing at selected locations and evidence for global lower tropospheric cooling following the eruptions of El-Chichon and Pinatubo. *Geophysical Research Letters*, **19**, 2313–2316, doi:10.1029/92GL02495.

Edwards, D. P., L. K. Emmons, J. C. Gille, A. Chu, J. L. Attie, L. Giglio, S. W. Wood, J. Haywood, M. N. Deeter, S. T. Massie, D. C. Ziskin, and J. R. Drummond, 2006: Satellite-observed pollution from southern hemisphere biomass burning. *Journal of Geophysical Research*, **111**, doi:10.1029/2005JD006655.

Engelen, R. J., A. S. Denning, K. R. Gurney, and T. Modelers, 2002: On error estimation in atmospheric CO_2 inversions. *Journal of Geophysical Research*, **107**, doi:10.1029/2002JD002195.

Etiope, G. and A. V. Milkov, 2004: A new estimate of global methane flux from onshore and shallow submarine mud volcanoes to the atmosphere. *Environmental Geology*, **46**, 997–1002, doi:10.1007/s00254-004-1085-1, 16th Congress of International-Union-for-Quaternary-Research, Reno, NV, JUL, 2003.

Feng, L., P. I. Palmer, H. Boesch, and S. Dance, 2009: Estimating surface CO_2 fluxes from space-borne CO_2 dry air mole fraction observations using an ensemble kalman filter. *Atmospheric Chemistry and Physics*, **9**, 2619–2633.

Feng, W., M. P. Chipperfield, S. Dhomse, B. M. Monge-Sanz, X. Yang, K. Zhang, and M. Ramonet, 2011: Evaluation of cloud convection and tracer transport in a three-dimensional chemical transport model. *Atmospheric Chemistry and Physics*, **11**, 5783–5803, doi:10.5194/acp-11-5783-2011.

Ferretti, D. F., J. B. Miller, J. W. C. White, D. M. Etheridge, K. R. Lassey, D. C. Lowe, C. M. M. Meure, M. F. Dreier, C. M. Trudinger, T. D. van Ommen, and R. L. Langenfelds, 2005: Unexpected changes to the global methane budget over the past 2000 years. *Science*, **309**, 1714 –1717, doi:10.1126/science.1115193.

Fiore, A. M., L. W. Horowitz, E. J. Dlugokencky, and J. J. West, 2006: Impact of meteorology and emissions on methane trends, 1990-2004. *Geophysical Research Letters*, **33**, doi:10.1029/2006GL026199.

Forster, P., V. Ramaswamy, P. Artaxo, T. Berntsen, R. Betts, D. W. Fahey, J. Haywood, J. Lean, D. C. Lowe, G. Myhre, J. Nganga, R. Prinn, G. Raga, M. Schulz, and R. Van Dorland, 2007: Changes in atmospheric constituents and in radiative forcing. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]]*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Fung, I., J. John, J. Lerner, E. Matthews, M. Prather, L. P. Steele, and P. J. Fraser, 1991: Three-dimensional model synthesis of the global methane cycle. *Journal of Geophysical Research*, **96**, 13033–13065.

Gedney, N., P. M. Cox, and C. Huntingford, 2004: Climate feedback from wetland methane emissions. *Geophysical Research Letters*, **31**, doi:10.1029/2004GL020919.

Giering, R., 2000: Tangent linear and adjoint biogeochemical models. *Inverse Methods in Global Biogeochemical Cycles*, American Geophysical Union.

Gilbert, J. and C. Lemarechal, 1989: Some numerical experiments with Variable-Storage Quasi-Newton algorithms. *Mathematical Programming*, **45**, 407–435, doi:10.1007/BF01589113.

Gloor, M., E. Dlugokencky, C. Brenninkmeijer, L. Horowitz, D. F. Hurst, G. Dutton, C. Crevoisier, T. Machida, and P. Tans, 2007: Three-dimensional SF_6 data and tropospheric transport simulations: Signals, modeling accuracy, and implications for inverse modeling. *Journal of Geophysical Research*, **112**, 17 PP., doi:200710.1029/2006JD007973.

Gunn, L. N., 2008: *A new retrieval scheme to determine tropospheric composition from the global ozone monitoring experiment*. Ph.D. thesis, University of Leeds, Leeds, UK.

Hall, T. M. and D. W. Waugh, 1998: Influence of nonlocal chemistry on tracer distributions: Inferring the mean age of air from SF_6. *Journal of Geophysical Research*, **103**, PP. 13,327–13,336, doi:199810.1029/98JD00170.

Hansen, J., R. Ruedy, J. Glascoe, and M. Sato, 1999: GISS analysis of surface temperature change. *Journal of Geophysical Research*, **104**, 30997–31022, doi:10.1029/1999JD900835.

Hao, W. and D. Ward, 1993: Methane production from global biomass burning. *Journal of Geophysical Research*, **98**, 20657–20661, doi:10.1029/93JD01908.

Harnisch, J. and A. Eisenhauer, 1998: Natural CF_4 and SF_6 on earth. *Geophysical Research Letters*, **25**, PP. 2401–2404, doi:199810.1029/98GL01779.

Harwood, R. and J. Pyle, 1980: The dynamical behavior of a two-dimensional model of the stratosphere. *Quarterly Journal of the Royal Meteorological Society*, **106**, 395–420, doi:10.1256/smsqj.44902.

Heard, D. E. and M. J. Pilling, 2003: Measurement of OH and HO_2 in the troposphere. *Chemical Reviews*, **103**, 5163–5198, doi:10.1021/cr020522s.

Heimann, M., 2011: Atmospheric science: Enigma of the recent methane budget. *Nature*, **476**, 157–158, doi:10.1038/476157a.

Hein, R., P. J. Crutzen, and M. Heimann, 1997: An inverse modeling approach to investigate the global atmospheric methane cycle. *Global Biogeochemical Cycles*, **11**, 43–76, doi:10.1029/96GB03043.

Henze, D. K., A. Hakami, and J. H. Seinfeld, 2007: Development of the adjoint of GEOS-Chem. *Atmos. Chem. Phys.*, **7**, 2413–2433, doi:10.5194/acp-7-2413-2007.

Holton, J., P. Haynes, M. McIntyre, A. Douglass, R. Rood, and L. Pfister, 1995: Stratosphere-Troposphere exchange. *Reviews of Geophysics*, **33**, 403–439, doi:10.1029/95RG02097.

Holtslag, A. and B. Boville, 1993: Local versus nonlocal Boundary-Layer diffusion in a global climate model. *Journal of Climate*, **6**, 1825–1842, doi:10.1175/1520-0442(1993)006.

Hooghiemstra, P. B., M. C. Krol, J. F. Meirink, P. Bergamaschi, G. R. van der Werf, P. C. Novelli, I. Aben, and T. Rckmann, 2011: Optimizing global CO emission estimates using a four-dimensional variational data assimilation system and surface network observations. *Atmospheric Chemistry and Physics*, **11**, 4705–4723.

Hossaini, R., M. P. Chipperfield, B. M. Monge-Sanz, N. A. D. Richards, E. Atlas, and D. R. Blake, 2010: Bromoform and dibromomethane in the tropics: a 3-D model study of chemistry and transport. *Atmospheric Chemistry and Physics*, **10**, 719–735.

Houghton, J., Y. Ding, D. Griggs, N. M., P. van der Linden, X. Dai, K. Maskell, and C. Johnson, 2001: *Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, 2007*. Cambridge University Press.

Hourdin, F. and O. Talagrand, 2006a: Eulerian backtracking of atmospheric tracers. I: Adjoint derivation and parametrization of subgrid-scale transport. *Quarterly Journal of the Royal Meteorological Society*, **132**, 567–583.

— 2006b: Eulerian backtracking of atmospheric tracers. II: Numerical aspects. *Quarterly Journal of the Royal Meteorological Society*, **132**, 585–603.

Houweling, S., T. Kaminski, F. Dentener, J. Lelieveld, and M. Heimann, 1999: Inverse modeling of methane sources and sinks using the adjoint of a global transport model. *Journal of Geophysical Research*, **104**, 160.

Houweling, S., T. Roeckmann, I. Aben, F. Keppler, M. Krol, J. F. Meirink, E. J. Dlugokencky, and C. Frankenberg, 2006: Atmospheric constraints on global emissions of methane from plants. *Geophysical Research Letters*, **33**, doi:10.1029/2006GL026162.

Huijnen, V., J. Williams, M. van Weele, T. van Noije, M. Krol, F. Dentener, A. Segers, S. Houweling, W. Peters, J. de Laat, F. Boersma, P. Bergamaschi, P. van Velthoven, P. Le Sager, H. Eskes, F. Alkemade, R. Scheele, P. Ndlec, and H. Ptz, 2010: The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0. *Geoscientific Model Development*, **3**, 445–473, doi:10.5194/gmd-3-445-2010.

Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential and variational. *Journal of the Meteorological Society of Japan*, **75**, 181–189, 2nd WMO International Symposium on Assimilation of Observations in Meteorology and Oceanography, Tokyo, Japan, Mar 13-17, 1995.

Ito, A., 2010: Changing ecophysiological processes and carbon budget in East Asian ecosystems under near-future changes in climate: implications for long-term monitoring from a process-based model. *Journal of Plant Research*, **123**, 577–588, doi:10.1007/s10265-009-0305-x.

Jacob, D., 1999: *Introduction to Atmospheric Chemistry*. Princeton University Press, Princeton, New Jersey.

Jacob, D., J. Crawford, H. Maring, A. Clarke, J. Dibb, L. Emmons, R. Ferrare, C. Hostetler, P. Russell, H. Singh, A. Thompson, G. Shaw, E. McCauley, J. Pederson, and J. Fisher, 2010: The Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission: design, execution and first results. *Atmospheric Chemistry and Physics*, **10**, 5191–5212, doi:10.5194/acp-10-5191-2010.

Jansen, E., J. Overpeck, K. R. Briffa, J. Duplessy, F. Joos, V. Masson-Delmotte, D. Olago, B. Otto-Bliesner, W. R. Peltier, S. Rahmstorf, R. Ramesh, D. Raynaud, D. Rind, O. Solomina, R. Villalba, and D. Zhang, 2007: Palaeoclimate. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]]*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Johnson, D. E. and G. M. Ward, 1996: Estimates of animal methane emissions. *Environmental Monitoring and Assessment*, **42**, 133–141.

Kai, F. M., S. Tyler, and J. Randerson, 2010: Modeling methane emissions from rice agriculture in china during 1961-2007. *Journal of Integrative Environmental Science*, **7**, 49–60.

Kalman, R., 1960: A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, **82**, 35—45.

Kaminski, T., P. J. Rayner, M. Heimann, and I. G. Enting, 2001: On aggregation errors in atmospheric transport inversions. *Journal of Geophysical Research - Atomspheres*, **106**, 4703–4715, doi:10.1029/2000JD900581.

Kanakidou, M., F. J. Dentener, and P. J. Crutzen, 1995: A global three-dimensional study of the fate of HCFCs and HFC-134a in the troposphere. *Journal of Geophysical Research*, **100**, PP. 18,781–18,801, doi:199510.1029/95JD01919.

Karl, D. and B. Tilbrook, 1994: Production and transport of methane in oceanic particulate organic-matter. *Nature*, **368**, 732–734, doi:10.1038/368732a0.

Keppler, F., J. T. G. Hamilton, M. Brass, and T. Rockmann, 2006: Methane emissions from terrestrial plants under aerobic conditions. *Nature*, **439**, 187–191, doi:10.1038/nature04420.

Khalil, M. A. K., C. L. Butenhoff, and R. A. Rasmussen, 2007: Atmospheric methane: Trends and cycles of sources and sinks. *Environmental Science & Technology*, **41**, 2131–2137, doi:10.1021/es061791t.

Klauda, J. B. and S. I. Sandler, 2009: Global distribution of methane hydrate in ocean sediment. *Energy and Fuels*, **19**, 459–470.

Krol, M., P. J. v. Leeuwen, and J. Lelieveld, 1998: Global OH trend inferred from methylchloroform measurements. *Journal of Geophysical Research*, **103**, 711.

Kvenvolden, K. A. and B. W. Rogers, 2005: Gaia's breath - global methane exhalations. *Marine and Petroleum Geology*, **22**, 579–590, doi:10.1016/j.marpetgeo.2004.08.004, symposium on Near-Surface Hydrocarbon Migration, Vancouver, Canada, Apr 07-10, 2002.

Lambert, G. and S. Schmidt, 1993: Reevaluation of the oceanic flux of methane - uncertainties and long-term variations. *Chemosphere Global Change Science*, **26**, 579–589, doi:10.1016/0045-6535(93)90443-9, NATO Advanced Research Workshop on Atmospheric Methane : Sources, Sinks and Role in Global Change, Mt Hood, OR, Oct 07-11, 1991.

Law, K. S. and J. A. Pyle, 1993: Modeling trace gas budgets in the troposphere 1. Ozone and odd nitrogen. *Journal of Geophysical Research*, **98**, 400.

Lawrence, M. G., P. Jckel, and R. von Kuhlmann, 2001: What does the global mean OH concentration tell us? *Atmospheric Chemistry and Physics*, **1**, 37–49.

Levin, I., T. Naegler, R. Heinz, D. Osusko, E. Cuevas, A. Engel, J. Ilmberger, R. L. Langenfelds, B. Neininger, C. v. Rohden, L. P. Steele, R. Weller, D. E. Worthy, and S. A. Zimov, 2010: The global SF_6 source inferred from long-term high precision atmospheric measurements and its comparison with emission inventories. *Atmospheric Chemistry and Physics*, **10**, 2655–2662.

Li, C. S., J. J. Qiu, S. Frolking, X. M. Xiao, W. Salas, B. Moore, S. Boles, Y. Huang, and R. Sass, 2002: Reduced methane emissions from large-scale changes in water management of China's rice paddies during 1980-2000. *Geophysical Research Letters*, **29**, doi:10.1029/2002GL015370.

Lintner, B. R., W. Buermann, C. D. Koven, and I. Y. Fung, 2006: Seasonal circulation and mauna loa CO_2 variability. *Journal of Geophysical Research*, **111**, 18 PP., doi:200610.1029/2005JD006535.

Lobert, J., D. Scharffe, W. Hao, T. Kuhlbusch, R. Seuwen, P. Warneck, and P. Crutzen, 1991: Experimental evaluation of biomass burning emissions - Nitrogen and carbon containing compounds. *Global Biomass Burning: Atmospheric, Climatic, and Biospheric Implications*, J. Levine, ed., M.I.T Press, 289–304, williamsberg, VA, Mar 19-23, 1990.

Louis, J., 1979: Parametric model of vertical eddy fluxes in the atmosphere. *Boundary-Layer Meteorology*, **17**, 187–202, doi:10.1007/BF00117978.

Maiss, M. and C. A. M. Brenninkmeijer, 1998: Atmospheric SF_6: trends, sources, and prospects. *Environmental Science & Technology*, **32**, 3077–3086, doi:10.1021/es9802807.

Matthews, E. and I. Fung, 1987: Methane emission from natural wetlands: Global distribution, area, and environmental characteristics of sources. *Global Biogeochemical Cycles*, **1**, PP. 61–86, doi:198710.1029/GB001i001p00061.

Mauldin III, R. L., F. L. Eisele, C. A. Cantrell, E. Kosciuch, B. A. Ridley, B. Lefer, D. J. Tanner, J. B. Nowak, G. Chen, L. Wang, and D. Davis, 2001: Measurements of OH aboard the NASA P-3 during PEM-Tropics B. *Journal of Geophysical Research*, **106**, PP. 32,657–32,666, doi:200110.1029/2000JD900832.

Meehl, G. A., T. F. Stocker, W. D. Collins, P. Friedlingstein, A. T. Gaye, J. M. Gregory, A. Kitoh, R. Knutti, J. M. Murphy, A. Noda, S. C. B. Raper, I. G. Watterson, A. J. Weaver, and Z. Zhao, 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]]*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Meirink, J. F., P. Bergamaschi, C. Frankenberg, M. T. S. d'Amelio, E. J. Dlugokencky, L. V. Gatti, S. Houweling, J. B. Miller, T. Roeckmann, M. G. Villani, and M. C. Krol, 2008a: Four-dimensional variational data assimilation for inverse modeling of atmospheric methane emissions: Analysis of SCIAMACHY observations. *Journal of Geophysical Research - Atomspheres*, **113**, doi:10.1029/2007JD009740.

Meirink, J. F., P. Bergamaschi, and M. C. Krol, 2008b: Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: Method and comparison with synthesis inversion. *Atmospheric Chemistry and Physics*, **8**, 6341–6353.

Meirink, J. F., H. J. Eskes, and A. P. H. Goede, 2006: Sensitivity analysis of methane emissions derived from SCIAMACHY observations through inverse modelling. *Atmospheric Chemistry and Physics*, **6**, 1275–1292.

Mikaloff Fletcher, S. E., P. P. Tans, L. M. Bruhwiler, J. B. Miller, and M. Heimann, 2004: CH_4 sources estimated from atmospheric observations of CH_4 and its 13C/12C isotopic ratios: 2. Inverse modeling of CH_4 fluxes from geographical regions. *Global Biogeochemical Cycles*, **18**, doi:10.1029/2004GB002224.

Milkov, A. V., R. Sassen, T. V. Apanasovich, and F. G. Dadashev, 2003: Global gas flux from mud volcanoes: A significant source of fossil methane in the atmosphere and the ocean. *Geophysical Research Letters*, **30**, doi:10.1029/2002GL016358.

Mitchell, T. D. and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*, **25**, 693–712, doi:10.1002/joc.1181.

Monge-Sanz, B. M., M. P. Chipperfield, A. J. Simmons, and S. M. Uppala, 2007: Mean age of air and transport in a CTM: Comparison of different ECMWF analyses. *Geophysical Research Letters*, **34**, L04801.

Monteil, G., S. Houweling, E. J. Dlugockenky, G. Maenhout, B. H. Vaughn, J. W. C. White, and T. Rockmann, 2011: Interpreting methane variations in the past two decades using measurements of CH_4 mixing ratio and isotopic composition. *Atmospheric Chemistry and Physics Discussions*, **11**, 6771–6803, doi:10.5194/acpd-11-6771-2011.

Montzka, S. A., M. Krol, E. Dlugokencky, B. Hall, P. Joeckel, and J. Lelieveld, 2011a: Small interannual variability of global atmospheric hydroxyl. *Science*, **331**, 67–69, doi:10.1126/science.1197640.

Montzka, S. A., M. Krol, E. Dlugokencky, B. Hall, P. Jckel, and J. Lelieveld, 2011b: Small interannual variability of global atmospheric hydroxyl. *Science*, **331**, 67 –69, doi:10.1126/science.1197640.

Montzka, S. A., C. M. Spivakovsky, J. H. Butler, J. W. Elkins, L. T. Lock, and D. J. Mondeel, 2000: New observational constraints for atmospheric hydroxyl on global and hemispheric scales. *Science*, **288**, 500 –503, doi:10.1126/science.288.5465.500.

Morris, R. A., T. M. Miller, A. A. Viggiano, J. F. Paulson, S. Solomon, and G. Reid, 1995: Effects of electron and ion reactions on atmospheric lifetimes of fully fluorinated compounds. *Journal of Geophysical Research*, **100**, PP. 1287–1294, doi:199510.1029/94JD02399.

Nazareth, L., 1979: Relationship between the BFGS and conjugate gradient algorithms and its implications for new algorithms. *SIAM Journal on Numerical Analysis*, **16**, 794–800, doi:10.1137/0716059.

Nehrkorn, T., G. D. Modica, M. Cerniglia, F. H. Ruggiero, J. G. Michalakes, and X. Zou, 2006: MM5 adjoint development using TAMC: Experiences with an automatic code generator. National Center for Atmospheric Research, Boulder, CA, 2002, P6.18.

Nocedal, J., 1980: Updating Quasi-Newton matrices with limited storage. *Mathematics of Computation*, **35**, 773–782, doi:10.2307/2006193.

O'Connor, F. M., O. Boucher, N. Gedney, C. D. Jones, G. A. Folberth, R. Coppell, P. Friedlingstein, W. J. Collins, J. Chappellaz, J. Ridley, and C. E. Johnson, 2010: Possible role of wetlands, permafrost, and methane hydrates in the methane cycle under future climate change: A review. *Reviews of Geophysics*, **48**, doi:10.1029/2010RG000326.

Olivier, J., 2002: Part III: Greenhouse gas emissions: 1. Shares and trends in greenhouse gas emissions; 2. Sources and methods; greenhouse gas emissions for 1990 and 1995. *CO_2 emissions from fuel combustion 1971 - 2000*, International Energy Agency (IEA), Paris, 2002 edition edition, III.1–III.31.

Olivier, J. and J. Berdowski, 2001: Global emissions sources and sinks. *The Climate System*, A. A. Balkem Publishers/Swets & Zeitlinger Publishers, Lisse, The Netherlands, 33–78, edited by J. Berdowski, R. Guicherit and B. J. Heij.

Palmer, P. I., P. Suntharalingam, D. B. A. Jones, D. J. Jacob, D. G. Streets, Q. Fu, S. A. Vay, and G. W. Sachse, 2006: Using CO_2:CO correlations to improve inverse analyses of carbon fluxes. *Journal of Geophysical Research*, **111**, D12318, doi:10.1029/2005JD006697.

Patra, P. K., S. Houweling, M. Krol, P. Bousquet, D. Belikov, D. Bergmann, H. Bian, P. Cameron-Smith, M. P. Chipperfield, K. Corbin, A. Fortems-Cheiney, A. Fraser, E. Gloor, P. Hess, A. Ito, S. R. Kawa, R. M. Law, Z. Loh, S. Maksyutov, L. Meng, P. I. Palmer, R. G. Prinn, M. Rigby, R. Saito, and C. Wilson, 2011: TransCom model simulations of CH4 and related species: linking transport, surface flux and chemical loss with CH4 variability in the troposphere and lower stratosphere. *Atmospheric Chemistry and Physics Discussions*, **11**, 18767–18821, doi:10.5194/acpd-11-18767-2011.

Patra, P. K., M. Takigawa, K. Ishijima, B. Choi, D. Cunnold, E. J. Dlugokencky, P. Fraser, A. J. Gomez-Pelaez, T. Goo, J. Kim, P. Krummel, and R. Langenfelds, 2009: Growth rate, seasonal, synoptic, diurnal variations and budget of methane in the lower atmosphere. *Journal of the Meteorological Society of Japan*, **87**, 635–663, doi:10.2151/jmsj.87.635.

Peters, W., M. C. Krol, E. J. Dlugokencky, F. J. Dentener, P. Bergamaschi, G. Dutton, P. v. Velthoven, J. B. Miller, L. Bruhwiler, and P. P. Tans, 2004: Toward regional-scale modeling using the two-way nested global model TM5: characterization of transport using SF_6. *Journal of Geophysical Research*, **109**, 17 PP., doi:200410.1029/2004JD005020.

Peters, W., J. B. Miller, J. Whitaker, A. S. Denning, A. Hirsch, M. C. Krol, D. Zupanski, L. Bruhwiler, and P. P. Tans, 2005: An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations. *Journal of Geophysical Research*, **110**, D24304.

Prather, M., 1986: Numerical advection by conservation of 2nd-Order moments. *Journal of Geophysical Research*, **91**, 6671–6681, doi:10.1029/JD091iD06p06671.

Prather, M., M. McElroy, S. Wofsy, G. Russell, and D. Rind, 1987: Chemistry of the global troposphere - Fluorocarbons as tracers of air motion. *Journal of Geophysical Research*, **92**, 6579–6613, doi:10.1029/JD092iD06p06579.

Prigent, C., F. Papa, F. Aires, W. B. Rossow, and E. Matthews, 2007: Global inundation dynamics inferred from multiple satellite observations, 1993-2000. *Journal of Geophysical Research*, **112**, doi:10.1029/2006JD007847.

Prinn, R. G., J. Huang, R. F. Weiss, D. M. Cunnold, P. J. Fraser, P. G. Simmonds, A. McCulloch, C. Harth, S. Reimann, P. Salameh, S. O'Doherty, R. H. J. Wang, L. W. Porter, B. R. Miller, and P. B. Krummel, 2005: Evidence for variability of atmospheric hydroxyl radicals over the past quarter century. *Geophysical Research Letters*, **32**, doi:10.1029/2004GL022228.

Prinn, R. G., J. Huang, R. F. Weiss, D. M. Cunnold, P. J. Fraser, P. G. Simmonds, A. McCulloch, C. Harth, P. Salameh, S. O'Doherty, R. H. J. Wang, L. Porter, and B. R. Miller, 2001: Evidence for substantial variations of atmospheric hydroxyl radicals in the past two decades. *Science*, **292**, 1882 –1888, doi:10.1126/science.1058673.

Prinn, R. G., R. F. Weiss, B. R. Miller, J. Huang, F. N. Alyea, D. M. Cunnold, P. J. Fraser, D. E. Hartley, and P. G. Simmonds, 1995: Atmospheric trends and lifetime of CH_3CCI_3 and global OH concentrations. *Science*, **269**, 187 –192, doi:10.1126/science.269.5221.187.

Raper, J. L., M. M. Kleb, D. J. Jacob, D. D. Davis, R. E. Newell, H. E. Fuelberg, R. J. Bendura, J. M. Hoell, and R. J. McNeal, 2001: Pacific exploratory mission in the tropical pacific: PEM-Tropics b, March-April 1999. *Journal of Geophysical Research*, **106**, 425.

Ravishankara, A. R., S. Solomon, A. A. Turnipseed, and R. F. Warren, 1993: Atmospheric lifetimes of long-lived halogenated species. *Science*, **259**, 194 –199, doi:10.1126/science.259.5092.194.

Reddmann, T., R. Ruhnke, and W. Kouker, 2001: Three-dimensional model simulations of SF_6 with mesospheric chemistry. *Journal of Geophysical Research*, **106**, PP. 14,525–14,537, doi:200110.1029/2000JD900700.

Reshetnikov, A. I., N. N. Paramonova, and A. A. Shashkov, 2000: An evaluation of historical methane emissions from the soviet gas industry. *Journal of Geophysical Research - Atomspheres*, **105**, 3517–3529, doi:10.1029/1999JD900761.

Ridgwell, A. J., S. J. Marshall, and K. Gregson, 1999: Consumption of atmospheric methane by soils: A process-based model. *Global Biogeochemical Cycles*, **13**, 59–70.

Rigby, M., R. G. Prinn, P. J. Fraser, P. G. Simmonds, R. L. Langenfelds, J. Huang, D. M. Cunnold, L. P. Steele, P. B. Krummel, R. F. Weiss, S. O'Doherty, P. K. Salameh, H. J. Wang, C. M. Harth, J. Muehle, and L. W. Porter, 2008: Renewed growth of atmospheric methane. *Geophysical Research Letters*, **35**, doi:10.1029/2008GL036037.

Ringeval, B., N. d. Noblet-Ducoudr, P. Ciais, P. Bousquet, C. Prigent, F. Papa, and W. B. Rossow, 2010: An attempt to quantify the impact of changes in wetland extent on methane emissions on the seasonal and interannual time scales. *Global Biogeochemical Cycles*, **24**, 12 PP., doi:201010.1029/2008GB003354.

Russell, G. and J. Lerner, 1981: A new Finite-Differencing scheme for the tracer Transport-Equation. *Journal of Applied Meteorology*, **20**, 1483–1498, doi:10.1175/1520-0450(1981)020.

Sackett, W. and J. Brooks, 1975: Origin and distributions of low-molecular-weight hydrocarbons in Gulf of Mexico coastal waters. *Abstracts of Papers of the American Chemical Society*, 17.

Scheehle, E. A., W. N. Irving, and D. Kruger, 2002: Global anthropogenic methane emissions. *Non-CO2 Greenhouse Gases Scientific Understanding, Control Options and Policy Aspects*, J. VanHam, A. P. M. Baede, R. Guicherit, and J. G. F. WilliamsJacobse, eds., Millpress Science Publishers, PO BOX 84118, 3009 CC Rotterdam, Netherlands, 257–262, 3rd International Symposium on Non-CO2 Greenhouse Gases, Maastricht, Netherlands, Jan 21-23, 2002.

Scranton, M. and P. Brewer, 1977: Occurence of methane in near-surface waters of western subtropical north-Atlantic. *Deep-Sea Research*, **24**, 127–138, doi:10.1016/0146-6291(77)90548-3.

Shewchuck, R., J., 1994: An introduction to the conjugate gradient method without the agonizing pain. www.cs.cmu.edu/˜jrs/jrspapers.html.

Simpson, I. J., D. R. Blake, F. S. Rowland, and T. Y. Chen, 2002: Implications of the recent fluctuations in the growth rate of tropospheric methane. *Geophysical Research Letters*, **29**, doi:10.1029/2001GL014521.

Simpson, I. J., F. S. Rowland, S. Meinardi, and D. R. Blake, 2006: Influence of biomass burning during recent fluctuations in the slow growth of global tropospheric methane. *Geophysical Research Letters*, **33**, doi:10.1029/2006GL027330.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, 2007: *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007*. Cambridge University Press.

Spahni, R., J. Chappellaz, T. F. Stocker, L. Loulergue, G. Hausammann, K. Kawamura, J. Flckiger, J. Schwander, D. Raynaud, V. Masson-Delmotte, and J. Jouzel, 2005: Atmospheric methane and nitrous oxide of the late pleistocene from antarctic ice cores. *Science*, **310**, 1317 –1321, doi:10.1126/science.1120132.

Spivakovsky, C. M., J. A. Logan, S. A. Montzka, Y. J. Balkanski, M. Foreman-Fowler, D. B. A. Jones, L. W. Horowitz, A. C. Fusco, C. A. M. Brenninkmeijer, M. J. Prather, S. C. Wofsy, and M. B. McElroy, 2000: Three-dimensional climatological distribution of tropospheric OH: update and evaluation. *Journal of Geophysical Research*, **105**, PAGES 89318980.

Stockwell, D. Z. and M. P. Chipperfield, 1999: A tropospheric chemical-transport model: Development and validation of the model transport schemes. *Quarterly Journal of the Royal Meteorological Society*, **125**, 1747–1783, doi:10.1256/smsqj.55713.

Sugimoto, A., T. Inoue, I. Tayasu, L. Miller, S. Takeichi, and T. Abe, 1998: Methane and hydrogen production in a termite-symbiont system. *Ecological Research*, **13**, 241–257, doi:10.1046/j.1440-1703.1998.00262.x.

Talagrand, O. and F. Bouttier, 1999: Internal diagnostics of data assimilation systems. Workshop on Diagnostics of Models and Data Assimilation Systems, ECMWF, 407–409.

Tan, D., I. Faloona, J. B. Simpas, W. Brune, J. Olson, J. Crawford, M. Avery, G. Sachse, S. Vay, S. Sandholm, H. Guan, T. Vaughn, J. Mastromarino, B. Heikes, J. Snow, J. Podolske, and H. Singh, 2001: OH and HO_2 in the tropical pacific: Results from PEM-Tropics B. *Journal of Geophysical Research*, **106**, PP. 32, 667–32, 681, doi:200110.1029/2001JD900002.

Tarantola, A. and B. Valette, 1982a: Generalized non-linear inverse problems solved using the least-squares criterion. *Reviews of Geophysics*, **20**, 219–232, doi:10.1029/RG020i002p00219.

— 1982b: Inverse problems = Quest for information. *Journal of Geophysics*, **50**, 159–170.

Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in Large-Scale models. *Monthly Weather Review*, **117**, 1779–1800, doi:10.1175/1520-0493(1989)117.

Tremolet, Y., 2007: Model-error estimation in 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, **133**, 1267–1280, doi:10.1002/qj.94.

Trenberth, K. E. and A. Dai, 2007: Effects of Mount Pinatubo volcanic eruption on the hydrological cycle as an analog of geoengineering. *Geophysical Research Letters*, **34**, doi:10.1029/2007GL030524.

van der Werf, G. R., J. T. Randerson, L. Giglio, G. J. Collatz, P. S. Kasibhatla, and J. Arellano, 2006: Interannual variability in global biomass burning emissions from 1997 to 2004. *Atmospheric Chemistry and Physics*, **6**, 3423–3441.

Velders, G., 1995: Description of the RIVM 2-dimensional stratosphere model. Technical Report RIVM Report 722201002.

Villani, M. G., P. Bergamaschi, M. Krol, J. F. Meirink, and F. Dentener, 2010: Inverse modeling of european CH_4 emissions: sensitivity to the observational network. *Atmospheric Chemistry and Physics*, **10**, 1249–1267.

Walter, B. P. and M. Heimann, 2000: A process-based, climate-sensitive model to derive methane emissions from natural wetlands: Application to five wetland sites, sensitivity to model parameters, and climate. *Global Biogeochemical Cycles*, **14**, 745–765, doi:10.1029/1999GB001204.

Walter, K. M., L. C. Smith, and I. I. I. Chapin, 2007: Methane bubbling from northern lakes: Present and future contributions to the global methane budget. *Philosophical Transactions of the Royal Society*, **365**, 1657–1676, doi:10.1098/rsta.2007.2036, discussion Meeting on Trace Gas Biogeochemistry and Global Change, Royal Soc, London, England, Nov 13-14, 2006.

Wang, K. Y., J. A. Pyle, M. G. Sanderson, and C. Bridgeman, 1999: Implementation of a convective atmospheric boundary layer scheme in a tropospheric chemistry transport model. *Journal of Geophysical Research*, **104**, 23729–23745, doi:10.1029/1999JD900383.

Wheeler, G. S., M. Tokoro, R. H. Scheffrahn, and N. Y. Su, 1996: Comparative respiration and methane production rates in nearctic termites. *Journal of Insect Physiology*, **42**, 799–806, doi:10.1016/0022-1910(96)00002-9.

Wolin, M., 1981: Fermentation in the rumen and human large-intestine. *Science*, **213**, 1463–1468, doi:10.1126/science.7280665.

Wuebbles, D. J. and K. Hayhoe, 2002: Atmospheric methane and global change. *Earth-Science Reviews*, **57**, 177–210, doi:10.1016/S0012-8252(01)00062-9.

Yan, X., H. Akiyama, K. Yagi, and H. Akimoto, 2009: Global estimations of the inventory and mitigation potential of methane emissions from rice cultivation conducted using the 2006 intergovernmental panel on climate change guidelines. *Global Biogeochemical Cycles*, **23**, doi:10.1029/2008GB003299.

Yver, C. E., I. C. Pison, A. Fortems-Cheiney, M. Schmidt, F. Chevallier, M. Ramonet, A. Jordan, O. A. Svde, A. Engel, R. E. Fisher, D. Lowry, E. G. Nisbet, I. Levin, S. Hammer, J. Necki, J. Bartyzel, S. Reimann, M. K. Vollmer, M. Steinbacher, T. Aalto, M. Maione, J. Arduini, S. O'Doherty, A. Grant, W. T. Sturges, G. L. Forster, C. R. Lunder, V. Privalov, N. Paramonova, A. Werner, and P. Bousquet, 2011: A new estimation of the recent tropospheric molecular hydrogen budget using atmospheric observations and variational inversion. *Atmos. Chem. Phys.*, **11**, 3375–3392.

Zimov, S. A., Y. V. Voropaev, I. P. Semiletov, S. P. Davidov, S. F. Prosiannikov, F. S. Chapin, M. C. Chapin, S. Trumbore, and S. Tyler, 1997: North siberian lakes: A methane source fueled by pleistocene carbon. *Science*, **277**, 800–802, doi:10.1126/science.277.5327.800.