THE UNIVERSITY OF READING

Department of Meteorology

The Tropopause and the Climate Impact of Water Vapour Emissions from Aircraft

Laura Jenny Wilcox

A thesis submitted for the degree of Doctor of Philosophy

December 2010

Declaration

I confirm that this is my own work and that the use of all material from other sources has been properly and fully acknowledged.

Laura Jenny Wilcox

Abstract

The overall aim of this thesis is to improve the estimate of the climate forcing associated with the accumulation of water vapour in the stratosphere due to aviation emissions.

Three emissions inventories have been used with a fully Lagrangian trajectory model to calculate the accumulation in the stratosphere of water vapour emissions from aircraft, and their resulting radiative forcing.

The annual and global-mean radiative forcing due to present-day aviation water vapour emissions has been found to be 0.86 ± 0.48 mW m^{-2}. This forcing is sensitive to the vertical distribution of the emissions, and, to a lesser extent, interannual variability in meteorology. It is an order of magnitude smaller than recent estimates of the radiative forcing due to aviation CO_2. A recently published assessment gave a best estimate of water vapour forcing of 2.8 mW m^{-2}, and a 90% likelihood range from 0.4 to 20 mW m^{-2}. The detailed calculations in this thesis indicate that even this best estimate is a significant overestimate, and the 90% likelihood range is too large.

The calculated accumulation of stratospheric emissions, and therefore the estimate of the climate impact, is sensitive to the tropopause definition used. A new tropopause definition has been developed for use in this work, and a tropopause climatology based on this definition is presented.

Global and regional trends have been evaluated for the new tropopause definition using the ERA-Interim analysis. The global tropopause height is rising at a rate of 47 m decade^{-1}, with temperature falling by -0.18 K decade^{-1} (all 1989-2007). Trends of almost double the global values were identified in the subtropics, indicating a widening of the tropical region, found to be in the range 0.9° decade^{-1} to 2.2° decade^{-1}.

I would like to thank my supervisors, Keith Shine and Brian Hoskins, for their insight and patience over the last three years. Thanks also to my monitoring committee, Lesley Gray and Anthony Illingworth, for their advice.

The work in this thesis has been dependent on data from many different sources. I am grateful to the Volpe Centre for producing the AEDT on a resolution useful for this work, to ECMWF for allowing me to access ERA data, and to Paul Berrisford for archiving it a little closer to home. Thanks to Steven Baughcum for providing a copy of the NASA inventory, and to Christine Frömming and Michael Ponater for providing a copy of the TRADEOFF inventory. Additional thanks go to Christine for allowing me to use her data for comparison with my results. I am also grateful to the AERO2k and QUANTIFY projects for access to their inventories.

John Methven provided invaluable instructions on the use of his trajectory model. Amanda Maycock spent many hours working to get the Edwards and Slingo radiation code up and running, and was ever helpful when it crashed.

Thanks to Marc Stringer for all his help, advice, and perseverance when the calculation of the trajectories turned out to be too much for the department computers. Thanks to David Spence for helping us to keep things running on the Campus Grid, and thanks to the National Grid Service for allowing me to use their computers. I am also extremely grateful to Andy Heaps for his help with IDL, especially in those first few months.

I would also like to thank my friends for making my time in Reading an enjoyable one, and to my family for their interest and support. In particular, I am grateful to Luke for his incredible capacity to remember things about meteorology, and for his enthusiasm about my work. Thanks to those among you who were brave enough to read part of this thesis.

Finally, thank you to the U.S. Federal Aviation Administration for providing the funding that allowed this work to happen.

To Daddy, for all those Saturdays in the rain

Contents

Chapter 1

Introduction

Aviation currently accounts for almost 5% of total anthropogenic climate forcing, depending on the metric used (Lee et al., 2009). This percentage is expected to increase as a result of increasing passenger traffic, and decreases in the contributions from other sectors. Around 50% of the total forcing due to aviation is from the emissions of non-CO_2 gases and particles, and the formation of contrails, yet many of these components have low levels of scientific understanding (e.g. Sausen et al., 2005, Lee et al., 2009). The radiative forcing due to water vapour emissions in particular has a very large 90% confidence interval relative to its best estimate. A new estimate of the radiative forcing due to aviation water vapour emissions, calculated using high resolution Lagrangian trajectories, is presented in this thesis.

The climate impact of aviation has been of scientific interest for several decades, beginning with an interest of the potential impact of contrails in the late 1960s (Reinking, 1968; Murcray, 1970). In the 1970s interest in the potential impacts of a hypothetical fleet of supersonic aircraft began (e.g. Barrie et al., 1976), which later expanded into an interest in futuristic aircraft (e.g. Gauss et al., 2003). In recent years, there has been a series of comprehensive assessments of the climate impact of aviation (Brasseur et al., 1998; IPCC, 1999, Sausen et al., 2005; Lee et al., 2009, 2010), and an increased focus on providing information for policy and mitigation (e.g. Fuglestvedt et al., 2010, Forster

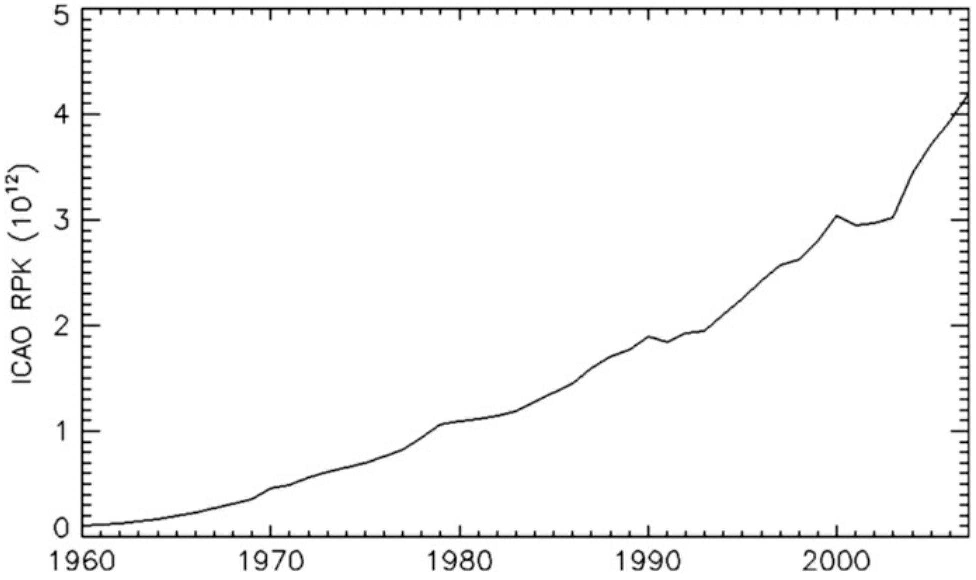


Figure 1.1: *Scheduled civil air traffic growth from 1960 to 2007 in billions of revenue passenger kilometres (RPK), from Lee et al. (2010).*

et al., 2006).

1.1 Aviation emissions

Aviation is a unique source of anthropogenic emissions. It is the only regular source of emissions above the atmospheric boundary layer. Additionally, the transport sector is the only major economic sector in Europe where greenhouse gas emissions are still increasing, and the largest contributor to this increase is from aviation (Lee et al., 2010). The aviation sector is continuing to grow, despite the increasing price of fuel and the threat of terrorism (see Figure 1.1).

Aviation can impact climate in four ways:

1. Direct emission of greenhouse gases, e.g. CO_2, H_2O
2. Indirect emission of greenhouse gases, e.g. the production of ozone and reduction of methane resulting from NO_x emissions

3. The emission of aerosols and aerosol precursors, which have a direct radiative impact
4. Changes to clouds as a result of emissions

Most of the components of an aircraft's exhaust plume are products of the combustion of kerosene, although there may be additional components from engine erosion and oil. The primary products of the combustion of kerosene are carbon dioxide (CO_2) and water vapour (H_2O). Trace species, including oxides of nitrogen (NO_x), sulphur dioxide (SO_2), unburnt hydrocarbons, and soot aerosols are also produced. Aircraft exhaust may also trigger the formation of contrails, if certain temperature and supersaturation conditions are met, which can go on to form cirrus clouds.

Although the formation of contrails can be initiated by water vapour emissions from aircraft, the contrail itself is formed by the condensation of water vapour already in the atmosphere. Only the direct radiative effects of water vapour emissions are considered in this thesis.

There has been a series of comprehensive assessments of the climate impact of aviation in recent years including an Intergovernmental Panel on Climate Change (IPCC) report (IPCC, 1999), and the assessments of Sausen et al., 2005 and Lee et al., 2009 and 2010.[1] The estimates from Lee et al. (2009) are shown in Figure 1.2. The best estimate of radiative forcing due to water vapour emissions is 0.0028 W m^{-2}, which is an order of magnitude smaller than the best estimate of radiative forcing due to CO_2 emissions. However, there is a low level of scientific understanding associated with the radiative forcing due to aviation water vapour emissions, and a large 90% confidence interval. If water vapour radiative forcing were to lie at the upper end of this range (0.02 Wm^{-2}) it could potentially make a larger contribution to total aviation radiative forcing than that from NO_x or linear contrails. Therefore, it is important to improve understanding of how water vapour emissions impact climate, and constrain the large uncertainty associated with the radiative forcing due to aviation water vapour emissions.

---

[1] All acronyms used are expanded in the list of acronyms given at the end of this thesis.

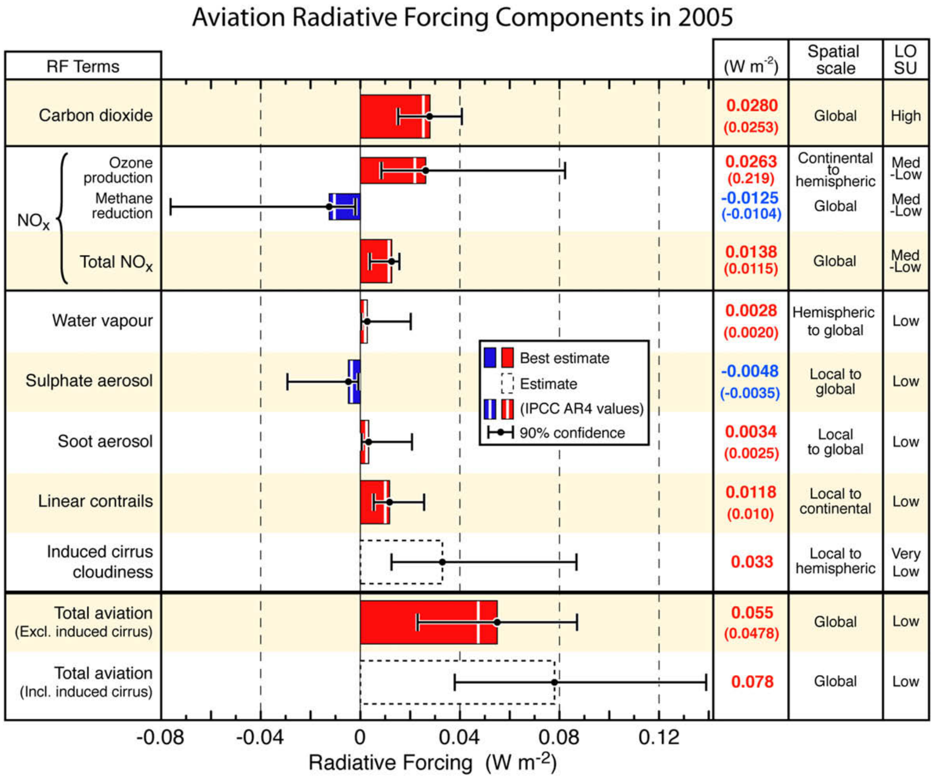


Figure 1.2: *Radiative forcing components from global aviation, from Lee et al. (2009). The coloured bars indicate the best estimate, and the white lines show the IPCC AR4 estimate. Numerical values of these estimates are given on the right, IPCC AR4 estimates are in brackets. The bars indicate the 90% confidence interval. 'LOSU' is the level of scientific understanding.*

1.2 Why water vapour emissions?

Aviation water vapour emissions have not recently been investigated in detail, but mitigation strategies need to be judged on their total effect. This is difficult to do if there are large uncertainties associated with certain emissions. Water vapour emissions have often been assumed to be negligible (e.g. Sausen et al., 2005). However, there is a large uncertainty associated with the best estimate of the radiative forcing, which suggests that there is potential for the climate impact to be much larger than the current best estimate (Figure 1.2). Hence, it is important to either revise the best estimate, or reduce the uncertainty about the best estimate of the climate impact of water vapour emissions from aircraft, despite the apparent smallness of that estimate.

Changes to aircraft design to mitigate their climate impact include a move towards the use of lighter materials in the construction of the aircraft, or a change to the shape of the aircraft, from something as small as the addition of a winglet to a complete redesign e.g. blended wing body aircraft. Changes can also be made to the fuel used. Research into the use of biofuels in aviation is currently ongoing (e.g. Beyersdorf and Anderson, 2010; Miake-Lye et al., 2010), and there has been some speculation about a change to liquid hydrogen fuelled aircraft (e.g. Gauss et al., 2003; Ponater et al., 2006). Such changes alter the composition of the emissions. Water vapour will be the main emission from a fleet of liquid hydrogen fuelled aircraft, which is estimated to have a smaller overall climate impact should they ever come into service (Ponater et al., 2006). Additionally, the altitude dependence of the climate impact of water vapour emissions made water vapour the dominant forcing component for a hypothetical supersonic fleet (e.g. IPCC, 1999). Therefore, it is important to understand how aviation water vapour emissions affect climate if the potential climate impact of these hypothetical fleets is reliably to be assessed.

The majority of studies focusing solely on water vapour emissions do so for a hypothetical fleet of either supersonic or liquid hydrogen aircraft. There are few estimates of the climate impact of water vapour emissions calculated using present-day global emissions from aviation, and none that use high resolution calculations. Both horizontal and vertical resolution have been shown to be important when calculating stratosphere to troposphere exchange (e.g. Kentarchos et al., 2000; Land et al., 2002). The representation of stratosphere to troposphere exchange is likely to be of importance when evaluating the climate impact of water vapour emissions. Only stratospheric emissions are likely significantly to perturb the background humidity, and hence have a radiative effect, because of their persistence and the low ambient humidity there.

1.3 The tropopause

The lifetime of water vapour emissions in the stratosphere, and their climate impact, has been shown to be dependent on their initial height above the tropopause (e.g. Gettelman, 1998; Grewe and Stenke, 2008). As the height of the tropopause can vary substantially with changes in tropopause definition, estimates of the climate impact of water vapour emissions are also expected to be sensitive to the tropopause definition. There is no universally correct definition of the tropopause, and several different definitions have been used in previous work. In this thesis, a new definition is developed for use in the calculations of the climate impact of aviation water vapour emissions, and a tropopause climatology based on this definition is presented.

In addition to being important to the evaluation of the climate impact of water vapour emissions, variables on the tropopause can provide a unique insight into atmospheric dynamics (Hoskins et al., 1985). In recent decades the height of the global tropopause has been observed to be increasing. It is thought that the primary cause of this increase is increases in well-mixed greenhouse gases (Santer et al., 2004; Santer et al., 2003a; Sausen and Santer, 2003). It has been suggested that tropopause trends provide a sensitive indicator of climate change (e.g. Randel et al., 2000) as the smaller natural variability at the tropopause compared to the surface means that changes there can be detected sooner (Sausen and Santer, 2003). Therefore, this thesis also assesses trends in tropopause height, using the new definition and recent atmospheric data sets.

1.4 Thesis outline

The overall aim of this thesis is to present a new estimate of the radiative forcing due to aviation water vapour emissions. This estimate is produced using recent emissions inventories, a fully Lagrangian trajectory model using winds and tropopause location given by reanalysis data, and a radiative transfer model.

The thesis is structured as follows:

- Chapter 2: An introduction to the emissions inventories and meteorological reanalyses used in this work
- Chapter 3: A discussion of different tropopause definitions and their respective merits, followed by the introduction of a new tropopause definition
- Chapter 4: Climatology based on the new tropopause definition
- Chapter 5: Trends in tropopause variables
- Chapter 6: A discussion of the direct deposition of emissions into the stratosphere, an introduction to the trajectory model, and preliminary estimates of the stratospheric lifetime of emissions
- Chapter 7: High resolution calculations of the accumulation of emissions in the stratosphere, and a revised estimate of the stratospheric lifetime of emissions
- Chapter 8: Calculations of the radiative forcing due to aviation water vapour emissions
- Chapter 9: Summary, conclusions, and suggestions for future research

Chapter 2

Datasets

Aircraft make frequent records of their position and speed throughout each flight. They also record external parameters such as air temperature and wind speed. Ideally, this data would be used in this thesis. Unfortunately, airlines' data are commercially sensitive and not readily available. Aviation emissions inventories are used in place of actual airline data, as they are more readily available and provide global coverage. These inventories are partly based on actual data, including radar tracking data and airline schedules, but they rely on a series of assumptions in regions where little data are available, and also to make the task more computationally simple.

In this Chapter an overview of the four main data sets used in this project; AERO2k, QUANTIFY, Aviation Environment Design Tool (AEDT), and the European Centre for Medium-range Weather Forecasts (ECMWF) re-analyses (ERA); is presented. AERO2k, QUANTIFY and AEDT are global aviation emissions inventories that provide information about aircraft movements, and the associated emissions, for a given time period. The ERA reanalyses used in this work are ERA-40 (1957-2002) and ERA-Interim (1989-2007). They provide spatial and temporal information about a range of meteorological parameters, including pressure, temperature, wind and potential vorticity.

2.1 The AERO2k emissions inventory

AERO2k (Eyers et al., 2004) is a global gridded aviation emissions inventory for the year 2002. It also includes a forecast of emissions for the year 2025. The inventory was developed under the European Commission Framework Programme 5 AERO2k project and covers both civil and military aviation. Only the inventory for 2002 civil aviation will be considered in this thesis.

For each month of 2002, monthly totals of fuel burn, distance flown, carbon dioxide (CO_2), water vapour (H_2O), oxides of nitrogen (NO_x), carbon monoxide (CO), hydrocarbon and particulate emissions are provided for each box in a three dimensional grid covering all longitudes and latitudes, and altitudes from 0 to 50,000 ft. The grid resolution is $1^\circ \times 1^\circ \times 500$ ft$\times 1$ month.

2.1.1 Making AERO2k

According to Eyers et al. (2004), the inventory was compiled using a combination of radar, flight plan and schedules data. Although the inventory is presented as monthly totals for the year 2002, it is based on only 6 weeks of actual data. Data were collected for the United States and Europe for the second weeks of February, April, June, September and October, and the first week of December. For the rest of the world, data were collected on only one day from each of these months. While the choice of week within each month was arbitrary, the months themselves were chosen to reflect the annual variation in air traffic (see Figure 2.1). February and April capture the spring increase in traffic, June and September represent the summer maximum, and October and December capture the winter decrease in traffic. Annual trends were then taken from Official Airline Guide (OAG) schedules, and applied to the collected data to give monthly totals.

Four-dimensional data was used for the US and Europe, compiled from Air Traffic Flow Management Modelling Capability (AMOC) flight plan data and Enhanced Traffic Man-

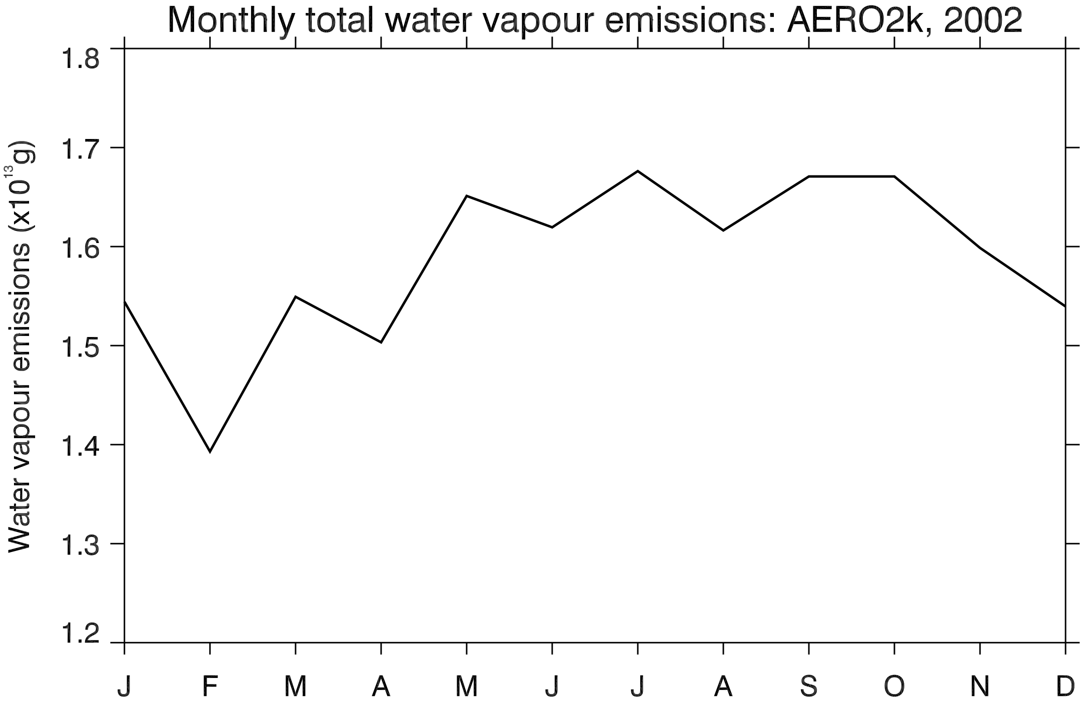


Figure 2.1: *Monthly global total water vapour emissions (proportional to fuel burn) for the year 2002 from the AERO2k inventory. The monthly variation is a result of scaling of the sampled data.*

agement System (ETMS) radar data.[1] The regions covered by AMOC and ETMS overlap, and where there was a difference between the two trajectories for a single aircraft, the trajectories were combined to give the longest possible final trajectory for the aircraft (see Figure 2.2). For the rest of the world, schedule data from the Back Aviation commercial database is used, which details only the times and locations of arrivals and departures for each aircraft. When schedule data is used, aircraft are assumed to follow great circle routes, i.e. they take the shortest possible route between two points. All of the movement data used to compile the AERO2k inventory is based on Instrument Flight Rules (IFR) rated aircraft. Flights by Visual Flight Rules (VFR) aircraft are neglected as estimates from Boeing for 1992 showed that they account for only 2% of all air traffic (Mortlock and Van Alstyne, 1998). [2]

There were approximately 300 different aircraft types, and their variants, in operation in 2002. In AERO2k these are modelled using 40 representative aircraft, which were chosen by grouping aircraft by seat capacity, engine technology, maximum take-off weight and configuration (Eyers et al., 2004). The numbers of aircraft within each seat capacity and technology category were then determined, with the most frequently-used aircraft being selected as the representative aircraft for that category, on the condition that the type was available within the Project Interactive ANalysis and Optimization (PIANO) aircraft design model. PIANO was then used to calculate fuel use per flight. In addition to representative aircraft, a standard flight speed was assumed for modelling purposes. All flights were assumed to travel at Long Range Cruise (LRC) speed. This is a commonly-used compromise speed, where aircraft cruise slightly faster than their most efficient speed, but are then only able to cover 99% of the maximum possible distance.

Fuel burn and emissions were calculated on a flight-by-flight basis (Eyers et al., 2004).

---

[1] ETMS is the US Federal Aviation Administration's (FAA) record of flight position and flight plan information used for air traffic management. ETMS covers North America, Western Europe and the North Atlantic. It captures every flight that files a flight plan, and every flight within the range of the FAA radars

[2] There are two different categories of flight rules, VFR and IFR, under which a pilot or aircraft can operate. These rules are determined by the weather conditions and the methods of navigation used by the pilot. Under IFR pilots rely on information displayed on instruments or equipment within the aircraft for navigation. The majority of IFR flights are required to file flight plans with Air Traffic Control before departure. Under VFR pilots navigate using visual information, following roads, rivers or other landmarks. VFR flights are not normally required to file flight plans.

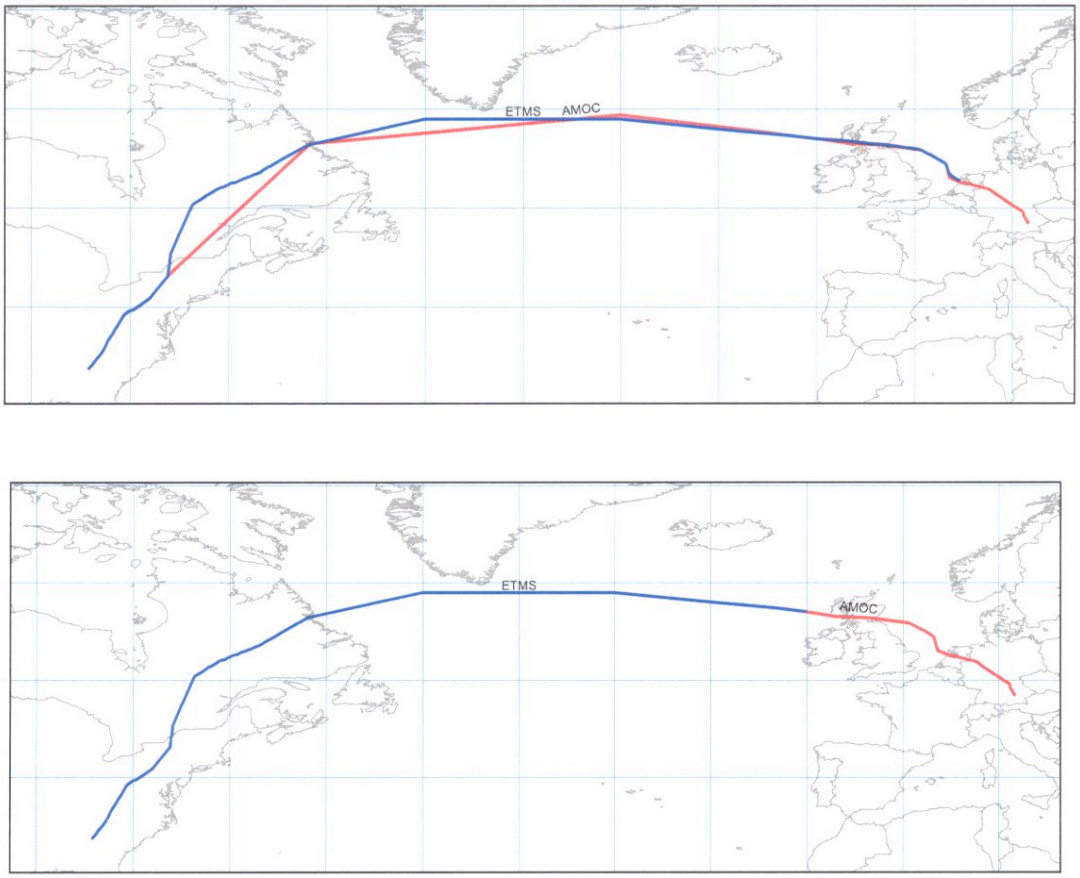


Figure 2.2: *Treatment of four dimensional flight data in AERO2k. The top panel shows the same flight identified using ETMS (blue) and AMOC (red). The bottom panel shows the flight path used in the inventory, after the two sources have been combined. (From Eyers et al., (2004))*

Fuel data was assigned to flight profiles by breaking each flight into 6 major sections:

- Take-off and climb-out: The take-off weight for each aircraft was assumed to be 60.9% of its maximum, following the average world-wide payload for scheduled airlines (ICAO, 2002). Emissions from taxi, take-off and climb-out up to 3000 ft, were taken from the Aircraft Engine Exhaust Emissions Databank (QINETIQ, 2004) and lookup tables. All emissions associated with these early stages of each flight were allocated to the latitude and longitude of the airport. All airports were assumed to be at mean sea level.
- Climb: The fuel used in the remainder of the climb (above 3000 ft) was found from climb data tables. Initial fuel was calculated from the profile, and the aircraft mass and distance flown was calculated at the top of the climb.
- Cruise: Fuel flow in the cruise portion of the flight was found from cruise data tables for the altitude, Mach number and aircraft mass. Distance flown and aircraft mass were continually re-evaluated throughout the cruise segment. Aircraft speed was kept fixed throughout the cruise portion of each flight, and altitude remained fixed between way-points. [3]
- Mid-cruise adjustments: Step climbs or mid-cruise descents were applied where necessary. Any adjustments to altitude were made at the first way-point after the change occurred.
- Descent: Fuel use for the descent to 3000 ft was calculated from descent lookup tables.
- Landing: Data for landing was taken from the Emissions Databank, as for take-off.

The main products of the combustion of hydrocarbon fuel are CO_2 and H_2O. The emission of these products is directly coupled to fuel mass flow. This allows an emissions index (a ratio of the emitted species to the total fuel burnt) to be used to calculate the

---

[3] Each flight that incorporated radar data was represented using 20-30 way-points. At each of these points the speed and 4-D position of the aircraft was found from the radar data.

mass of water vapour emitted. Following Rachner (1998), Eyers et al. (2004) applied an emissions index of 1.237 kg kg^{-1} for water vapour to all flights in the AERO2k inventory.

2.1.2 Evaluation of AERO2k

Where schedules data is used in AERO2k, cancelled flights are included and unscheduled flights are not. Eyers et al. (2004) concluded that it would be difficult to estimate the resulting uncertainty as no information was available on the total number of world flights. Reliance on schedules data means that little information is known about the actual route taken by aircraft. In AERO2k it was assumed that aircraft followed great circle routes between way-points. This results in areas of little detail when there are large distances between way-points. It also means that rapid changes in direction and altitude are not captured. Eyers et al. (2004) note that as such rapid changes normally occur in the climb and descent segments of a flight, this assumption causes a larger error in short flights. Radar tracked flights account for 70% of all flights in the AERO2k inventory. According to Eyers et al., each of these flights is modelled using 20–30 way-points. This is many more points than are used for flights based on schedule data. However, as it is still assumed that aircraft follow great circle routes between these points there is still an underestimate in fuel burn when data points are sparse. Eyers et al. compared AERO2k to operational flight data and found that the use of great circle routing results in a 4.5% underestimate in mean true distance travelled. Additionally, they found that the omission of non-IFR flights lead to an underestimate in fuel burn of around 2% of total civil aviation fuel burn. The emissions totals presented in AERO2k have not been adjusted to take this into account as most non-IFR flights are completed by light aircraft, which would be difficult to model, and have very different spatial distribution to commercial flights.

Eyers et al. considered the effect of minor manoeuvres in flight to be small over the total distance of a flight. They also assumed that the increased fuel burn from making step climbs tended to be offset by the reduced fuel consumption at the higher altitude.

The resolution of the movement data used in AERO2k meant that it was not possible for Eyers et al. to identify the exact point where step climbs were made. Step climbs were therefore modelled at the first way-point after the actual climb was made. Eyers et al. state that 'any error introduced by not including a specific climb fuel allowance will be smaller than the error from not changing the cruise flight altitude at the appropriate point in the cruise.' They found that both errors from these manoeuvres were considerably smaller than the errors that would have resulted from using an assumed cruise altitude.

The representative engines used in AERO2k were chosen on the basis of their NO_x emissions. Eyers et al. state that this methodology assumes that competitive engine types used on the same aircraft will have similar fuel consumption (within 1-2%). They also note that uncertainties over operational practice, actual idle settings in particular, are likely to be greater than any averaging effect resulting from the choice of representative engines.

Eyers et al. were restricted in their choice of representative aircraft by the types of aircraft contained in PIANO, the performance tool used in the AERO2k project. As PIANO contains few models of Russian built aircraft, these aircraft were represented by similar western aircraft. Eyers et al. state that 'validation using one of the few aircraft to have a model in PIANO showed a fuel burn agreement within 10%'. As Russian built aircraft perform only $\sim$1% of global aircraft movements, they concluded that this approximation is unlikely to have a large impact on the inventory.

PIANO was found by Eyers et al. to produce results that compared well with Eurocontrol's Base of Aircraft DAta (BADA) and airline data. Eyers et al. found that trip fuel burn calculated using PIANO was within 3% on average of model data from both British Airways and Virgin. PIANO fuel burn was also no more than 5% higher than fuel burn from BADA on average. Eyers et al. calculated that the mean of PIANO calculated fuel burn was equal to 97.777% of that from operational data, with an 'especially good match at cruise altitude'. Operational flight data was recorded by on-board flight data recorded at intervals of approximately 4 seconds throughout each flight. The data included records of several parameters including altitude, aircraft weight and flight speed.

According to Eyers et al., the discrepancy between results from PIANO and operational data can be attributed to factors not accounted for by PIANO, such as manoeuvres in flights and engine and airframe deterioration; the use of different aircraft variants and weights; and sensor error in the operational flight data, and modelling inaccuracies in the other prediction programs.

Eyers et al. found that the assumption of a 60.9% payload used in AERO2k gave the best fit with airline data. However, Lee et al. (2009) quote the global average load factor as between 68% (1989) and 76% (in 2006).

For a given flight speed, wind alters an aircraft's speed relative to the ground. Flying into a headwind means that the aircraft effectively flies slower relative to the ground, taking longer to reach its destination and burning more fuel. Tailwinds have the opposite effect, while side winds will push an aircraft off course resulting in greater fuel burn through compensation. When radar track data is used in AERO2k these effects are intrinsically incorporated. However, for the rest of the world still air was assumed. From analysis of operational flight data, Eyers et al. found that 'winds at cruise altitude can affect the equivalent still air distance by as much as ±20%.' This impact will be particularly pronounced near to jet stream altitudes, which is anticipated to be a region of interest in this work. However, Eyers et al. found that the mean difference between AERO2k distance and operational distance was -1.5%.

In addition to the assumption of still air, International Standard Atmosphere (ISA) standard day temperatures were used in the compilation of AERO2k. The ambient air temperature has an impact on aircraft performance, and therefore has an impact on fuel burn. Eyers et al. observed considerable deviation from ISA temperatures in parts of the atmosphere, with some aircraft flying in regions with temperatures of ISA±20°C. However, they found that most flights lay within regions where temperatures were equal to ISA±5°C. They concluded that this meant that the difference between actual fuel burn and fuel burn derived using the ISA assumption would be small. However, the difference is likely to be largest near the jet streams, and consequently the altitudes expected to be of interest in this thesis.

There is limited coverage of holding in the AERO2k inventory.[4] According to Eyers et al., even though radar data includes holding, the low time resolution of the data relative to the holding time makes modelling impossible. Data obtained from British Airways by Eyers et al. from 3 major UK airports for the year 2000-2001 showed that holding resulted in a 1.2% increase in fuel burn. However, Eyers et al. highlight that this data included London Heathrow, which is noted for its long holding times.

Airframe and engine components deteriorate with use, causing the aircraft to be less efficient and increasing its fuel consumption. The EC 4^{th} framework project Aircraft Environmental Impacts and Certification Criteria (AEROCERT) (Norman, 2001) found that engine deterioration resulted in a 3-4% increase in typical maximum fuel consumption. Mykoniatis and Martin (1998) found an increase of fuel burn between 2 and 6% due to deterioration. However, aircraft are rarely flown with several deteriorated engines, and the overall impact is likely to be less than the above figures suggest. Eyers et al. assumed that the effects of deterioration were adequately represented in AERO2k as the increased fuel consumption is implicitly included in both data from airlines and published performance data.

Tankering, the practice of carrying more fuel than is required for a flight so that it can be used for subsequent flights, is not represented in AERO2k as Eyers et al. thought it difficult to quantify. The IPCC (1999) report suggested that tankering may result in an additional fuel burn of $\sim$0.5% globally, which is small compared to the errors arising from some of the other assumptions behind AERO2k.

2.2 The QUANTIFY emissions inventory

QUANTIFY (Lee et al., 2005) is a European Union Integrated Project based on funding from the European Commission under Framework Programme 6. The main aim of the QUANTIFY project is to determine the climate impact of global transport systems

---

[4] Also known as stacking, holding is a feature of congested airspace. If there are no landing slots available for an aircraft it is directed into a ‘holding pattern’ by air traffic control, where it will fly circuits at a constant altitude away from the main descent track until its allocated landing time.

by individually assessing the impact of various modes of transport, including shipping, aviation and land surface transport. The QUANTIFY aviation inventory covers civil aviation for the year 2000.

For each month of the year 2000, monthly totals of distance travelled, fuel burn, H_2O, CO_2, NO_x, and black carbon, are provided for each grid box in a three dimensional grid covering all latitudes and longitudes, and heights from 0 to 47,000 ft. The grid resolution is $1^\circ \times 1^\circ \times 610$ m$\times 1$ month.

2.2.1 Making QUANTIFY

All movements data used in the compilation of QUANTIFY were taken from OAG global schedules, and all flights were assumed to follow great circle routes. Following AERO2k, only flights by IFR rated aircraft were considered.

The Future Aviation Scenario Tool (FAST), which was designed around the methodology used in the development of the Abatement of Nuisances Caused by Air Transport ANCAT/EC1 and ANCAT/EC2 inventories (see IPCC (1999) for further information on ANCAT inventories), is based on a data set of aircraft movements indicating the frequency of flights of aircraft of certain types between city pairs for a given year. According to Lee et al. (2005), in the compilation of the QUANTIFY aviation emissions inventory, aircraft types were grouped from this database and assigned representative aircraft types based on those used in AERO2k. A total of 43 representative aircraft types were used in QUANTIFY.

The PIANO tool was used to provide fuel flow information for each aircraft-route combination, with the assumption that all flights were carrying 70% of their maximum payload. Fuel consumption for each flight was calculated between the arrival and departure locations, assuming that the locations were linked by the great circle distance and using a modelled cruise altitude. The modelled cruise altitude was based on an analysis of preliminary AERO2k data for 16 representative aircraft types. This analysis gave an average cruise altitude for each aircraft type for each route, and also a maximum cruise altitude.

An emissions index of 1.237 kg kg^{-1} was used to calculate water vapour emissions from fuel burn.

2.2.2 Evaluation of QUANTIFY

The exclusive use of schedules data in QUANTIFY was found by Lee et al. (2005) to result in an underestimate of fuel burn of approximately 9% compared to that calculated from full movements data. They note that this may even be a conservative estimate, with the difference being as large as -50% for UK and EU traffic. By comparing Eurocontrol data using actual routes, and Eurocontrol data with great circle routes substituted in, they also found that the assumption that all flights follow great circle routes causes an underestimate in fuel burn of 12-13% in the EU. A similar comparison outside the Eurocontrol traffic domain was difficult as most of the original Eurocontrol data in this region is itself based on great circle routes.

Lee et al. found that the representative aircraft used in QUANTIFY cause proportionally greater levels of uncertainty with increasing distance travelled. They found that the rate of increase in uncertainty was different for the different representative types. A 5% difference in fuel consumption between the A330-200 and the A330-300 was found after 4800 nm. For the B737-300, B737-400 and B737-500 the difference was 10% after only 1600 nm. Uncertainties due to the use of representative engine types were not investigated by Lee et al.

A 70% load factor is assumed in QUANTIFY. Lee et al. calculated that a payload variation of 10% about this value would change fuel consumption by $\sim$2.5%. It is possible that tankering could cause a payload variation outside this range. However, Lee et al. found the significance of tankering to be ‘difficult if not impossible to quantify on the scale necessary for this study. The IPCC (1999) report suggested that, based upon information from one airline, additional fuel burn as a result of tankering may be of the order of 0.5% of total aircraft fuel consumption.’ Lee et al. report that information from an unspecified European airline showed that 13.6% of short haul and 5.3% of long haul flights tanker fuel. When tankering occurred on short haul flights, the tankered fuel could be as much as 100% of the fuel necessary for completion of the flight. For long haul flights the amount of tankered fuel was less, around 50% of the required fuel on average. These figures come with the caveat that the extent of tankering is dependent on several factors, including airline practices and costs, absolute market fuel prices, regional market dislocations and local distribution costs.

2.3 The AEDT emissions inventory

AEDT is a computer model used to 'predict' aircraft fuel burn and emissions for all civil aviation in a given year. AEDT was developed by the US Federal Aviation Administration (FAA) with support from the Volpe Centre of the US Department of Transportation, the Massachusetts Institute of Technology and the Logistics Management Institute. It has been used to generate global inventories for the years 2000 to 2006, with work ongoing to extend this range to the present day.

Data from AEDT have been provided for use in this thesis by the FAA/Volpe Centre for February 2004, and for the year 2006.[5] Total fuel burn, distance flown, H_2O, NO_x and particulate emissions are given for each grid box in a three dimensional grid covering all latitudes, longitudes and altitudes. For use in this work, the data for February 2004 was provided at two different resolutions: $1^\circ \times 1^\circ \times 1$ km$\times$6 hr and $1^\circ \times 1^\circ \times 152$ m$\times$6 hr. 2006 data was provided with a resolution of $1^\circ \times 1^\circ \times 152$ m$\times$4 hr, and was interpolated onto a 6 hr grid to allow it to be used with 6 hourly ERA data.

The vertical levels on which AEDT is archived are pressure-height levels, which are converted from pressure levels using the assumption of a standard atmosphere. AEDT has an altitude cap of $\sim$13.5 km. Some flights are modelled above this altitude, but they are assumed to be erroneous (pers. comm. Gregg Fleming).

---

[5] Originally data was expected for the years 2000 to 2005. Non-delivery of this data has resulted in some inconsistencies between the dates of the meteorology and the emissions used in some of the calculations in this thesis.

2.3.1 Making AEDT

The AEDT inventory is compiled using a combination of radar data from ETMS and schedule data from the OAG. 4D data is taken from ETMS for the US and Western Europe, while schedule data is used for the rest of the world. Data from ETMS is filtered so that AEDT only includes scheduled (IFR) flights. In contrast to AERO2k and QUANTIFY, great circle routes are not assumed when schedule data is used. Instead, a trajectory generation method is used to create a more realistic spread of trajectories about a great circle route. The trajectory generation method is based on cruise altitudes and horizontal tracks seen in ETMS flights from May and October of 2000 and 2003 (Kim et al., 2005).

Unlike AERO2k and QUANTIFY, AEDT does not rely on representative aircraft for fuel use calculations. In-cruise aircraft performance and operation parameters are taken from BADA, which can be expanded to cover 295 types of aircraft using matching tables. Performance data for landing and take-off is taken from the Integrated Noise Model (INM). The use of a single cruise speed for every flight is also avoided by taking aircraft speeds from the BADA airline procedure models, which detail typical speeds for each altitude band.

Standard assumptions about aircraft weight are avoided in AEDT. According to Kim et al. (2005), an aircraft's take-off weight is a function of the distance of the route it is about to travel. In AEDT, each modelled trip is split into stages, with each stage corresponding to a given distance. The number of stages that an aircraft will traverse in a flight is then used to assign an equivalent take-off weight. However, to account for fuel tankering, all take-off weights in AEDT were increased by the equivalent of two stage lengths. The example of a Boeing 747-200 is given by Kim et al. (2005). An aircraft flying a 5 stage trip would have a take-off weight of 302,000 kg, the same aircraft flying a 7 stage trip would have a take-off weight of 352,000 kg. The maximum take-off weight of a B747-200 is 378,000 kg, so B747-200 flying a 5 stage trip would have its take-off weight scaled from 80% to 93% of its maximum weight to account for tankering.

Taxi times and delays are also modelled in AEDT, giving the inventory gate to gate coverage. Unscheduled and cancelled flights are accounted for through the use of scaling factors, which, according to Kim et al., (2005), ‘generally have the effect of increasing the number of flights’. Like AERO2k, AEDT is compiled using single flight modelling, with each flight modelled as a series of finite sections: departure gate, departure taxi, departure runway roll, takeoff and climb-out (up to 3000 ft), cruise, approach (below 3000 ft), arrival, runway roll, arrival taxi, arrival gate. Fuel burn is modelled for the duration of each flight to avoid dependence on lookup tables.

2.3.2 Evaluation of AEDT

Malwitz et al. (2004) present details of the AEDT assessment, which was performed in two stages: modular evaluation to establish the reliability of key subroutines in the AEDT model, and a full system evaluation to assess the trajectory modelling components of the model as a whole.

Malwitz et al. (2004) give detailed information on the evaluation of three modules: the calculation of landing and take off cycles (LTO) cycles, the aerodynamic components of AEDT, and the fuel flow component. Malwitz et al. state that Integrated Noise Model (INM) data and methods were implemented for LTO modelling as these have been extensively tested, and are internationally accepted. They found that comparison of the horizontal distance travelled in take-off in INM and AEDT agree to within 10-20%, with no bias in either direction. They attribute this uncertainty to the the fact that there is no standard procedure followed by all airlines.

Malwitz et al. found that the general shape of Lift-Drag ratio curves produced by AEDT compare well with published data, and that the AEDT generated values were within $\pm14\%$ of NASA data. This difference was attributed to the use of BADA data in the calculation of the profiles. BADA data has a fixed set of coefficients and is therefore insensitive to factors influencing the ratio, including flight speed and ambient temperature changes.

The fuel flow component of AEDT was evaluated by Malwitz et al. via a comparison with high resolution computer flight recorder data provided by a major US carrier for ten flights, two for each of five different types of aircraft, and with data from the NASA test aircraft (a B757-200). They found that the mean percentage difference in fuel flow between AEDT and the carrier data was 6.95%, with a standard deviation of 36.7%. The mean percentage difference between AEDT and data from the NASA test aircraft was only -0.24%, with a standard deviation of 37.3%. Further comparisons between AEDT fuel consumption and that from Jane's Aero engines showed that for most engines the difference in specific fuel consumption from the two sources was within 10%. However, for some engines Malwitz et al. found that the difference was in excess of 20%.

Malwitz et al. performed a full system evaluation using data from a US carrier, and data from two Asian carriers. This data included aggregate flight-level fuel burn for each of the airlines' flights for October 2000 and 2003 respectively. For October 2000, the data from the US carrier contained 44,000 flights that appeared in both ETMS and the OAG. Malwitz et al. used AEDT to calculate fuel burn for these flight using only ETMS data, and then only OAG data, rather than the combination normally used in the model. They found that the percentage difference between the AEDT fuel burn and actual fuel burn values was -0.25%, with a standard deviation of 13.04% for the ETMS calculations, and -11.32% with a standard deviation of 10.66% for the OAG calculations. When both ETMS and OAG data was used in the calculations, the percentage difference between AEDT and actual fuel burn values was -2.62% with a standard deviation of 13.69%, as AEDT gives preference to ETMS data over OAG data. Malwitz et al. therefore concluded that AEDT 'can predict fleet fuel burn with less than a 5% error on average'. Comparison of AEDT fuel burn (calculated using ETMS combined with OAG data) with actual fuel burn from the two Asian carriers gave a difference of 0.42% with a standard deviation of 21.50%.

The fuel burn errors show that the biggest uncertainty in AEDT fuel burn comes from flights where only OAG data is available. Malwitz et al. state that this may be due to the 'created flight tracks and speed data' used in the OAG. Malwitz et al. also found that errors in fuel burn were greatest for short flights. They believe that this is due to a greater degree of aircraft substitutions occurring on short flights, and also to the decreased accuracy of the trajectory modelling for short flights.

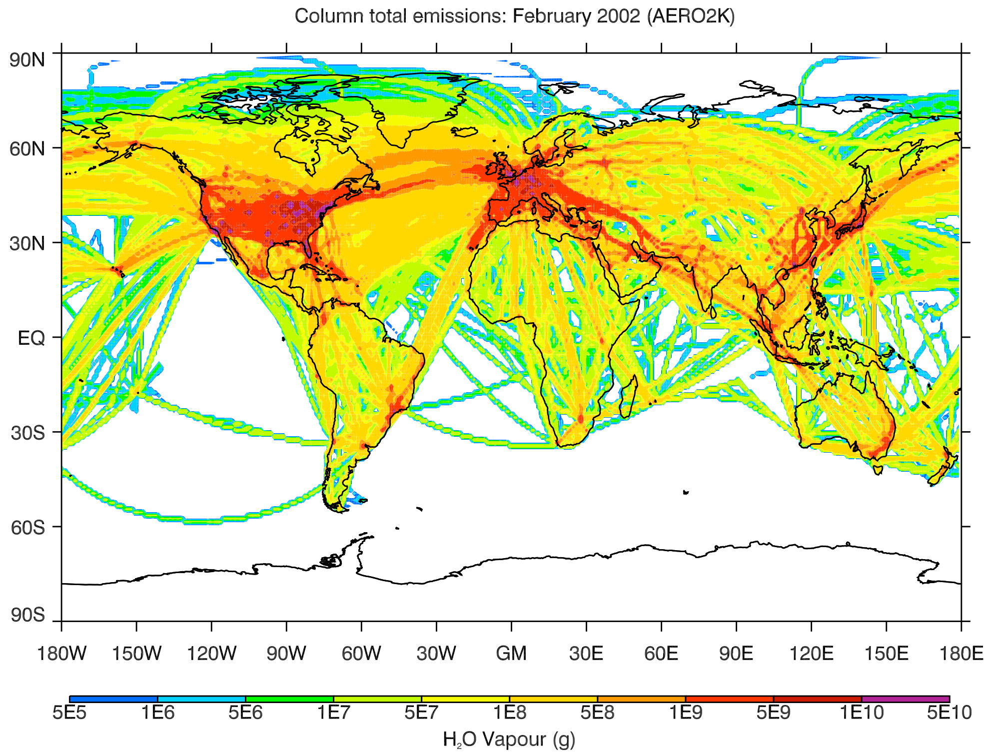
Malwitz et al. concluded that some of the error in fuel burn calculations was attributable to the lack of weather in AEDT. Full meteorological data was included in AEDT for one day, and the calculated fuel burn was compared to actual fuel burn provided by the US carrier. Malwitz et al. found that the mean percentage difference between AEDT fuel burn and actual fuel burn was only 2.70% with a standard deviation of 12.59% when weather was included in AEDT, compared to -7.87% with a standard deviation of 13.75% when it was not.

2.4 Inventory comparisons

The differences in the methodologies described in the previous Sections contribute to discrepancies between the three emissions inventories. Many of the differences found between AERO2k, QUANTIFY and AEDT are large, and therefore have the potential to have a significant influence on any results that depend on the use of an emissions inventory. Some of the more striking differences, and suggestions for their origin, are now presented.

The horizontal distributions of water vapour emissions given by the three inventories are shown in Figure 2.3 (AERO2k and QUANTIFY) and Figure 2.4 (AEDT). For each inventory the column total emissions are shown for February. It is important to note that each inventory represents data from different years (2000 for QUANTIFY, 2002 for AERO2k, and 2004 for AEDT), and that any comparisons referred to in this Section are not direct as a result.

The horizontal distribution of emissions in both QUANTIFY and AERO2k is largely determined by the use of great circle routes (Figure 2.3). Every route in QUANTIFY is assumed to be a great circle, along with a large number of routes in AERO2k. The result of the combination of ETMS and AMOC trajectories over North America and Europe,


Column total emissions: February 2002 (AERO2K)

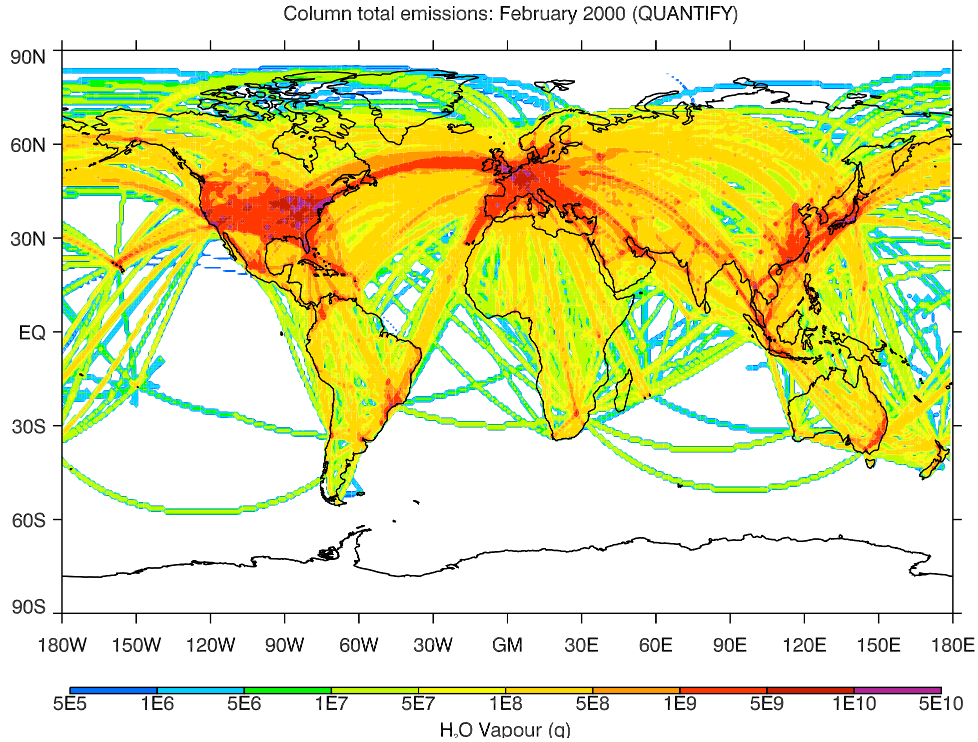

Column total emissions: February 2000 (QUANTIFY)

Figure 2.3: *Column total water vapour emissions for the month of February. The top panel shows the emissions from the AERO2k inventory, for 2002; the bottom panel shows the emissions from the QUANTIFY inventory, for 2000.*

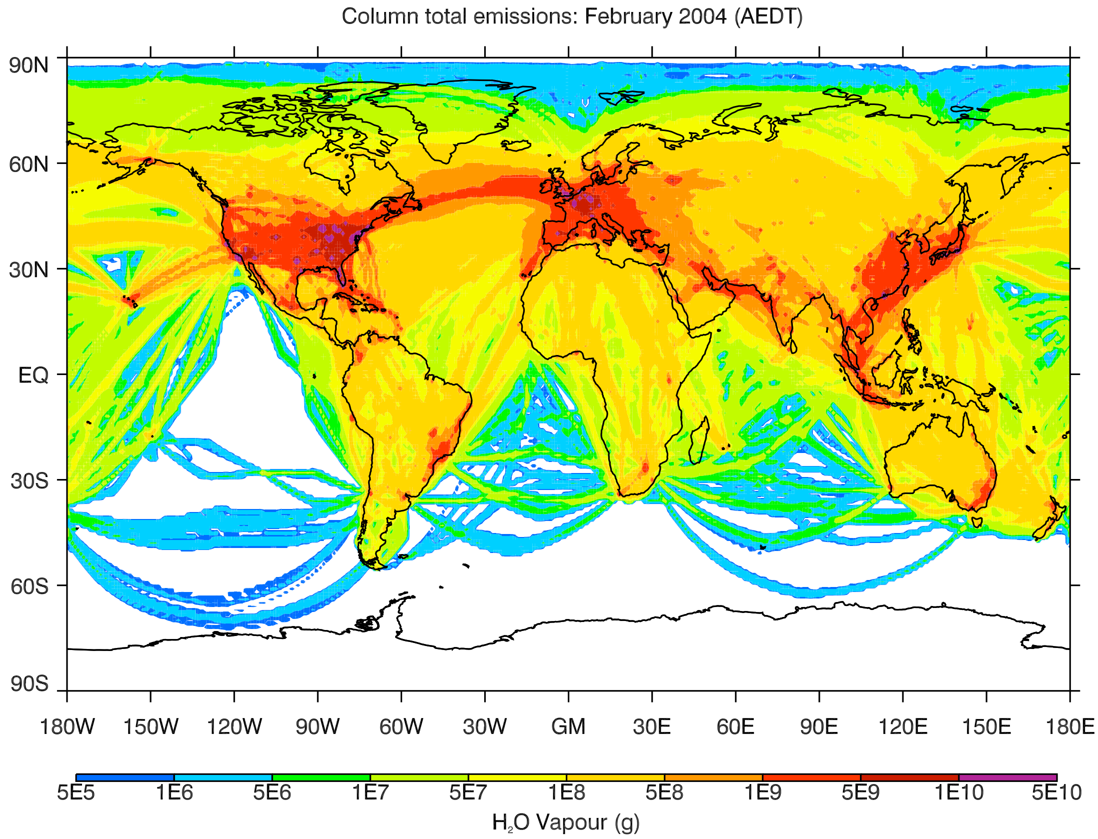


Figure 2.4: *Column total water vapour emissions for February 2004 using emissions from the AEDT inventory.*

and the use of great circles between way-points compared to their use for whole trips, can clearly be seen in Figure 2.3. While many of the routes in the Southern Hemisphere appear to be very similar in AERO2k and QUANTIFY, there are significant differences between the distribution of emissions from AERO2k and those from QUANTIFY in the Northern Hemisphere. All tracks from QUANTIFY have a distinctive curvature associated with great circle routing, but many of the Northern Hemisphere tracks from AERO2k appear as straight lines on the map projection, those traversing the North Atlantic, and over Eastern Europe, in particular.

Comparing the horizontal distribution of emissions from AEDT (Figure 2.4) to those from AERO2k and QUANTIFY (Figure 2.3) reveals an even more striking contrast than that illustrated within Figure 2.3. The horizontal spread in emissions forced by the use of a probability distribution about great circle routes in AEDT is clearly visible in the Southern Hemisphere, in strong contrast to the great circle routes used in AERO2k and QUANTIFY. However, there is also greater horizontal spread in the Northern Hemisphere emissions from AEDT compared to those from AERO2k and QUANTIFY. The water vapour emissions also extend further north in the AEDT inventory. This is likely to be due, in part, to the use of radar data instead of flight plan data.

Comparison of Figures 2.3 and 2.4 reveals a difference in the magnitude of emissions between the three inventories, with column total emissions being visibly higher in the AEDT inventory than either AERO2k or QUANTIFY. Global total water vapour emissions for February in QUANTIFY (2000), AERO2k (2002) and AEDT (2004) are 1.42×10^{13}g, 1.40×10^{13}g, and 1.72×10^{13}g respectively. There are 23% more water vapour emissions in AEDT for February 2004 than there are in AERO2k for February 2002, and 21% more water vapour emissions in AEDT for February 2004 than there are in QUANTIFY for February 2000. Lee et al. (2009) found that there was an average increase of 5.3% per annum in passenger traffic between 2000 and 2007. Taking this into account, there is still a difference of $\sim 13\%$ between the total water vapour emissions in AEDT and AERO2k, but only a $\sim 1\%$ difference between AEDT and QUANTIFY.

There is a number of possible contributing factors to the large difference in total water vapour emissions between AEDT and AERO2k. The differences may be accounted for by:

1. The use of the probability distribution in AEDT, which means that aircraft fly further than they would on the same route in AERO2k
2. The scaling factor used in AEDT to account for unscheduled flights is known to result in a greater number of total flights
3. The issues of representative vs. real aircraft, scheduled vs. actual flight trajectories and the use of different aircraft performance information.

The scaling used to account for tankering in AEDT may also contribute. However, it is difficult to determine the difference that this will make in the average compared with AERO2k and QUANTIFY. Although the scaling means that all AEDT aircraft are assumed to be carrying surplus fuel, there is no general assumption made about aircraft weight in AEDT, as there is in AERO2k and QUANTIFY. This means that some aircraft in AEDT may be much lighter than the 70% of maximum weight assumed in AERO2k and QUANTIFY, and that the average aircraft weight in AEDT may lie close to 70% of the maximum.

While the absence of a like-for-like comparison, and the use of only one month of data, make a quantitative comparison impossible here, the relatively large values of total emissions found in the AEDT inventory have also been identified in other studies. Kim et al. (2005a) found that values of AEDT fuel burn, which is proportional to water vapour emissions, were 30% greater than could be expected from the trends exhibited in previous studies. Studies by the International Civil Aviation Organization Committee on Aviation Environmental Protection (ICAO/CAEP) (CAEP, 2007) have also identified large differences in global total fuel burn between emissions inventories, with investigations into the underlying causes still ongoing.

When comparing the magnitude of the emissions in the three inventories it is important to note that all of the assumptions behind them, with the exception of scaling to account for unscheduled flights, and modelling of tankering in AEDT cause a decrease in inventory emissions with respect to actual emissions. The anomalously high values of fuel burn and emissions seen in AEDT when compared to other inventories may therefore be a reflection of the fact that it relies on fewer assumptions, and of its improved agreement with airline data compared to other inventories.

The relatively long lifetime of water vapour emissions in the stratosphere has been shown in previous studies (Gettelman (1998), for example) to be sensitive to the height of the emissions above the tropopause. This sensitivity means that differences in the vertical distribution of water vapour emissions between inventories have the potential to produce a large influence on results. The vertical distribution of water vapour emissions in AERO2k, QUANTIFY and AEDT for February are compared here for several regions. The vertical distribution of global emissions, US emissions, and North Atlantic emissions can be seen in Figure 2.5.

Vertical profiles of water vapour emissions from all three inventories have two main maxima: one at cruise altitude (typically 9-12 km), and another at ground level. In both the global distribution and the US distribution (Figure 2.5) it can be seen that the gate to gate modelling used in AEDT results in a smaller percentage of the column total emissions of water vapour being emitted at, or near to, ground level than the parameterization used in AERO2k and QUANTIFY, which is known to be an overestimate (Eyers et al., 2004).

The largest maxima in the profiles from all three inventories is at cruise altitude, with up to 60% of the column total emissions being emitted at this level. For all the regions shown in Figure 2.5 the effect of the combination of great circle routing and specified cruise altitudes is reflected in the relatively narrow peaks, and greater percentage of total emissions, at cruise altitude in the QUANTIFY profiles. The cruise altitude maxima in the AERO2k and QUANTIFY profiles are found at the same altitude. This is likely to be due to the fact that the cruise altitudes modelled in QUANTIFY were calculated from a sample of AERO2k profiles.

In both the global and North Atlantic profiles (Figure 2.5), the cruise altitude in the AEDT inventory is 1 km higher than that in the AERO2k and QUANTIFY profiles. This is likely to be due primarily to the use of ETMS data in AEDT instead of the OAG data used in AERO2k and QUANTIFY. The combination of ETMS and OAG data in the

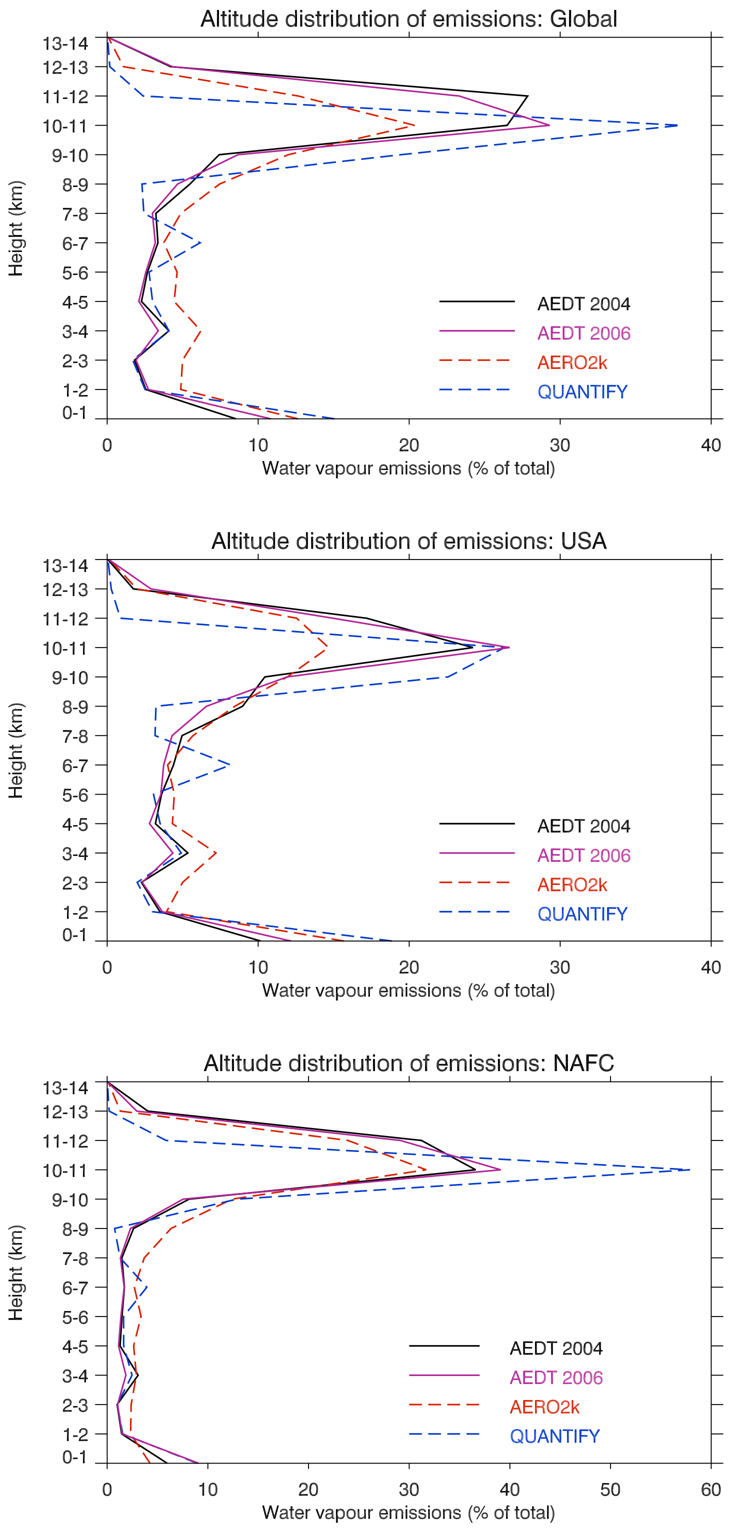


Figure 2.5: *Altitude distribution of monthly total water vapour emissions for February from AEDT, AERO2k and QUANTIFY for the years 2004, 2002 and 2000, respectively. The global emissions distribution is shown in the top panel, he vertical distribution of emissions over the continental USA is shown in the middle panel, and the distribution in the North Atlantic Flight Corridor (NAFC) is shown in the bottom panel. The NAFC has been defined in this thesis as the region bounded by 70°W, 0°W, 40°N, and 65°N.*

AERO2k inventory over the US and Europe means that AERO2k and AEDT use broadly the same input data over the US, so it is reasonable to expect the profiles to look similar. However, although the cruise altitude is the same in the two inventories, water vapour emissions are much more evenly distributed with altitude in AERO2k than they are in AEDT. This large spread in emissions about cruise altitude in AERO2k can be seen in all the profiles, the US profile (Figure 2.5) in particular. While the AEDT and QUANTIFY inventories both produce a sharp maximum at cruise altitude in all the profiles shown, emissions in AERO2k are spread over a large range of altitudes, resulting in a much smaller percentage of column total emissions being located at cruise altitude: the peak in the global profile (Figure 2.5) could be interpreted as extending from 7-13 km, compared to 10-13 km, and 9-12 km in the AEDT and QUANTIFY inventories respectively. The cause of the large vertical spread of emissions within AERO2k is unknown. However, it is thought that it is unlikely to be a true representation of the vertical distribution of actual emissions.

When discrepancies arise between the three inventories, AEDT will be regarded as the most reliable representation of actual aircraft movements and emissions. QUANTIFY will be regarded as the most simplistic. Much more radar data, and schedule data, is assimilated into the AEDT inventory than either AERO2k or QUANTIFY, removing the need for any scaling of the recorded data. Above all, of the three inventories AEDT fuel burn comes closest to matching records of operational fuel burn. However, all three are an underestimate relative to fuel sales. In some studies, e.g. Lee et al., (2009), this underestimate is accounted for by scaling the inventory emissions to IEA totals. A summary of the methods used in the construction of the three inventories is given in Table 2.1. The reported errors are shown in Table 2.2.

In AERO2k, radar data is combined with schedule data in the US and Europe to give the longest possible trajectory. This combination was performed for every flight. Malwitz et al. (2004) have since demonstrated that schedule data is a much larger source of error than radar data for this region. As a result, when radar data and schedule data are combined in AEDT, preference is given the radar data over schedule data. This gives a better comparison with operational data than that which can be achieved by using either radar data or schedule data alone. The method for combining radar and schedule data over the US and Europe used in AERO2k, where each is given equal weight, is therefore anticipated to be more prone to cause deviations from actual aircraft movements than the methodology used in AEDT. The great circle routes used in the QUANTIFY inventory for this region are known to cause large underestimates in the distance travelled, as is the exclusive use of schedules data (Lee et al., 2005).

All three inventories rely on schedules data outside the US and Europe. Although scaling is applied to take account of this in AEDT, there is no explicit treatment of unscheduled or cancelled flights in these regions in any of the inventories. Only six days of schedules data was used in the AERO2k inventory. The more frequent sampling rate used in AEDT means that more details of flights are known. When schedule data is used in the AERO2k and QUANTIFY inventories, aircraft are assumed to follow great circle routes at idealised altitudes. The assumption of great circle routes is also made in AEDT. However, the use of a probability distribution gives a more realistic spread of trajectories, both horizontally and vertically.

A more realistic representation of speed is used in AEDT than in AERO2k or QUANTIFY. In AERO2k it is assumed that an aircraft flies at a constant speed whilst cruising. In AEDT different cruise speeds are used for different altitudes, based on typical speeds from the BADA procedure models. As speed is related to fuel burn, and fuel burn is proportional to water vapour emissions, this is potentially important to the results of this thesis.

Component	AERO2k	QUANTIFY	AEDT
Radar	ETMS US/EU, but combined with AMOC flight plans	None	ETMS over US/EU
Schedules	OAG	OAG	OAG
Trajectories	US/EU: Combine ETMS with AMOC to give longest trajectory based on way-points. Great circle routes assumed between way-points. Rest: Great circle routes assumed.	Great circle routes. Modelled cruise based on analysis of 16 AERO2k representative aircraft to give mean and max cruise	US/EU from ETMS. Rest: Trajectory generation based on altitudes and horizontal tracks from ETMS flights for May and October 2000 and 2003.
LTO	Emissions below 3000 ft taken from data tables	Follows AERO2k methodology	Dynamically modelled using INM, including taxi
Representative aircraft	40, chosen by seat capacity, engine technology, max take off weight, and PIANO availability. Representative engines chosen on basis of NO_x emissions.	43, based on AERO2k	None. In cruise performance taken from BADA and matching tables. Cruise speed taken from BADA procedure models.
Atmosphere	ISA		ICAO SA
Load factor	60.9% payload	70% payload	Weight assigned based on length of trip
Assumptions	LRC Airports at mslp EI=1.237 kg kg^{-1}	EI=1.237 kg kg^{-1}	Add equivalent weight of two stage lengths for tankering

Table 2.1: *Summary of the methods and models used in the construction of the AERO2k, QUANTIFY and AEDT inventories.*

Source	AERO2k	QUANTIFY	AEDT
Great circles	-4.5% mean true distance travelled	-12 to -13% fuel burn for EU	
Schedules		-9% compared to operational. Up to 50% regionally	
No VFR	-2% world movements		
LTO	97.7777% of operational data		INM and AEDT agree to within 10-20% on horizontal distance travelled
ISA	-1.5% AERO2k distance and operational due to lack of wind. Small difference due to fixed temperatures		When weather is included there is a 2.7% difference between AEDT and actual data. When weather is not included the difference is -7.87%.
L/D curves			AEDT within ±14% of NASA data
Representative aircraft		Uncertainty increases with distance. Different levels of uncertainty for different aircraft types.	
PIANO	Within 3% of airline models, and no more than 5% higher than BADA		
Fuel flow	$<$3% lower than actual values due to the lack of deterioration		6.95% greater than carrier data, 0.24% less than NASA data, and within 10% of data from Jane's aero engines
Holding	$\leq$1.2% underestimate compared to actual data for not modelling holding		
Weight assumptions	60.9% load factor gives best agreement with airline data. $\leq$0.5% underestimate due to not accounting for tankering.	70% load factor is close to world average. Possible 10% variation in weight due to tankering may cause a 2.5% change in fuel burn.	

Table 2.2: *Summary of the errors reported in the evaluation of the AERO2k, QUANTIFY and AEDT inventories.*

2.5 The ECMWF Re-Analyses: ERA

A re-analysis is an analysis produced using a record of past observations and a fixed data assimilation system. Re-analyses tend to be used in research instead of operational analyses as they offer a higher level of consistency. The methods used to produce operational analyses change over time as problems are identified and subsequently fixed. This can introduce discontinuities into the analysis. According to Uppala et al. (2005), these problems are often compounded by poor communication. Users of the analyses experience difficulties in knowing what alterations have been made and when, and what impact they have had. Further problems arise as it is difficult to know what impact changes in the observations have had over a given period. These difficulties are avoided, or reduced, when a reanalysis is used: the method of data assimilation remains the same for the duration of the re-analysis, and changes in the observing system are recorded and their effects can be tested. Two re-analysis products are used in this project: ERA-40 and ERA-Interim. The ERA re-analyses are global, and describe the state of both the atmosphere and the Earth's surface. ERA-Operational data is used for calculations beyond the end of ERA-40 for comparison with ERA-Interim calculations, and for calculations performed prior to the release of ERA-Interim.

ERA-40 covers the period between September 1957 and August 2002 (Uppala et al., 2005). It has a 60 model-level vertical resolution, up to 0.1 hPa, and a horizontal resolution of T159. The reanalysis contains directly analysed variables, horizontal wind components, temperature, specific humidity, and ozone, on model levels, and also surface pressure. These variables, and derived upper-air variables, including geopotential and vertical velocity, have been interpolated by ECMWF onto 23 isobaric levels, 15 isentropic levels , and the $|PV|=2$ PVU surface, with a horizontal resolution of $1.125^\circ \times 1.125^\circ$. Spectral model level data is used to calculate Lagrangian trajectories (see Chapter 5) and gridded data is used to calculate tropopause variables (see Chapters 3 and 4).

ERA-Interim begins in January 1989 and runs at approximately 10 days behind real time (Simmons et al., 2007). In this thesis, the period from January 1989 to December 2007 will be used. ERA-Interim is calculated using a model with T799 horizontal resolution and 91 model levels, and archived with a horizontal resolution of T255 on 60 levels. As for ERA-40, directly analysed and derived variables have been interpolated by ECMWF onto regular grids. For ERA-Interim these have 37 isobaric levels and 15 isentropic levels, and are provided with a horizontal resolution of $0.703^\circ \times 0.703^\circ$. Most of the increase in the vertical resolution of the gridded isobaric data between ERA-40 and ERA-Interim is due to the inclusion of extra levels in near the surface, and in the vicinity of the tropopause.

The quality of a reanalysis is dependent on the quality of both the observation and assimilation systems. Uppala et al. (2005) observe that the general quality of ERA-40 improves over time as the number and type of observations increase (see Figure 2.6). There are more observations assimilated in the Northern Hemisphere compared to the Southern Hemisphere, and consequently data for the Northern Hemisphere is of higher quality than that for the Southern Hemisphere (Bromwich and Fogt, 2004). ERA-Interim uses the same observations that were used in ERA-40, but with some additions. These include the assimilation of Geostationary Operational Environmental Satellite (GOES) water vapour radiances, MODerate resolution Imaging Spectroradiometer (MODIS) winds, and Solar Backscatter Ultraviolet Radiometer (SBUV) and Global Ozone Monitoring Experiment (GOME) ozone data (ECMWF, 2010).

ERA-40 uses a three dimensional variational assimilation system (3D Var) (Uppala et al., 2005). In 3D Var, observations within ±3 hours of the analysis time are assumed to occur at the analysis time. They are then combined with background information from the forecast model to give an analysis. The observations and background are combined using statistical estimates of their errors.[6] 4D Var, in which observations are used at the time they were taken, is used in ERA-Interim (Andersson and Thèpaut, 2008).

The use of 4D Var in ERA-Interim resulted in systematically improved forecast performance compared to ERA-40, especially in the Southern Hemisphere. The use of

---

[6] In ERA this involves minimisation of the ‘cost function’. This is the sum of the squared deviations of analysed values from observations and forecast, weighted by the accuracy of the observations and forecast.

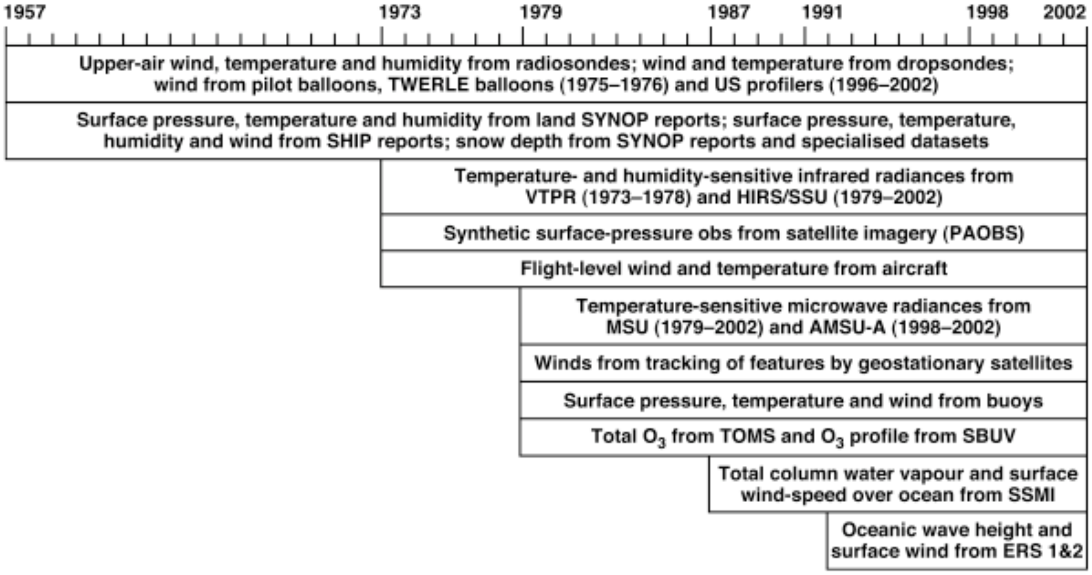


Figure 2.6: *Chronology of types of observations assimilated in ERA-40 from 1957 to 2002. (From Uppala et al. (2005))*

variational bias correction improves the consistency of the observations used in the analysis compared to ERA-40 (Uppala et al., 2008). ERA-Interim also benefits from a new humidity analysis (Simmons et al., 2007), and revised cloud physics and convection schemes, amongst other changes to the forecasting system (ECMWF, 2010).

The advantages of using ERA-Interim compared to ERA-40 are discussed in detail by Simmons et al. (2007). They include a significant reduction of the too-strong Brewer-Dobson circulation and a more realistic age of air in the stratosphere, key improvements in the water cycle, and systematic improvements in the Southern Hemisphere. The data are also calculated and archived at higher resolution, which is expected to be significant for this work.

Chapter 3

The Tropopause

The troposphere is generally characterised by rapid vertical mixing, and decreasing temperature with height (Figure 3.1). It contains most of the atmosphere's water vapour, and accounts for most of the total atmospheric mass, but only accounts for a small percentage of the total atmospheric depth. Conversely, the stratosphere is a very dry region, and its static stability means that little vertical mixing occurs. Temperature remains approximately constant with height in the lower layers of the stratosphere, while the influence of ozone causes an increase in temperature with height in the upper layers (Seinfeld and Pandis, 1998).

Because the troposphere and stratosphere have a number of contrasting properties, it is possible to define the tropopause, the boundary between the troposphere and stratosphere, using several different diagnostics, with no universally right or wrong choice. As such, there is little concurrence between the tropopause definitions used in different studies, even those evaluating the impacts of aviation on climate. Thermal definitions identify the tropopause from a change in lapse rate, or a temperature minimum; the change in stability from troposphere to stratosphere can be used to give a potential vorticity (PV) definition; and abrupt transitions in water vapour, ozone, and other trace gas concentrations can be used to define the tropopause as a chemical transition layer.

The concept of the tropopause has been evolving over the last century. In 1909 Gold used

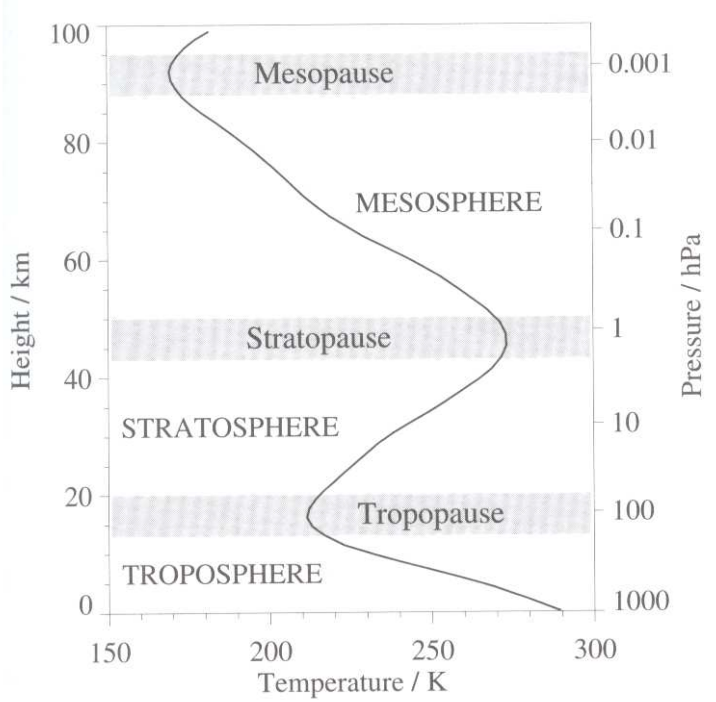


Figure 3.1: *Typical vertical structure of temperature in the lowest 100 km of the atmosphere. (From Andrews (2000))*

balloons to investigate the nature of the upper levels of the atmosphere. These experiments 'revealed the existence of an abnormal change in the vertical temperature gradient' that we now know to be associated with the tropopause. Gold (1909), in agreement with the earlier observations of Teisserenc de Bort in 1899, found that the temperature of the atmosphere decreased with height at a rate of approximately 6°C km^{-1}, before reaching a height above which the temperature remained almost constant (Figure 3.1). Gold observed similar temperature profiles in several locations in the Northern Hemisphere, allowing him to conclude that this was not simply a local phenomenon. The collection of profiles also revealed that the change in the vertical temperature gradient tended to occur at higher altitudes near to the equator (Figure 3.2).

Gold (1909) was able to reproduce many of the observed characteristics of the atmosphere using a simple radiative model, where radiative transfer prevented the establishment of a temperature gradient necessary for convection at upper levels, while convec-

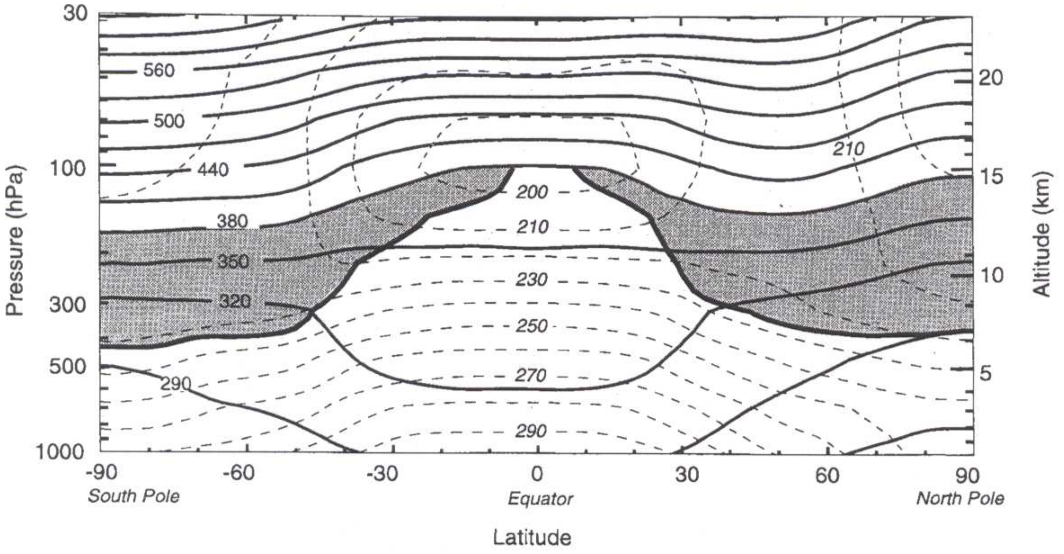


Figure 3.2: *Latitude-altitude cross section of the lower atmosphere in January 1993. Contours show longitudinally averaged temperature (dashed) and potential temperature temperature (solid). The heavy contour denotes a dynamic tropopause. (From Holton et al. (1995))*

tion was maintained at lower levels by the transfer of energy from the Earth's surface. Gold also concluded that the observed latitudinal variation in tropopause height (Figure 3.2) was due to variations in the atmospheric water vapour content. However, he was unable to reproduce the magnitude of the observed variation in height. Simpson (1928) found that the assumption that atmospheric water vapour can be treated as a grey body, as used by Gold (1909), tends to lead to poor agreement with observations. He suggested that the latitudinal variation in the height of the tropopause was likely to be controlled not by the thermodynamics of the stratosphere, but by the dynamics of the troposphere.

Several decades after Simpson's suggestion that the height of the tropopause may be dependent on tropospheric dynamics, Manabe and Strickler (1964) considered a series of radiative and thermal equilibrium models of the atmosphere. Like Gold (1909), they were able to reproduce the temperature inversion above the tropopause. They also realised the importance of the absorbing properties of ozone in this region, and computed a series of thermal equilibria for typical distributions of absorbers at different latitudes. They concluded that the latitudinal variation in the concentrations of both water vapour and ozone could be partly responsible for the latitudinal variation in the thickness of the isothermal layer. However, they found that they were still unable to reproduce the magnitude of the latitudinal variation in the height of the tropopause, and concluded that, as suggested by Simpson (1928), the 'dynamics of large scale motion' must therefore be important.

The characteristics of diabatic heating are very different in the tropics and the midlatitudes. Moist convection is dominant in the tropics, while large-scale eddies play a significant role in the midlatitudes. Held (1982) realised that this would act as a barrier to the construction of models of the global-mean vertical temperature profile, like the radiative-convective models used in earlier studies. In an attempt to provide a qualitative theory for the height of the tropopause that did not require detailed knowledge of the vertical structure of dynamical heating, Held suggested three constraints: a global radiative constraint, and a dynamical constraint for each of the tropics and the midlatitudes. Using the Charney (1947) conceptual model in midlatitudes, and a model of moist static stability in the tropics, Held was able to reproduce the observed variation in tropopause height.

The height of the tropopause tends to lie in the range 9 to 16 km. This range includes the typical cruise altitude of commercial aircraft, 10 to 11 km, and therefore the definition of the tropopause is important in determining the calculated proportion of emissions that enter the stratosphere directly (Gettelman and Baughcum, 1999, Forster et al., 2003). The tropopause is often significantly lower than 9 km in localised regions, and as a result the time spent by aircraft in the stratosphere can often be underestimated if a climatological tropopause is used (RCEP, 2002). Because of the potential sensitivity of the calculated climate impact of water vapour emissions to the location of the tropopause, as demonstrated by Gettelman (1998) for example, it is important that a well-based and consistent definition is used.

In this chapter a thermal and dynamic tropopause definition will be considered in detail in Sections 3.1 to 3.3 before a global tropopause definition is introduced in Section 3.4. The sensitivity of the tropopause definitions to the ERA data set used in their calculation is discussed in Section 3.5.

3.1 The thermal tropopause

The temperature profiles in the troposphere and stratosphere are very different: while temperature tends to decrease with height in the troposphere, with an average lapse rate, $\frac{dT}{dz}$, of -6.5 °C km^{-1}, it remains constant or even increases with height in the stratosphere. These differences in temperature profile, and the discontinuous change between them, mean that the temperature lapse rate can be used to define the tropopause. Conventionally, the tropopause defined by the World Meteorological Organization (WMO) in 1957 is used. It often forms the basis of studies comparing different tropopause definitions (Hoerling et al. (1991), Gettelman and Baughcum, 1999, Forster et al., 2003, for example), and is used to calculate the 'reported tropopause' from radiosonde ascents. The WMO defines the 'first tropopause' as "*the lowest level at which the average lapse rate* *between this level and all higher levels within 2 km does not exceed 2°C km^{-1}. When above the first tropopause, if the average lapse rate between any level and all higher levels within 1 km exceeds 3°C km^{-1}, then a 'second tropopause' can occur and is defined by the first criteria. This tropopause can either be above or within the 1 km layer. Further tropopauses may be defined similarly.*" In this study a thermal tropopause definition is based on the 'first tropopause' of the WMO definition: the 'thermal tropopause' is defined as the lowest level above 500 hPa at which the temperature lapse rate is greater than or equal to -2 °C km^{-1}, provided that the lapse rate between this level and all higher levels within 2 km has an average value of not less than -2 ° km^{-1}. 500 hPa is defined as the maximum thermal tropopause pressure to avoid a tropopause that either intersects the ground, or has an anomalously low altitude, following Gettelman and Baughcum (1999).

The thermal tropopause is calculated using ERA-40 data on pressure levels. Restricting the tropopause to these pressure levels would result in discontinuities, so linear interpolation is used to find the thermal tropopause. The lapse rate and height at the midpoint between each of the pressure levels is calculated. The height, z_{low}, at which the lapse rate, $\Gamma = \frac{dT}{dz}$, is equal to -2 ° km^{-1} can be found from these values using linear interpolation:

$$z_{low} = z_{-\frac{1}{2}} + \frac{(-2 - \Gamma_{-\frac{1}{2}})(z_{+\frac{1}{2}} - z_{-\frac{1}{2}})}{(\Gamma_{+\frac{1}{2}} - \Gamma_{-\frac{1}{2}})} \tag{3.1.1}$$

where subscripts of $-\frac{1}{2}$ and $+\frac{1}{2}$ denote the heights and lapse rates of the midpoints at the lower and upper boundaries of the interpolation region respectively. If the mean lapse rate is greater than or equal to -2 ° km^{-1} for a depth of 2 km or more above z$_{low}$, then z$_{low}$ is taken as the height of the thermal tropopause.

3.2 The dynamic tropopause

PV is the product of the thermal stability along the direction of absolute vorticity and the magnitude of the vorticity of an air parcel divided by the density. The Ertel form of

PV, which is relevant for adiabatic frictionless motion, is written as:

$$P = \frac{1}{\rho}\zeta_a \cdot \nabla\theta \qquad (3.2.1)$$

where ζ_a is the vertical component of absolute vorticity (Equation 3.2.2), θ is potential temperature (Equation 3.2.3), and ρ is the fluid density (Hoskins et al., 1985). Making the hydrostatic approximation, the vertical component of absolute vorticity is given by:

$$\zeta_a = f + \mathbf{k} \cdot (\nabla \times \mathbf{v}) \qquad (3.2.2)$$

where f is the Coriolis parameter, **k** is the unit vector in the vertical, and **v** is the horizontal wind vector. The potential temperature is the temperature that an air parcel would have if it were expanded or compressed adiabatically from its existing pressure and temperature to a standard pressure, p_0. For an ideal gas, potential temperature is given by the equation:

$$\theta = T\left(\frac{p_0}{p}\right)^{\kappa} \qquad (3.2.3)$$

where T is the temperature of the parcel, p is the environmental pressure, p_0 is a standard pressure, usually taken to be 1000 hPa, and κ is the ratio $\frac{R}{c_p} \approx 0.286$, where R is the gas constant and c_p is the specific heat capacity at constant pressure.

PV, as defined in Equation 3.2.1, is conserved following the motion for adiabatic, frictionless flow. Making the anelastic, Boussinesq, and small Rossby number approximations, it is possible to define quasi-geostrophic potential vorticity (QGPV).[1] QGPV is conserved following the horizontal geostrophic flow for frictionless , adiabatic motion, and is an approximation for the conservation of PV moving with the full flow. QGPV is defined as:

$$q = f_0 + \beta y + \nabla_h^2\psi_g + f_0^2\frac{\partial}{\partial z}\left(\frac{1}{N^2}\frac{\partial\psi_g}{\partial z}\right) \qquad (3.2.4)$$

---

[1] In the anelastic approximation it is assumed that variations in density and pressure are dominated by the exponential decrease with height necessary for hydrostatic balance in the atmosphere. In the Boussinesq approximation it is assumed that density is constant, except in the buoyancy term of the thermodynamic equation. This is equivalent to assuming that the height scale over which density varies is much more than the height scale of weather systems. The small Rossby number approximation (Ro<<1) is an assumption that systems are close to geostrophic balance.

where ζ_g is the vertical component of geostrophic absolute vorticity; f_0 is the Coriolis parameter, assumed to be constant; b' is buoyancy; N^2 is the Brunt-Väisälä frequency; ψ_g is the geostrophic streamfunction; and β is given by:

$$\beta = 2\left(\frac{\Omega}{a}\right) cos\phi_0 \tag{3.2.5}$$

where Ω is the Earth's angular velocity; a is the Earth's radius; and ϕ_0 is the latitude of the origin of the local Cartesian coordinates. Following Equation 3.2.4, the perturbation QGPV (from the basic state, f) can be written as:

$$q' = \bigtriangledown_h^2 \psi_g + \frac{f_0^2}{\rho}\frac{\partial}{\partial z}\left(\frac{\rho}{N^2}\frac{\partial \psi_g}{\partial g}\right) \tag{3.2.6}$$

Assuming that N^2 is a constant, and rescaling height so that $\hat{z} = \left(\frac{N}{f_0}\right) z$ allows Equation 3.2.6 to be written in the form of a Poisson equation:

$$q' = \bigtriangledown \psi_g \tag{3.2.7}$$

This means that if q' is known then the expression for q' can be integrated (inverted) to find ψ_g, given known boundary conditions, and hence find the circulation and potential temperature anomalies associated with that q'. It is the Lagrangian conservation properties and invertability of PV that allow it to be used to give a concise insight into atmospheric dynamics. Observations studies, Reed (1955) and Danielsen (1964) for example, have shown that PV is indeed a useful airmass tracer.

It can be found from Equation 3.2.1 that, on synoptic and larger scales where the vertical components dominate:

$$\frac{dP}{dt} \simeq -g(\zeta + f)\frac{\partial \dot{\theta}}{\partial p} \tag{3.2.8}$$

where $\dot{\theta}$ is the diabatic heating rate. From this equation, it can be seen that heating above or cooling below an air parcel (increasing the vertical gradient of diabatic heating) causes an increase in the PV of a parcel (Martin, 2006). It follows from the thermal structure of the atmosphere (see Figure 3.1) that stratospheric air, with its large stratification, $\frac{\partial \theta}{\partial z}$, should have much larger values of PV than tropospheric air. Reed (1955), and Reed and

Danielsen (1959), showed that the PV of stratospheric air may be as many as three orders of magnitude greater than that of tropospheric air. This difference allows a tropopause to be defined either using the vertical gradient of PV, or assigned to the position where PV reaches a chosen critical value.

Critical values of PV are often used to define a dynamic tropopause, and tend to be in the range $|PV|$=1.6 - 3.5 PVU (where 1 PVU = 1×10^6 K m^2 kg^{-1} s^{-1}), although there is no consensus as to where in this range the most useful value lies. Hoerling et al. (1991) concluded that a critical value of 3.5 PVU gave a dynamic tropopause that had the best agreement with observations of tropopause pressure extracted from rawinsonde data using the thermal definition, and also with the rms values of PV on the thermal tropopause poleward of 20° latitude. However, it should be noted that this conclusion was reached using relatively poor vertical resolution data with only 15 isobaric surfaces. Holton et al. (1995) found that the 2 PVU surface was ‘remarkably close’ to the observed WMO extratropical tropopause, while the WMO (1986) definition places the tropopause at a critical value of 1.6 PVU. Danielsen (1964), however, used simultaneous measurements of ozone and fallout from nuclear tests (both stratospheric tracers), and calculations of potential vorticity from observations of wind and potential temperature, to show that the lowest reported tropopause was associated with a critical value at or near to 1 PVU. When comparing such results, it should be noted that the results of model studies, like those of Hoerling et al. (1991), are sensitive to both the resolution of the data set used and the data set itself. Pan et al. (2004) found that the 2 PVU surface was generally $\sim$1 km lower than the thermal tropopause in the extratropics, but that it was in very good agreement with a tropopause identified from the ozone profile. They also identified a tropopause from the carbon monoxide profile, which was $\sim$1 km above the thermal tropopause. These results suggest that the tropopause lies between 1 and 4 PVU.

In this study the ‘dynamic tropopause’ is defined as the lower of the surface where the absolute value of potential vorticity is equal to 2 PVU, and the 380 K isentrope. The 2 PVU surface lies in the middle of the range of values found in observational studies (Danielsen, 1968, Shapiro, 1978, 1980) and follows the definition most commonly used in previous studies e.g. Hoskins (1988, 1991) and Forster et al., (2003). The definition is not based on the PV surface alone as the vertical component of absolute vorticity changes sign at the equator, causing the height of the surface to tend to infinity there. The 380 K isentrope, shown by Holton et al. (1995) to be a good approximation to the tropical tropopause, acts as a cap on the height of the tropopause in the equatorial regions where the PV definition is not applicable (see Figure 3.2).

Although 1-4 PVU is a clear transition region between the troposphere and the stratosphere, tropopauses defined using the lower values of PV in this range tend to give more detail of synoptic scale events, such as tropopause folding. As these events are thought to play a key role in stratosphere to troposphere exchange (e.g. Lamarque and Hess, 1994) it is necessary to define a tropopause that will resolve them when the removal of emissions from the stratosphere is being considered in Chapter 6.

ERA data is provided on the PV=2 PVU surface. For ERA-40, this data is the result of calculations by ECMWF using 60 model levels, while a tropopause placed at any other critical value would need to be calculated from the available pressure level data, as used to calculate the thermal tropopause, which has 23 vertical levels. As there is no strong case for not using a value of 2 PVU, this data is used to avoid performing vertical interpolation between relatively coarse isobaric levels.

3.3 Analysis of the thermal and dynamic tropopauses

The detailed properties of the thermal and dynamic tropopauses have been investigated and compared through two case studies, 1 January 2002, 0Z and 1 August 2002, 0Z, providing examples of the surfaces from different seasons.

Figure 3.3 illustrates the height above sea level of the thermal and dynamic tropopauses on 1 January 2002 at 0Z. The difference between the two surfaces is also shown. While some significant differences between the heights of the surfaces produced by the two definitions are apparent, mainly in the sub-tropics, the same synoptic features can be identified from both plots. However, more synoptic structure can be seen in the dynamic tropopause. In the Northern Hemisphere there are ridges over the western United States and Canada, the North Atlantic and central Asia. At most longitudes there is a steep gradient in tropopause height near 30°N associated with the sub-tropical jet. In the Southern Hemisphere there is a large area of low tropopause height to the west of the tip of South America, and another south of Africa. Regions of high tropopause height extend poleward south-west of Australia, and equatorward over the South Pacific and South Atlantic. The height gradient near 30°S is more gradual than that in the Northern Hemisphere. This is to be expected as temperature gradients are weaker in summer, resulting in a weaker jet stream, and weaker sub-tropical tropopause height gradients. The difference in these gradients is illustrated more clearly in the tropopause height - latitude section shown in Figure 3.4.

Cross sections through the tropopause along the Greenwich meridian are shown in Figure 3.4 for the January and August case studies. It can be seen in Figure 3.4 that there is reasonable agreement between the thermal and dynamic tropopause heights in both the Northern and Southern Hemisphere extratropics. The dynamic tropopause is generally a few hundred metres lower than the thermal tropopause, but at the same height in some places (see Figure 3.4). This height difference is much smaller than that found by Pan et al. (2004) in their study. In a reflection of its dynamical nature, the dynamic tropopause definition allows the resolution of greater small scale detail than the thermal tropopause definition. As both the thermal and dynamic tropopause definitions are arbitrary, it would not have been useful to sacrifice some of this resolution in order to use a higher value of PV and produce and better agreement in height between the two surfaces. It can be seen in Figure 3.4 that where there are depressions in the height of the tropopause in the extratropics the thermal tropopause definition tends to underestimate their depth relative to the dynamic definition. The very high thermal tropopause heights that can be seen in the high Southern Hemisphere latitudes in August are discussed later in this section.

The dynamic tropopause is capped at 380 K (approximately 17 km), which was shown by Holton (1995) to be a good approximation to the tropical tropopause. This cap means that the dynamic tropopause is almost flat across the equator. The discontinuity associ-

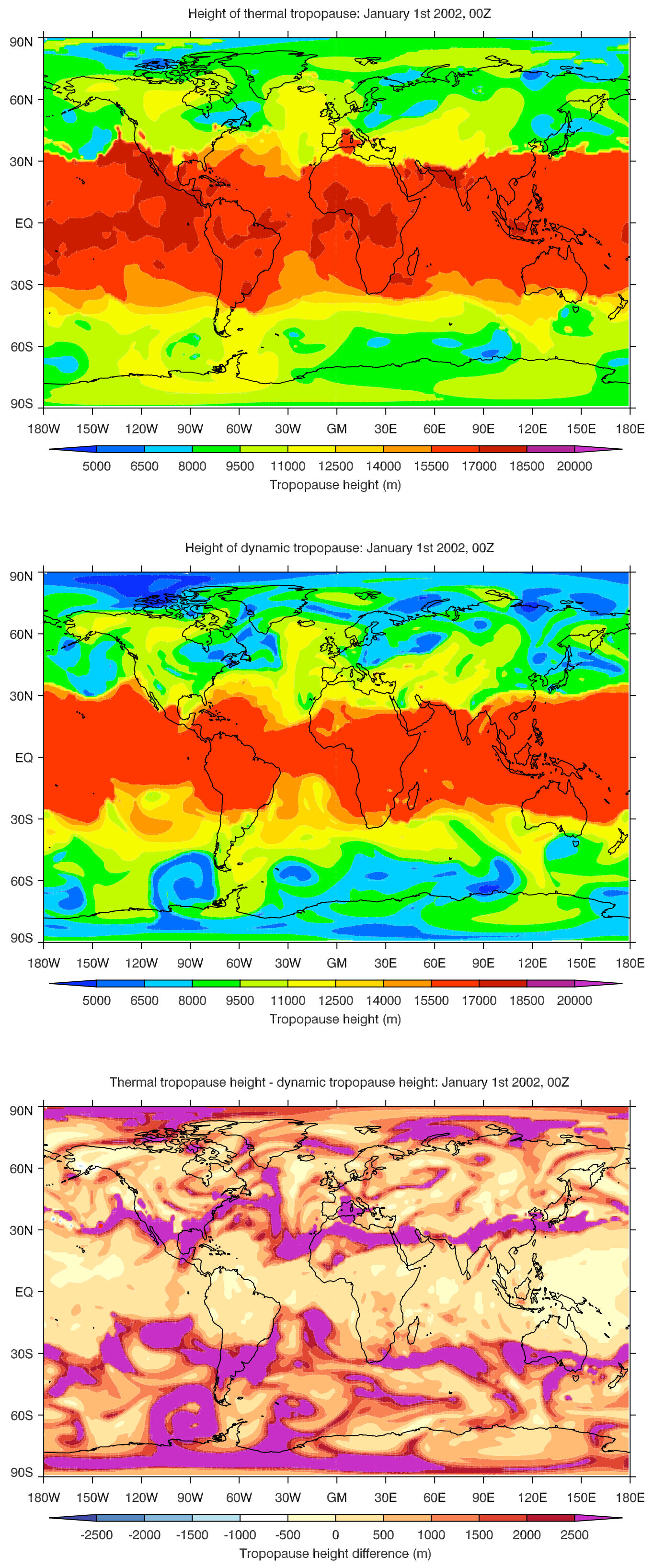
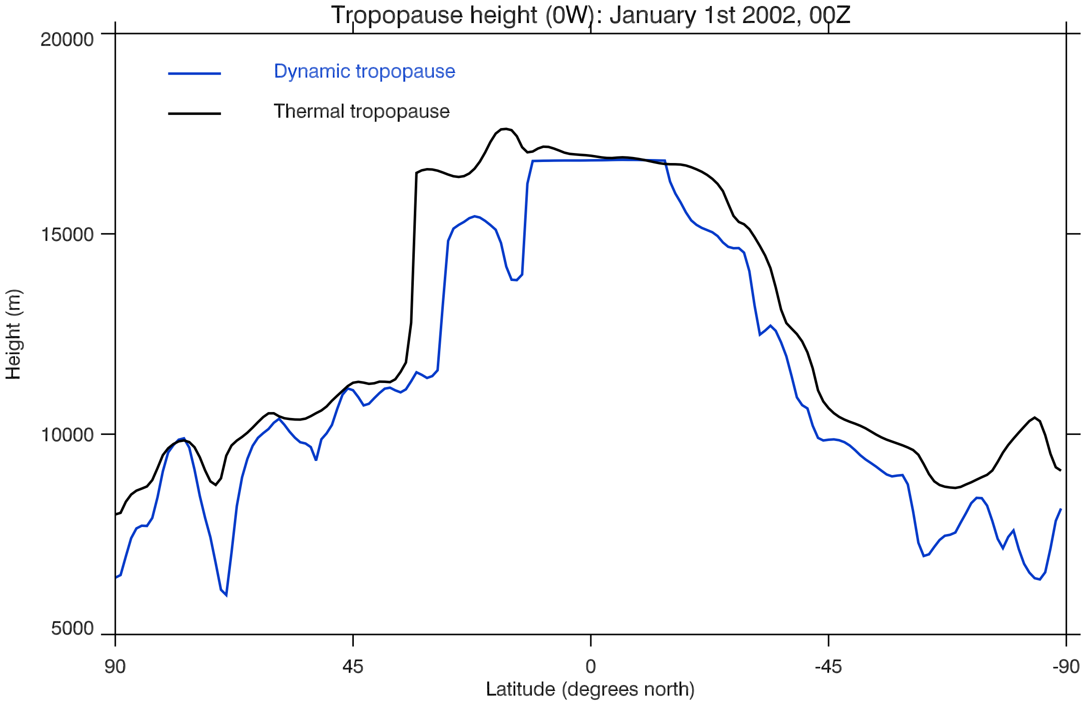


Figure 3.3: *Height (in metres) of the thermal (top) and dynamic (middle) tropopauses for 1 January 2002 at 0Z. The difference (in metres) between the two plots is shown in the bottom panel.*



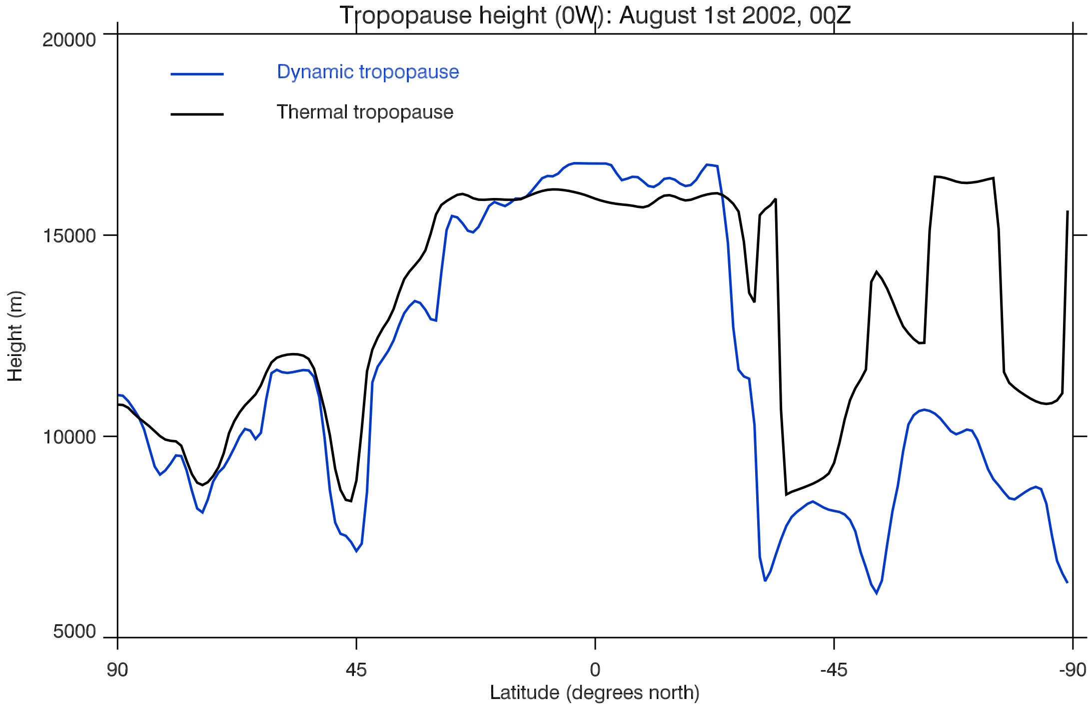


Figure 3.4: *Cross section along the Greenwich Meridian of the thermal and dynamic tropopauses for 1 January 2002 at 0Z (top) and 1 August 2002 at 0Z (bottom).*

ated with this can be seen in Figure 3.4. Comparison with the thermal tropopause shown in the same Figure shows that this discontinuity is much less severe than the one found naturally in the thermal tropopause at the tropopause break. It can also be seen that the altitude of the two tropopause surfaces are in good agreement in the tropics.

Figure 3.5 illustrates the height of the thermal and dynamic tropopauses on 1 August 2002 at 0Z. There are conspicuous differences between the two tropopauses in the Southern Hemisphere, but in general they show the same features. In the Northern Hemisphere there are troughs over Canada, Greenland and Russia, and a small region of low tropopause height extending over the British Isles and France. A large region of high tropopause height, extending from the Arabian Peninsula to Japan, is the dominant feature in the sub-tropics, and is present on both of the tropopauses. However, there are very sharp gradients on the northern edge of the feature in the thermal tropopause, suggesting that it may be multi-valued in this region. In the Southern Hemisphere there are large regions of low tropopause height over the south Indian Ocean, and to the west of the southern tip of South America.

As in the January case study, the dynamic tropopause resolves smaller scale features than the thermal tropopause. And again, there are steep height gradients near 30°S, which are slightly further poleward on the thermal tropopause.

The large regions of very high tropopause height over the Antarctic on the thermal tropopause are in marked contrast to the low dynamic tropopause, which has smooth height gradients. This difference arises as the thermal tropopause becomes unreliable in regions where there is little vertical mixing in the troposphere. In the absence of vertical mixing there is no strong change in lapse rate at heights where the tropopause is expected to lie, and so the WMO criteria are not met. When the average lapse rate does not have an average value of 2 °C km^{-1} of less for a depth of 2 km at any point in the profile, the thermal tropopause is placed at 500 hPa to ensure that a tropopause is defined at each grid point.

Anomalously high tropopause altitudes (as can be seen the Antarctic in Figure 3.5) arise when the WMO (1957) criteria are satisfied at atypically high altitudes. It is not possible

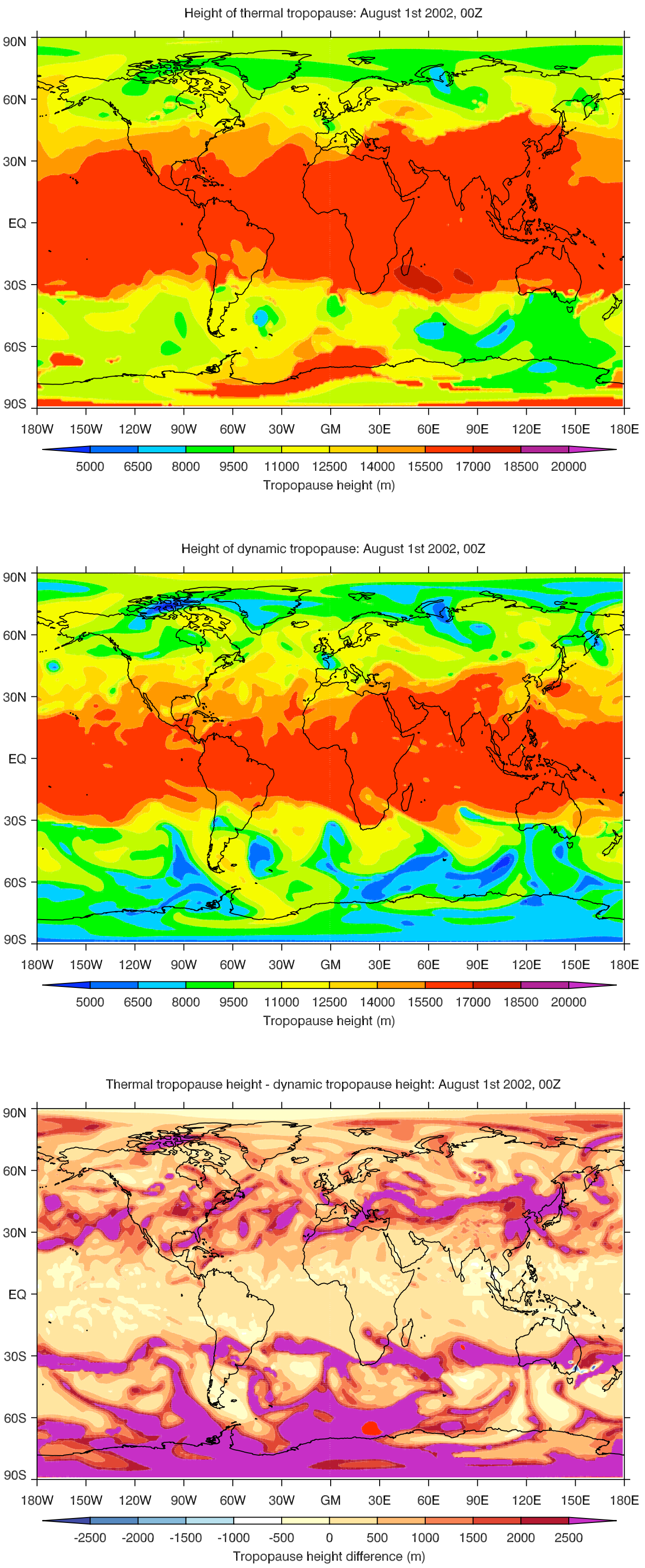


Figure 3.5: *Height (in metres) of the thermal and dynamic tropopauses for 1 August 2002 at 0Z. The thermal tropopause (top) is unreliable over the Antarctic, giving an artificially high tropopause, in strong contrast to the dynamic tropopause (middle). The difference between the two surfaces (in metres) is shown in the bottom panel.*

to prevent these anomalies by applying a global cap that forces the tropopause to the 500 hPa level if the WMO criteria places it at a very high altitude as the tropical tropopause is often at a higher altitude than the Antarctic anomalies. The magnitude of this problem is illustrated in Figures 3.4, and 3.5. In Figure 3.4 the thermal and dynamic tropopauses are in good agreement until 40°S when the thermal tropopause jumps to a height in excess of its height in the tropics, and more than 50% higher than the dynamic tropopause at the same point. A similar problem can occur in the Arctic region in boreal winter months, but the problem arises less frequently, and tends to affect only a small area.

Comparison of Figures 3.6 and 3.7 gives further insight into the difficulties of using the thermal tropopause in the polar night regions. Figure 3.6 shows temperature and lapse rate profiles for Brest, France, in January and August. Figure 3.7 is a similar figure for Davis, Antarctica. Sharp changes in both temperature and lapse rate gradient are observed near to 400 hPa in both sets of French profiles, and the January Antarctic profiles. The temperature profile for the August Antarctic ascent, by comparison, displays a much smoother transition to the isothermal layer. There is no clear lapse rate maximum below 100 hPa in the lower right panel of Figure 3.7. Both January Antarctic profiles have clear features associated with the tropopause near 400 hPa. Similar features do not occur in the August profiles at pressures greater than 10 hPa, and it is at these higher altitudes that the thermal tropopause is placed.

The PV profiles for Brest and Davis on 1 August 2002, 0Z, are shown in Figure 3.8. Tropospheric PV of $<$1 PVU can be seen in both profiles, along with a clear transition region between 1 and 6 PVU, before PV increases linearly with height in the stratosphere. The clarity and similarity of the both profiles, in contrast to the comparison between the lapse rate profiles, demonstrates that the dynamic tropopause is more suitable for use in the high latitudes.

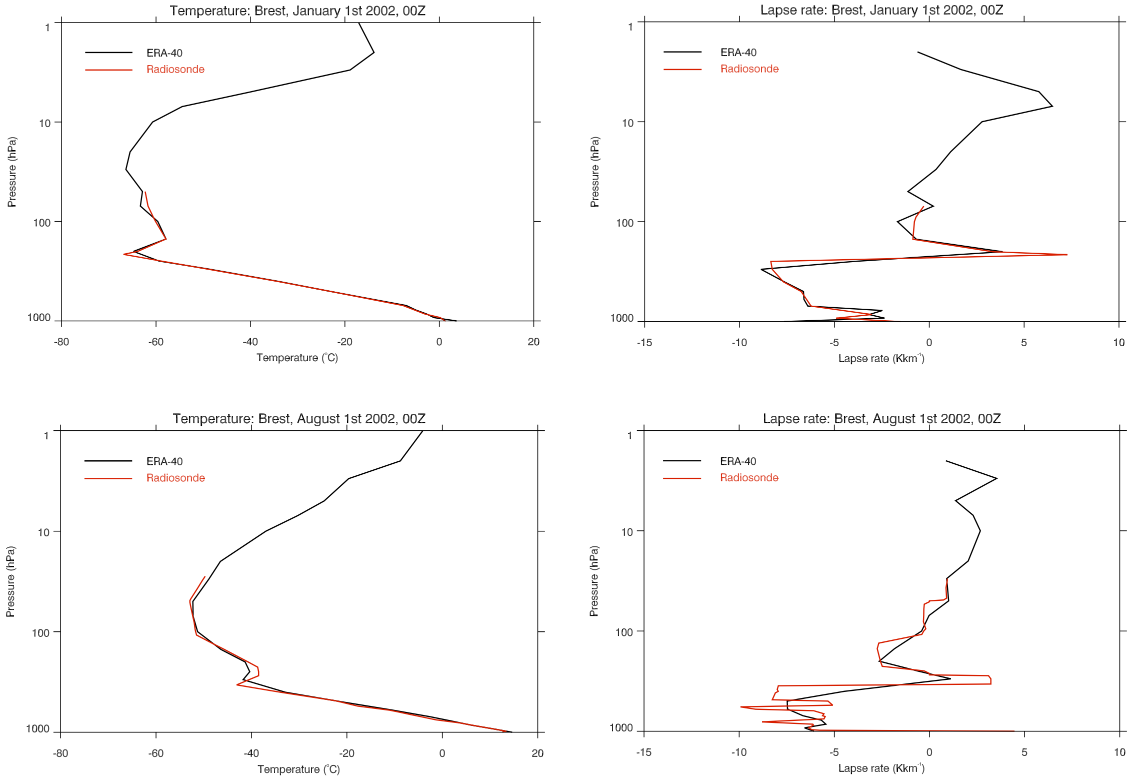


Figure 3.6: *Temperature (left) and lapse rate (right) profiles for Brest, France, generated from both the ERA-40 reanalysis and Met Office radiosonde data for 1 January 2002, 0Z, (top) and 1 August 2002, 0Z, (bottom).*

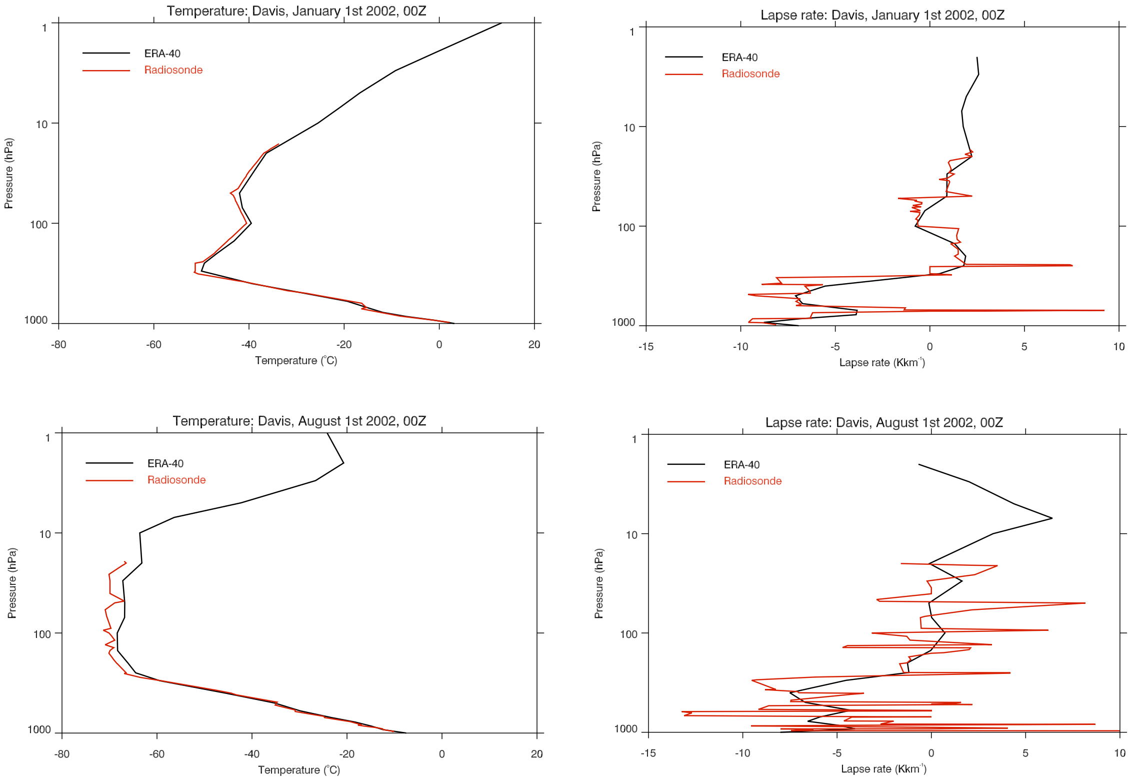


Figure 3.7: *Temperature (left) and lapse rate (right) profiles for Davis, Antarctica, generated from both the ERA-40 reanalysis and Met Office radiosonde data for 1 January 2002, 0Z (top) and 1 August 2002, 0Z, (bottom).*

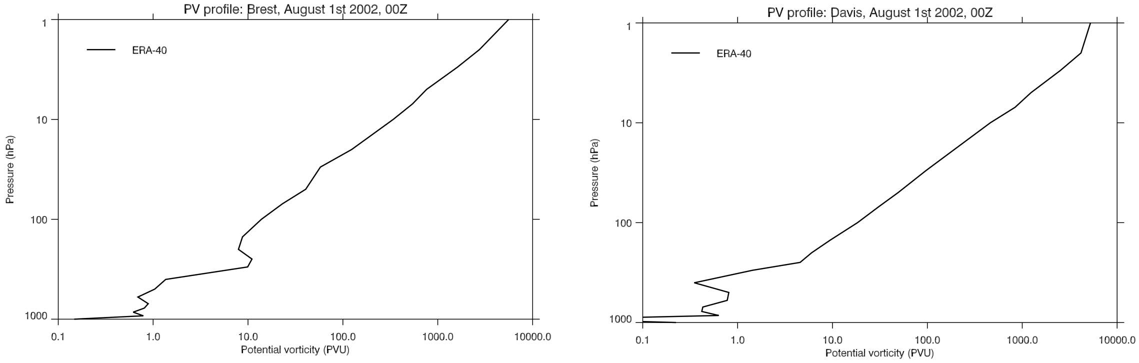


Figure 3.8: *PV profiles from ERA-40 for Brest (left) and Davis (right), 1 August 2002, 0Z.*

3.3.1 Advantages/disadvantages of the two definitions

Neither the thermal nor the dynamic tropopause can confidently be used globally, despite the thermal tropopause definition being globally applicable. As the 1 August 2002 case study demonstrated, the thermal tropopause definition becomes unreliable in the polar night regions (see Figures 3.4, 3.5, and 3.7). Although this problem is most apparent in the winter polar regions, the Antarctic in particular, sharp gradients in the height of the thermal tropopause can be found in regions of the extratropics. Wave disturbances causing convergence also cause a change in the static stability, leading to a displacement of the thermal tropopause as it jumps from the base of one stable layer to another (see Figure 3.3, (180E, 30N), or Figure 3.5, (150E, 30S), for example).

The dynamic tropopause definition includes a potential temperature cap close to the equator. The vertical component of absolute vorticity changes sign at the equator, causing the height of any constant PV surface to tend to infinity. Dameris (2003) concluded that, because of this, the dynamic tropopause is only really applicable poleward of 10° latitude. It can be seen in the 1 January 2002 cross section in Figure 3.4 that the 380 K cap is only used equatorward of 10° latitude. However, in the 1 August 2002 cross section the cap is used at around 20°S (Figure 3.4), suggesting that Dameris' conclusion may be slightly optimistic.

Holton et al. (1995), amongst others, note that the thermal tropopause definition, which does not evolve with airmasses, obscures the fact that the tropopause often behaves like a material surface. While the thermal tropopause is unable to characterise air masses because the vertical temperature gradient is not a conserved quantity, potential vorticity, and therefore the dynamic tropopause, is materially conserved in adiabatic motions. The behaviour of the tropopause as a material surface is consistent with the observation that the tropopause corresponds roughly to an isentropic surface in the tropics and a surface of constant potential vorticity in the extratropics. This property is utilised in the definition of a globally valid 'blended tropopause' in the following section.

Another advantage of the dynamic tropopause is the resolution of synoptic scale features in the extratropics, particularly the resolution of deep depressions in the tropopause (Hoerling et al., 1991). Danielsen (1968) also cites this as one of the main disadvantages of a dynamic tropopause definition compared to a thermal tropopause definition as, because it can have more than one value at a given point, the height of the tropopause cannot be projected onto a horizontal map for a tropopause folding event. However, the dynamic tropopause definition remains consistent in these situations, despite being multi-valued. When multiple |PV|=2 PVU surfaces are detected in the ERA analysis system, the data archived as the |PV|=2 PVU surface uses the highest altitude level. This problem does not arise for the thermal tropopause, which is defined as the lowest level where the lapse rate condition is met. However, the thermal tropopause can be difficult to identify at all in such situations, especially if the data has coarse vertical resolution.

It is possible that the resolution of synoptic scale features, such as tropopause folding events, may have an influence on the results of the trajectory analysis used in this work. It can be seen in Figure 3.9, which shows a cross section through a tropopause folding event, that maximum upper level winds occur at the focus of the tropopause fold. It is in this region that aircraft are likely to fly. The resolution of synoptic features will also be of importance for determining whether airmasses should be described as stratospheric or tropospheric when mixing occurs at the tropopause.

Previous studies have shown the thermal tropopause definition, and to some extent the dynamic tropopause definition, to be limited by the resolution of the data set used to calculate it. Hoerling et al. (1991) (for example) found that the thermal tropopause calculated from the ECMWF level IIIb data set differed from reported tropopause pressures, calculated from radiosonde observations using the same criteria, by as much as 80 mb. In this study, the resolution of the data set is likely to have more of an influence on the thermal tropopause than the dynamic tropopause, as the data used to calculate the dynamic tropopause has a much higher vertical resolution. The sensitivity of both tropopause definitions to the data set used to calculate them is discussed in Section 3.5.

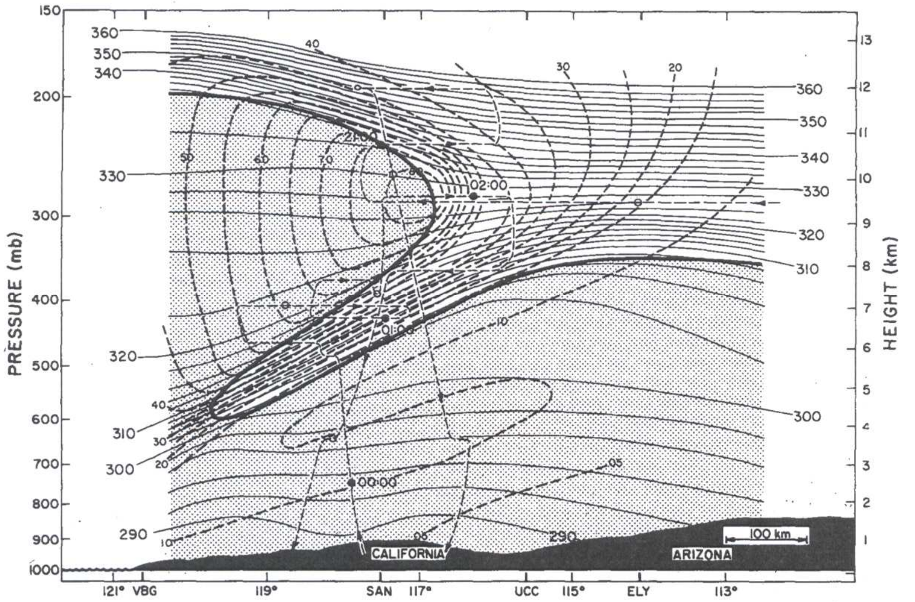


Figure 3.9: *Cross section through a tropopause folding event (13 March 1978), showing the 1 PVU surface (heavy solid contour), potential temperature (thin solid contours), wind speed (heavy dashed contours), the troposphere (stippled area) and a research aircraft flight track (thin dashed lines). (From Shapiro (1980))*

3.4 The blended tropopause

Held (1982) examined the different factors influencing tropopause position in the tropics and extratropics. He concluded that a theory for global-mean static stability would be of little use due to the differences in dynamics between the tropics, where moist convection dominates, and extratropics where synoptic-scale eddies are more significant. The results in the previous sections have highlighted some of the difficulties in applying a single tropopause definition globally. The dynamic tropopause definition includes a potential temperature cap in the deep tropics, and the thermal tropopause is sometimes unsuitable at high latitudes, and in active weather systems. A truly global tropopause analysis requires a combination of both the thermal and dynamic tropopause definitions in order to take these differences into account. A variety of combinations have been used in previous studies. They can be grouped into latitude-dependent combinations, and Lagrangian-style combinations where environmental tracers are used to delineate the tropopause.

Previous studies exploring the climate impact of aviation have tended to use a latitude-dependent tropopause definition (Hoinka, 1999, and Forster et al., 2003 for example). In these combinations, a dynamic definition is applied poleward of a given latitude ($\sim 30°$), and a thermal definition is applied equatorward of another latitude ($\sim 20°$), with linear interpolation used in between to give a smooth transition. However, even when a complicated weighting function, such as that developed by Hoerling et al. (1991), is used in the transition region, the latitude dependence of the definition is reflected in the tropopause surface itself.

Although they are simple and inexpensive to compute, latitude-dependent definitions can obscure the dynamic situation. For example, air can be transported between the extratropics and the tropics, and in these situations it is likely to be more appropriate to use a dynamic tropopause definition deeper into the tropics.

Schoeberl and Morris (2000) and Morris et al. (2003), avoid the problems of latitude dependence by defining the tropopause as the position of the height minimum of the thermal and dynamic tropopauses. This avoids the problem of the height of the dynamic tropopause tending to infinity at the equator, and is also beneficial in the polar night regions where the thermal tropopause tends to be artificially high.

In this study the interpolation region between the thermal and dynamic tropopauses is bounded by the point at which the dynamic tropopause intersects the 350 K and 370 K isentropic surfaces. This prevents the introduction of latitude dependence into the definition. The interpolation between the thermal and dynamic tropopauses is performed vertically, in potential temperature (θ) coordinates, so that no explicit latitude dependence is introduced through the computational method.

The steep gradients in tropopause height in the sub-tropics tend to be found further poleward when the thermal tropopause definition is used, compared to the dynamic tropopause definition, especially in the winter hemisphere. Additionally, there can be more than one latitude in each hemisphere where the PV=2 PVU surface crosses the 370 K surface for any given longitude. When both of these events occur at the same longitude, a simple combination of the dynamic and thermal tropopause definitions can result in discontinuities. Thus, in order for the blended tropopause to be both continuous and mathematically consistent, it has been defined as a combination of the dynamic tropopause and the lower of the dynamic and thermal tropopauses, h_{min} (Figure 3.10).

The height of the blended tropopause, $h_{blended}$, is given by:

$$h_{blended} = \begin{cases} h_{dynamic} & \text{where } \theta_{PV2} < 350 \text{ K} \\ \left(\frac{\theta_{PV2}-350}{20}\right) h_{min} + \left(1 - \frac{\theta_{PV2}-350}{20}\right) h_{dynamic} & \text{where } 350 \leq \theta_{PV2} \leq 370 \text{ K} \\ h_{min} & \text{where } \theta_{PV2} > 370 \text{ K} \end{cases}$$

where θ_{PV2} is potential temperature on the dynamic tropopause and $h_{dynamic}$ is the height of the dynamic tropopause.

There are very few obvious differences between the dynamic and blended tropopauses when a 1500 m contour interval is used. However, comparison of the surfaces using a finer scale shows differences in height in excess of 500 m equatorward of 30° latitude

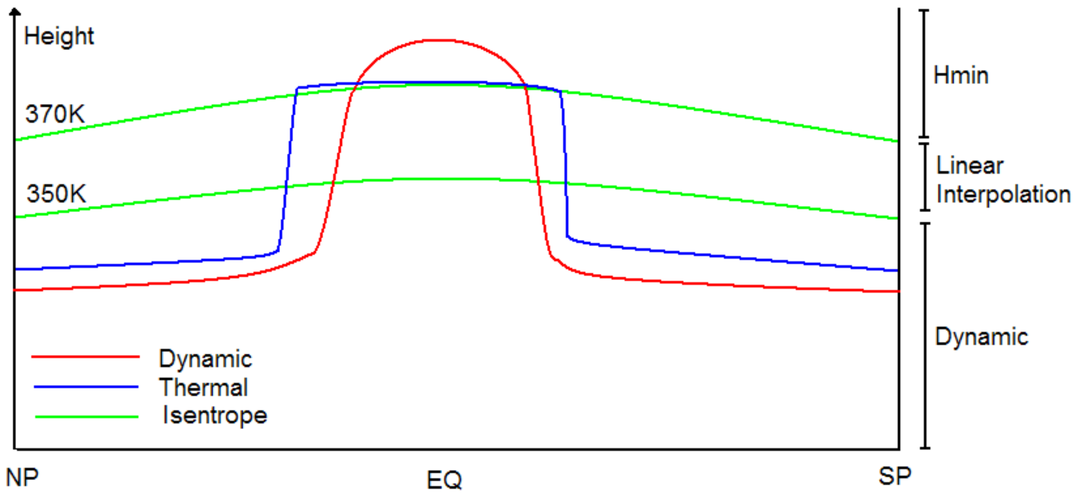


Figure 3.10: *Schematic showing cross sections of the dynamic and thermal tropopauses, and the 350 and 370K isentropes. The right hand axis indicates the composition of the blended tropopause.*

(see Figure 3.11 for the 1 August 2002 example). Over South America and the Indian Ocean the difference is in excess of 1250 m. It can be seen in Figure 3.11 that the blended definition is generally only different from the dynamic definition over a very narrow range of latitudes. This range is, however, very variable. The typical range of latitudes over which the blended definition is applied can be seen in Figure 4.2.

3.5 Sensitivity to ERA data set

It has been shown in previous studies that the tropopause is sensitive to the data set from which it is produced, and the resolution of that data set (e.g. Hoerling et al., 1991). Discussion of the tropopause in this Chapter so far has focused on the tropopause calculated using the ERA-40 dataset. However, the improved forecasting system used in ERA-Interim, along with an increased vertical resolution in both model level data and gridded data on pressure levels, is likely to result in some changes to the diagnosed tropopause.

It is expected that the change from ERA-40 to ERA-Interim will have the largest impact

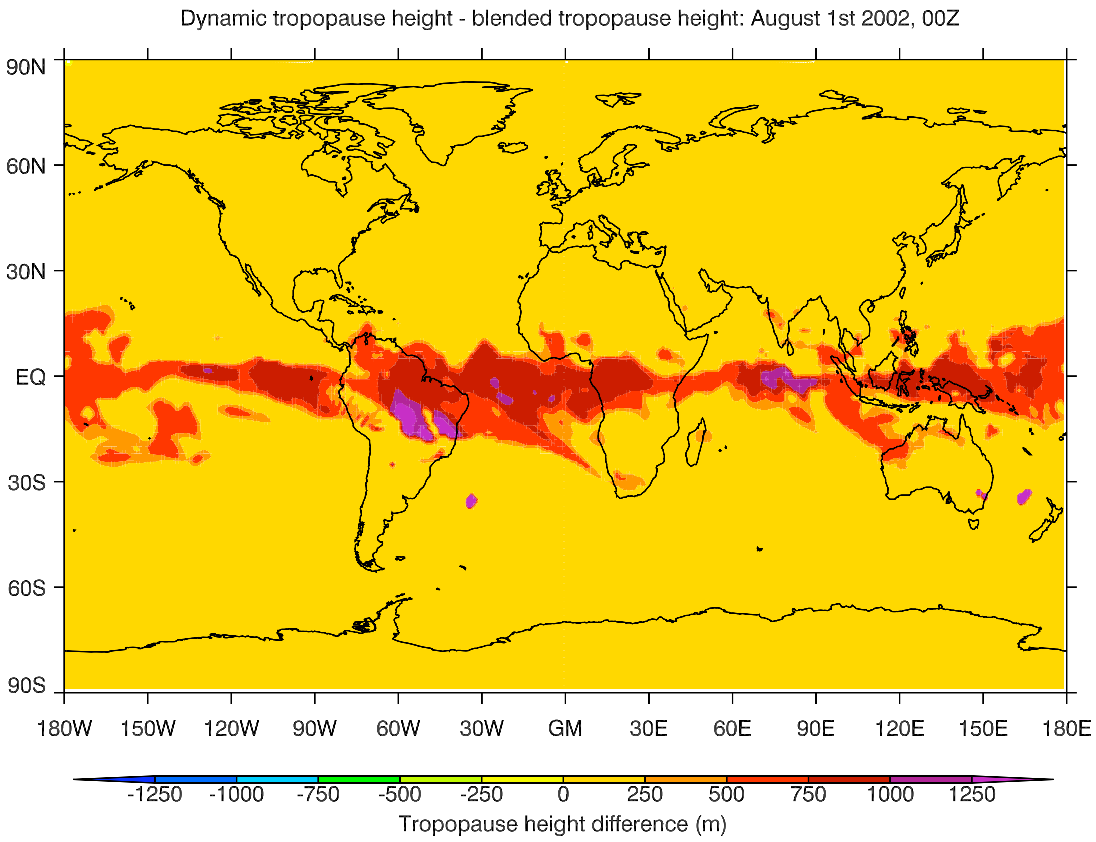


Figure 3.11: *The difference (in metres) between the dynamic and blended tropopauses on 1 August 2002, 0Z.*

on the thermal tropopause. The vertical interpolation, and the 2 km criterion, make the thermal tropopause particularly sensitive to changes in the vertical resolution of the data set from which it is calculated. There are 37 isobaric levels in ERA-Interim gridded data, compared to 23 in ERA-40, and a large number of these extra levels are included at typical tropopause pressures. If a standard atmosphere is assumed, the distance between pressure levels near 200 hPa decreases from 50 hPa (∼1500 m) in ERA-40 to 25 hPa (∼700 m) in ERA-Interim. This makes it more likely that the tropopause will be reliably identified when ERA-Interim data is used. The increased number of model levels used in ERA-Interim (91, compared to 60 in ERA-40) means that there will be a similar impact on the dynamic tropopause. However, it will be smaller in magnitude as the number of levels is large in both data sets, and the dynamic tropopause is not tied to a depth criterion.

In addition to the differences in data resolution, there are substantial differences between the forecasting systems used in ERA-40 and ERA-Interim (see Chapter 2). Changes likely to affect the tropopause surface are:

(i) The change from 3-D to 4-D Var (Andersson and Thépaut, 2008), which has resulted in systematically improved forecast performance, particularly in the Southern Hemisphere, where observations are often sparse.

(ii) Improvements in the treatment of humidity in ERA-Interim compared to ERA-40 have significantly reduced the overestimation of precipitation discussed in Hagemann et al. (2005), particularly over the tropical oceans, and the associated latent heating and circulation changes (Simmons et al., 2007).

(iii) Improvements in the representation of the stratosphere have resulted in a more realistic age of air in the stratosphere in ERA-Interim, and also increased agreement with observations of stratospheric humidity and temperature. However, where the stratosphere in ERA-40 has a significant cold bias compared to other reanalyses, there is a positive bias in ERA-Interim (Simmons et al., 2007).

(iv) The use of bias correction in ERA-Interim when assimilating satellite radiance

> data removes many of the conflicting bias corrections found in ERA-40, improving the fit to "conventional data" (Simmons et al., 2007).

For the consideration of the Antarctic tropopause, it is also worth noting that vertical temperature profiles in the polar stratosphere from ERA-40 often contained a "spurious oscillatory structure" (Simmons et al., 2007). This structure is reduced in amplitude in ERA-Interim.

The thermal tropopause calculated from ERA-40 and ERA-Interim data, and the difference between the two surfaces, is shown in Figures 3.12 and 3.13 for two case studies of 1 January 2002 and 1 August 2002 respectively. It can be seen in Figure 3.12 that the two surfaces are generally within 500 m of each other, which is a third of the typical distance between pressure levels in ERA-40 at typical tropopause altitudes. There are larger differences in the high latitudes and the sub-tropics. The change in the Northern Hemisphere high latitudes is likely to be due mostly to the increased vertical resolution of ERA-Interim compared to ERA-40, making it easier to identify the thermal tropopause in the polar night.

Differences between the tropopause surfaces in the Southern Hemisphere (see Figure 3.12) are likely to be due to a combination of the increased resolution and improved forecasting system in ERA-Interim. However, the extension of large differences into the extratropics suggests that it is changes to the IFS that are the dominant cause of the differences.

The differences between the two thermal tropopause surfaces in the sub-tropics on 1 January 2002 cannot necessarily be attributed to an improvement of the reliability of the tropopause with the use of ERA-Interim. Such large differences could easily be caused by a slight latitudinal shift in the position of the large gradients in the tropopause typical of this region, and it is difficult to attribute such a shift to either a change in the forecasting system, or the resolution of the data.

As in the January case study, the thermal tropopauses calculated using ERA-40 and ERA-Interim data for 1 August 2002, 0Z, are generally within 500 m of each other (Fig-

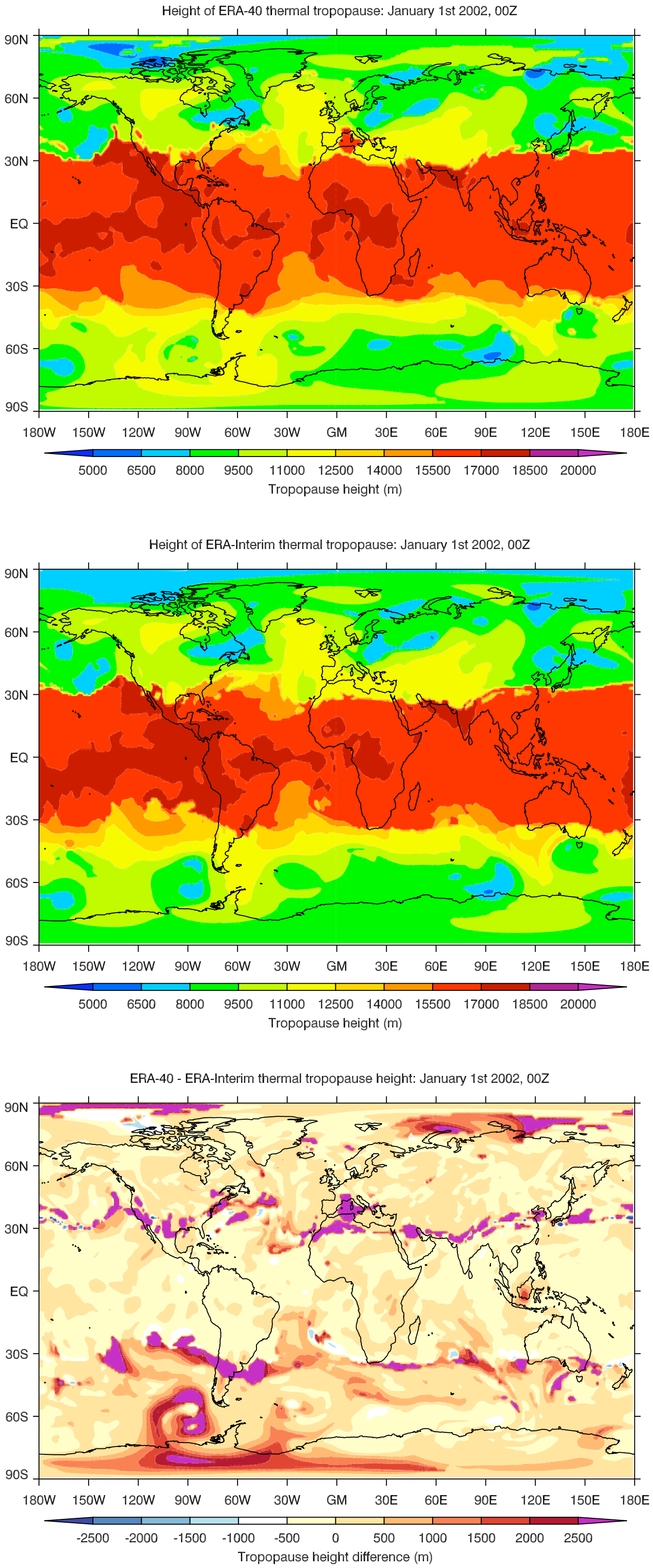
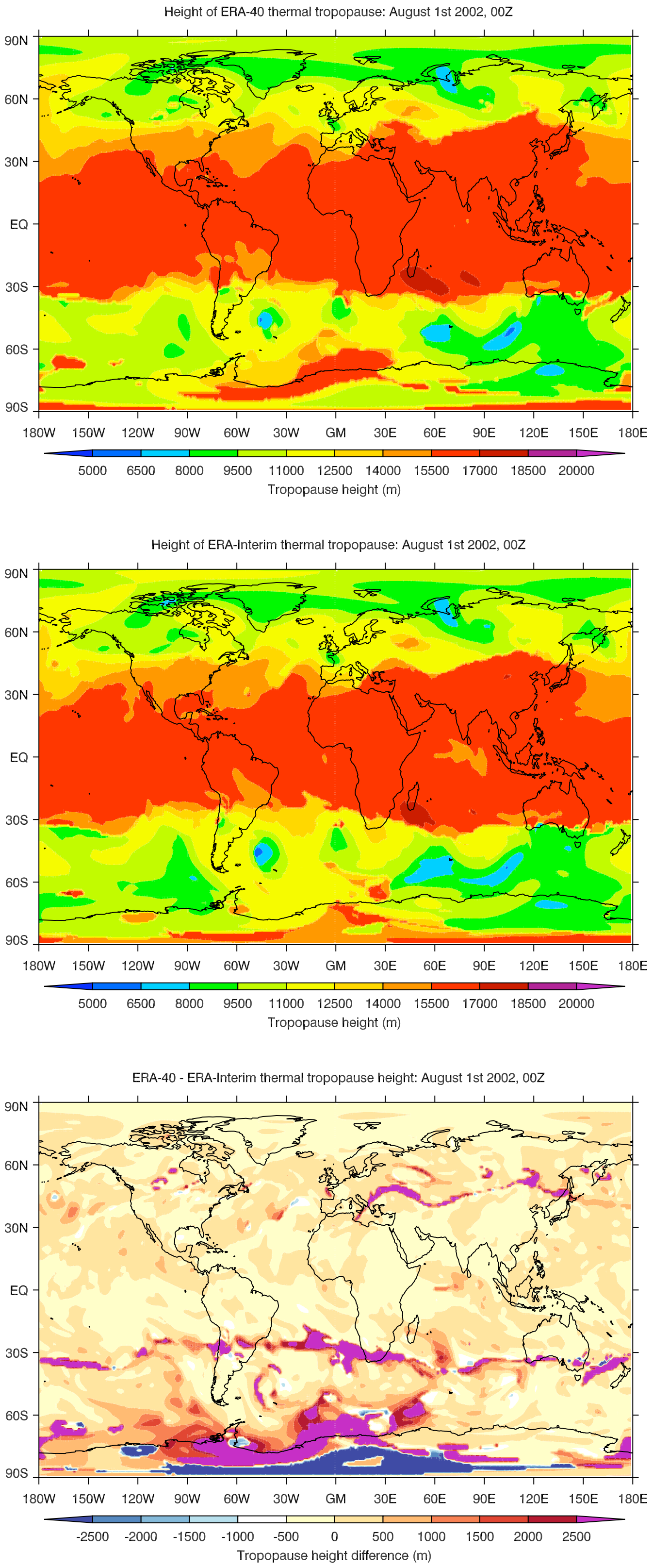


Figure 3.12: *Height (in metres) of the thermal tropopause on 1 January 2002, 0Z, calculated from ERA-40 data (top) and ERA-Interim data (middle). The difference between the two surfaces (in metres) is shown in the bottom panel.*



Height of ERA-40 thermal tropopause: August 1st 2002, 00Z

Height of ERA-Interim thermal tropopause: August 1st 2002, 00Z

ERA-40 - ERA-Interim thermal tropopause height: August 1st 2002, 00Z

Figure 3.13: *Height (in metres) of the thermal tropopause on 1 August 2002, 0Z, calculated from ERA-40 data (top) and ERA-Interim data (middle). The difference between the two surfaces (in metres) is shown in the bottom panel.*

ure 3.13). However, the differences between the two surfaces are larger in the Southern Hemisphere compared to the January example. This is to be expected as it is in the Austral winter months that the thermal tropopause is particularly difficult to identify. It can be seen from comparison of the ERA-40 and ERA-Interim tropopauses that the occurrence of an anomalously high thermal tropopause in the Southern Hemisphere high latitudes is reduced in ERA-Interim.

The blended tropopause calculated from ERA-40 and ERA-Interim data, and the difference between the two surfaces, is shown in Figures 3.14 and 3.15 for 1 January 2002 and 1 August 2002 respectively. As expected, the differences between the ERA-40 and ERA-Interim blended tropopauses are smaller than those for the thermal tropopause. It can be seen in both Figure 3.14 and 3.15 that the large differences between the ERA-40 and ERA-Interim surfaces are more evenly distributed across the globe than in the thermal tropopause cases. However, there is still a suggestion of larger differences in the Southern Hemisphere high latitudes and sub-tropics compared to other regions.

The only clear pattern of differences between the blended tropopause surfaces calculated from ERA-40 and ERA-Interim data occurs on 1 January (Figure 3.14). Here, the altitude of the tropopause in the polar regions is lower when ERA-40 is used.

Although differences can be seen between ERA-40 and ERA-Interim tropopauses using case studies, the comparison of 6 hourly snapshots cannot necessarily provide a reliable indication of the sensitivity of the tropopause to the data set used in its calculation. This sensitivity is better examined using a comparison of tropopause climatologies, which will reveal where there are consistent differences between the surfaces. The differences between ERA-40 and ERA-Interim tropopause climatologies are discussed in Section 4.4.

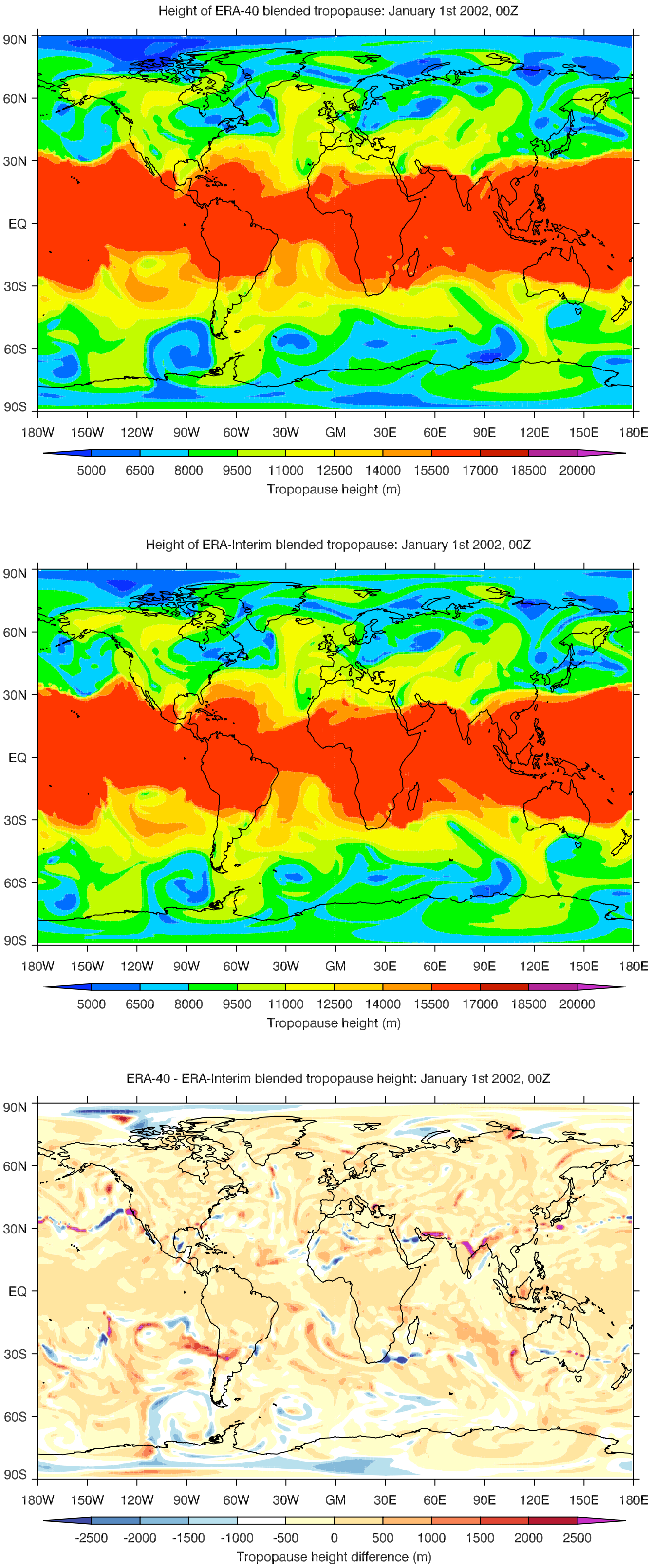


Figure 3.14: *Height (in metres) of the blended tropopause on 1 January 2002, 0Z, calculated from ERA-40 data (top) and ERA-Interim data (middle). The difference between the two surfaces (in metres) is shown in the bottom panel.*

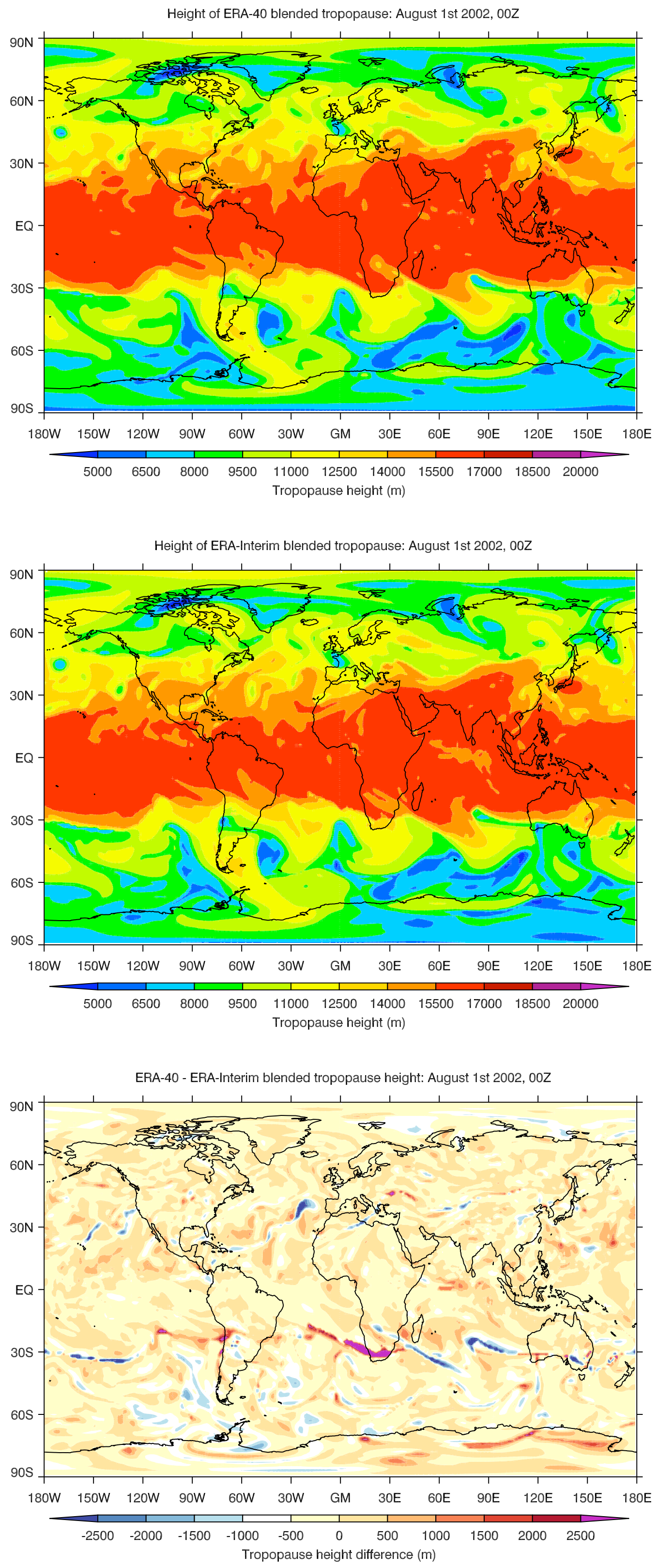


Figure 3.15: *Height (in metres) of the blended tropopause on 1 August 2002, 0Z, calculated from ERA-40 data (top) and ERA-Interim data (middle). The difference between the two surfaces (in metres) is shown in the bottom panel.*

3.6 Summary

There are several ways to define the tropopause, based on different atmospheric properties. Although some definitions are more commonly used than others, there is no 'correct' definition, as the choice is determined by the nature of an individual study. The most commonly used definitions are the WMO (1957) thermal tropopause, which is identified by changes in the temperature lapse rate, and a dynamic tropopause that is placed at a critical value of PV, often 2 PVU.

Neither the thermal or dynamic tropopause definitions can reliably be used globally. The thermal tropopause can be difficult to define in regions where there is little vertical mixing. Surfaces of constant PV become almost vertical near the equator, so the dynamic tropopause is capped using the 380 K surface. These issues make a dynamic tropopause a more reliable choice in the extratropics, and the thermal tropopause a better choice in the tropics. The two definitions are sometimes combined to give a blended tropopause definition that avoids the issues associated with the individual definitions, and also takes advantage of their strengths (e.g. improved resolution of dynamical features with strong tropopause gradients in the extratropics using the dynamic tropopause compared to the thermal).

Blended tropopauses are typically a latitude-dependent combination of the dynamic and thermal tropopauses. This can obscure the dynamic situation, making such a definition inappropriate for this work. A new flow-dependent blended tropopause definition has been introduced that uses isentropic surfaces to bound the region of interpolation between the thermal and dynamic definitions. This avoids the introduction of any latitude dependence, and uses the properties of local airmasses to determine when each component definition should be used.

Changes in the forecasting systems, and model and gridded data resolution between ERA-40 and ERA-Interim were found to have an impact on the tropopause surfaces, the thermal tropopause in particular. This sensitivity to the data set used will be investigated further using tropopause climatologies in the following Chapter.

Chapter 4

Tropopause climatologies

Global tropopause climatologies have been produced using ERA-40 and ERA-Interim data and the tropopause definitions introduced in Chapter 3. Dynamic, thermal and blended tropopause height, and potential temperature, pressure, specific humidity and tropopause-level wind components at the blended tropopause were stored for 6 hourly, monthly, seasonal and annual means, and monthly and seasonal climatologies were produced. [1]

Previously, only Hoinka (1998, 1999) has presented a global tropopause climatology, in which he uses a latitude-dependent tropopause definition. Generally tropopause climatologies are produced for specific regions, often using radiosonde data (e.g. Highwood and Hoskins, 1998).

In this Chapter key features in the ERA-Interim global climatologies will be discussed (Section 4.1), alongside more detailed regional analysis (Section 4.2). Seasonal cycles in tropopause variables are discussed in Section 4.3, and the difference between the ERA-Interim and ERA-40 climatologies will be presented in Section 4.4.

---

[1] Analysis of specific humidity is not performed here.

4.1 Global climatologies

Climatologies were produced using ERA-40 for the periods 1958-2001 (whole ERA-40), 1979-2001 (satellite era ERA-40), and January 1989- August 2002 (reanalysis overlap period), and using ERA-Interim for the periods 1989-2007 (whole ERA-Interim) and January 1989- August 2002. Although the period 1989-2002 is a short length of time for a climatology, it provides a useful point of comparison between ERA-40 and ERA-Interim. The ERA-Operational data set provides a continuation of ERA-40, but the lack of a fixed forecasting system means that it is not suitable for use in a climatology, particularly when trends are to be calculated.

For each time period, monthly, seasonal, and annual climatologies have been produced. An extended winter climatology, from December to March, was also calculated so that all months with an active Northern Hemisphere vortex could be considered together. In Chapter 3, the difference between ERA-Interim and ERA-40 blended tropopauses was shown generally to be small, but with substantial differences in some regions, particularly in the Southern Hemisphere high latitudes and the tropics. With this in mind, only the climatologies based on the more recent ERA-Interim will be presented in detail. The difference between ERA-40 and ERA-Interim for the overlap period is discussed in Section 4.4.

The zonal-mean annual-mean climatologies of blended, dynamic and thermal tropopause height are shown in Figure 4.1. The annual climatologies of the 350 and 370 K isentropic surfaces are also shown. In the extratropics the dynamic tropopause is always used in the blended definition, and is typically 500 to 1000 m lower than the thermal tropopause. The difference between the blended and thermal tropopauses is largest near the South Pole, where the anomalously high thermal tropopause that can occur in polar night, where there is no real lapse rate transition, biases the annual climatology. Equatorward of 10° latitude the dynamic tropopause is $\sim$100 m higher than the thermal and blended tropopauses. It can be seen in Figure 4.1 that the blended tropopause definition is a small but significant modification to the dynamic tropopause. The difference between the dynamic and blended tropopause climatologies is not as large as the differ-

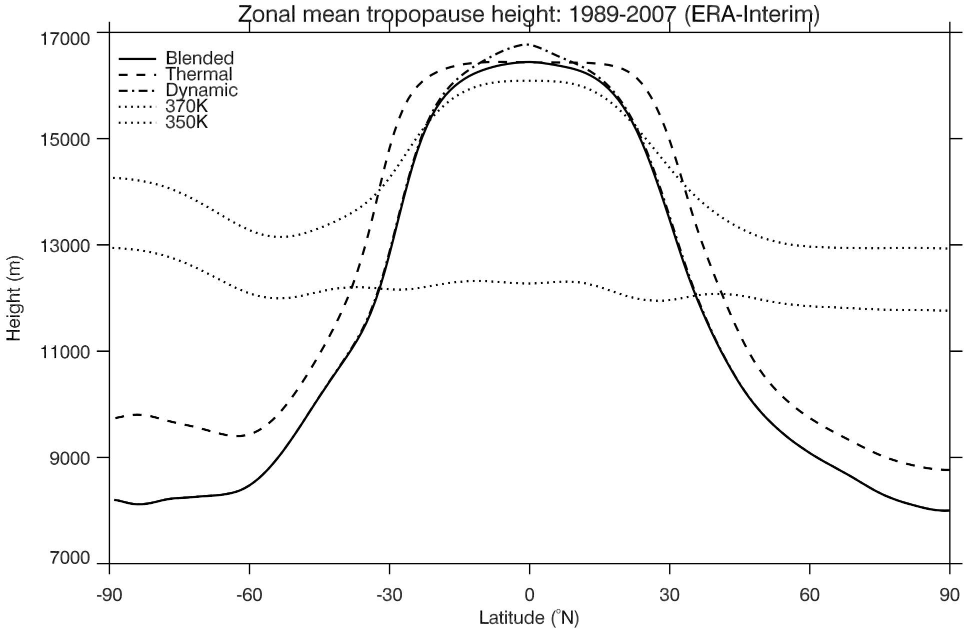


Figure 4.1: *Zonal-mean annual-mean tropopause height above mean sea level from the ERA-Interim (1989-2007) climatology using the blended (solid line), thermal (dashed line) and dynamic (dash-dotted line) tropopause definitions. The position of the zonal-mean annual-mean 350 and 370 K potential temperature surfaces are indicated by the dotted lines.*

ence between the surfaces in the tropics for some six hour snapshots, but the difference is consistent enough to have a clear impact on the climatology.

The average number of days per year in the climatology when the blended tropopause is different to the dynamic tropopause is shown in Figure 4.2. It can be seen in Figure 4.2 that the blended tropopause definition is used to some degree equatorward of 45° latitude. However, even in a 19-year climatology, there is still a clear deviation from a purely zonal structure in the sub-tropics (Figure 4.2), demonstrating the importance of using a tropopause definition that does not have a prescribed latitude dependence. The percentage of days when h_{min} is equal to the thermal tropopause in the blended tropopause definition is indicated by the contours in Figure 4.2. The likelihood of h_{min} being equal to the height of the thermal tropopause increases with proximity to the equator. This makes it unlikely that the thermal tropopause will be used for interpolation in the sub-tropics where it can be discontinuous, and also that the dynamic tropopause will be used in the region where it is capped at 380 K.

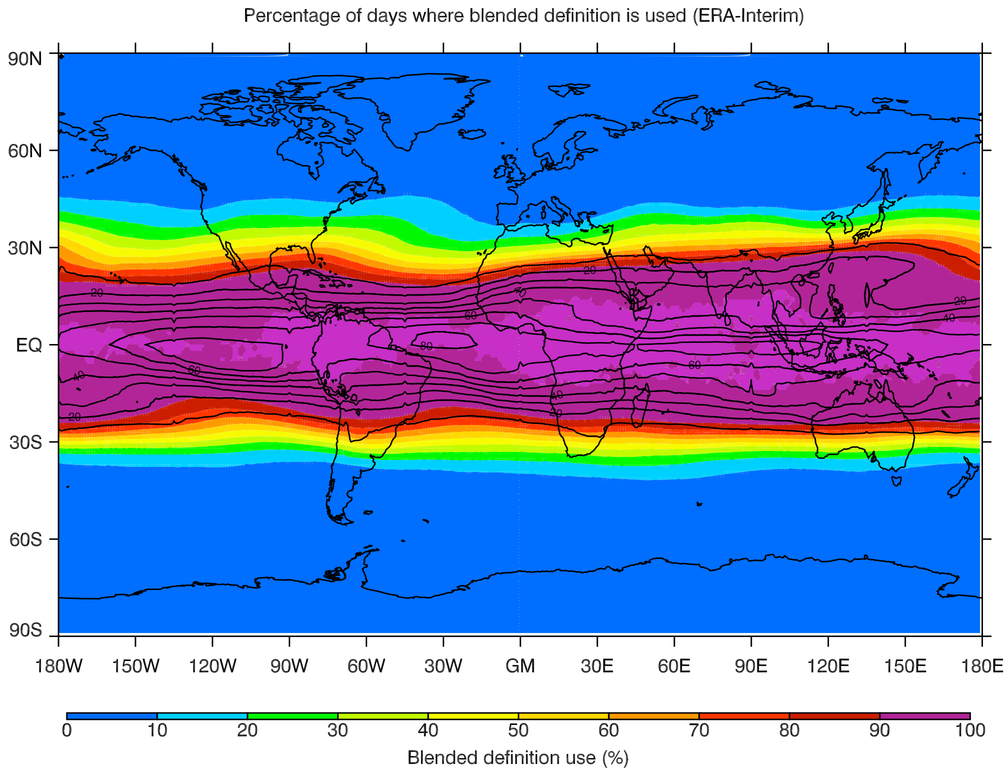


Figure 4.2: *The percentage of days in the ERA-Interim (1989-2007) climatology where the blended tropopause definition is used, instead of the dynamic tropopause definition. Line contours show the percentage of days when* h_{min} *is the height of the thermal tropopause. The contour interval is 10%.*

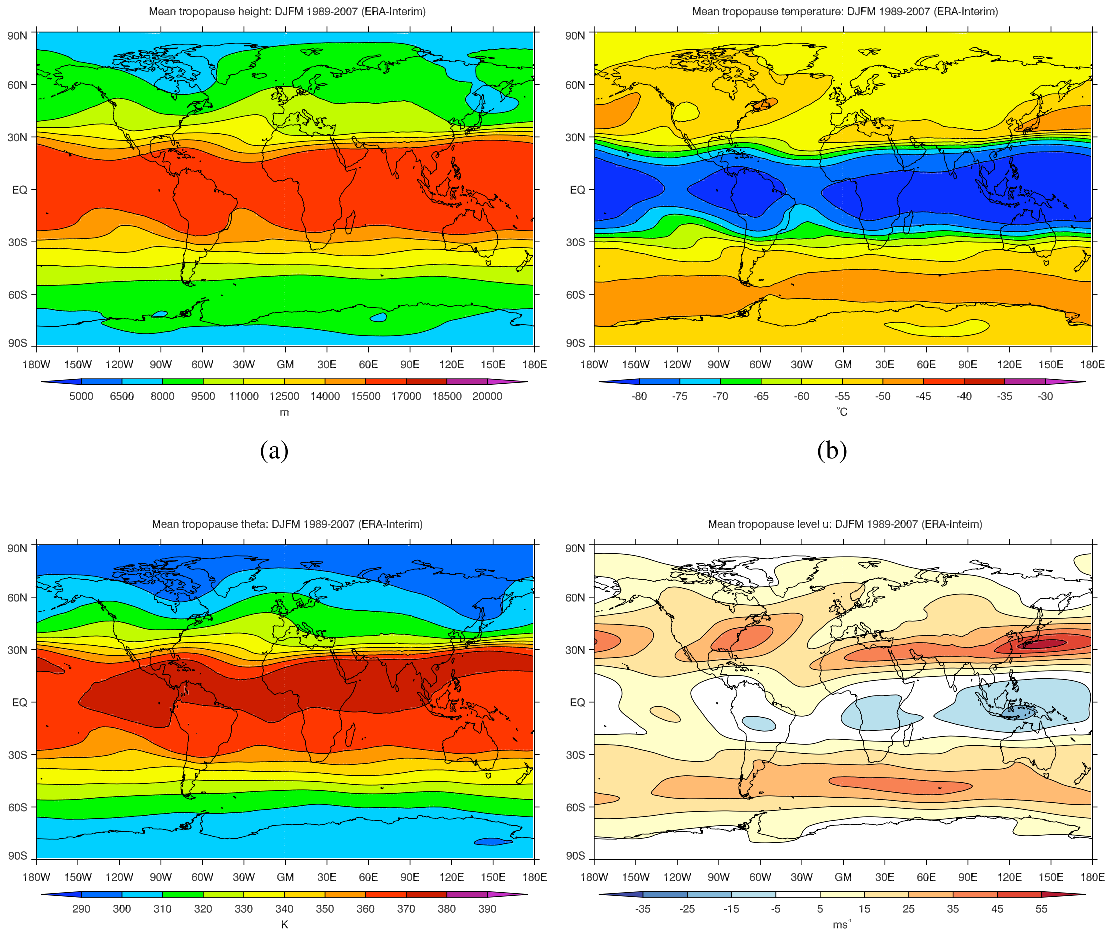
4.1.1 DJFM climatologies

In DJFM the Northern Hemisphere (NH) a polar tropopause height minimum extends over Greenland, Canada and Siberia (Figure 4.3(a)). There are large gradients in height in the sub-tropics, which are strongest in the winter hemisphere. The poleward shift of the tropopause height contours over the North Atlantic and North Pacific follow the paths of the jet streams, which are indicated by the maximum in zonal wind in Figure 4.3(d). Both the sub-tropical and mid-latitude jets are well depicted on this blended tropopause surface. In particular, the closed circuit around the Southern Hemisphere (SH), and spiral in the NH are apparent. The strongest westerly winds, of up to 59 m s^{-1}, are found in the NH over Japan and over the western half of the North Pacific, with a secondary maximum over the east coast of North America. There are easterlies over much of the tropics, which have a maximum of 16 m s^{-1} over Indonesia. The winds over the equatorial east Pacific are westerly in DJFM.

The minimum mean tropopause temperature of -86°C occurs over Indonesia in DJFM (Figure 4.3(b)), which corresponds to the position of the height maximum (16800 m). In DJFM the tropopause in the SH is warmer than in the NH. The temperature maxima in the extratropics correspond to the position of the poleward edge of the jet maxima (Figure 4.3(d)). Outside the tropics there are secondary temperature minima over Siberia and the eastern half of Antarctica.

Because of the 380 K cap on the dynamic tropopause, the blended tropopause potential temperature cannot exceed 380 K. In fact, the potential temperature has a maximum of just less than 380 K on the northern side of the Equator. It has a minimum 290 K over the Northern Hemisphere high latitude (Figure 4.3(c)). Large gradients in potential temperature are co-located with the large height and temperature gradients, and zonal wind maximum, over Japan. Globally, contours of potential temperature closely follow contours of height.

Zonal-mean sections from the ERA-Interim climatology are shown in Figure 4.4. In DJFM zonal-mean tropopause height is $\sim$8000 m at the poles (Figure 4.4(a)). In the NH,



(a) (b)

(c) (d)

Figure 4.3: *DJFM climatologies at the blended tropopause from ERA-Interim (1989-2007) (a): Tropopause height (m) (b): Tropopause temperature (°C) (c): Tropopause potential temperature (K) (d): Tropopause level zonal wind (m s^{-1})*

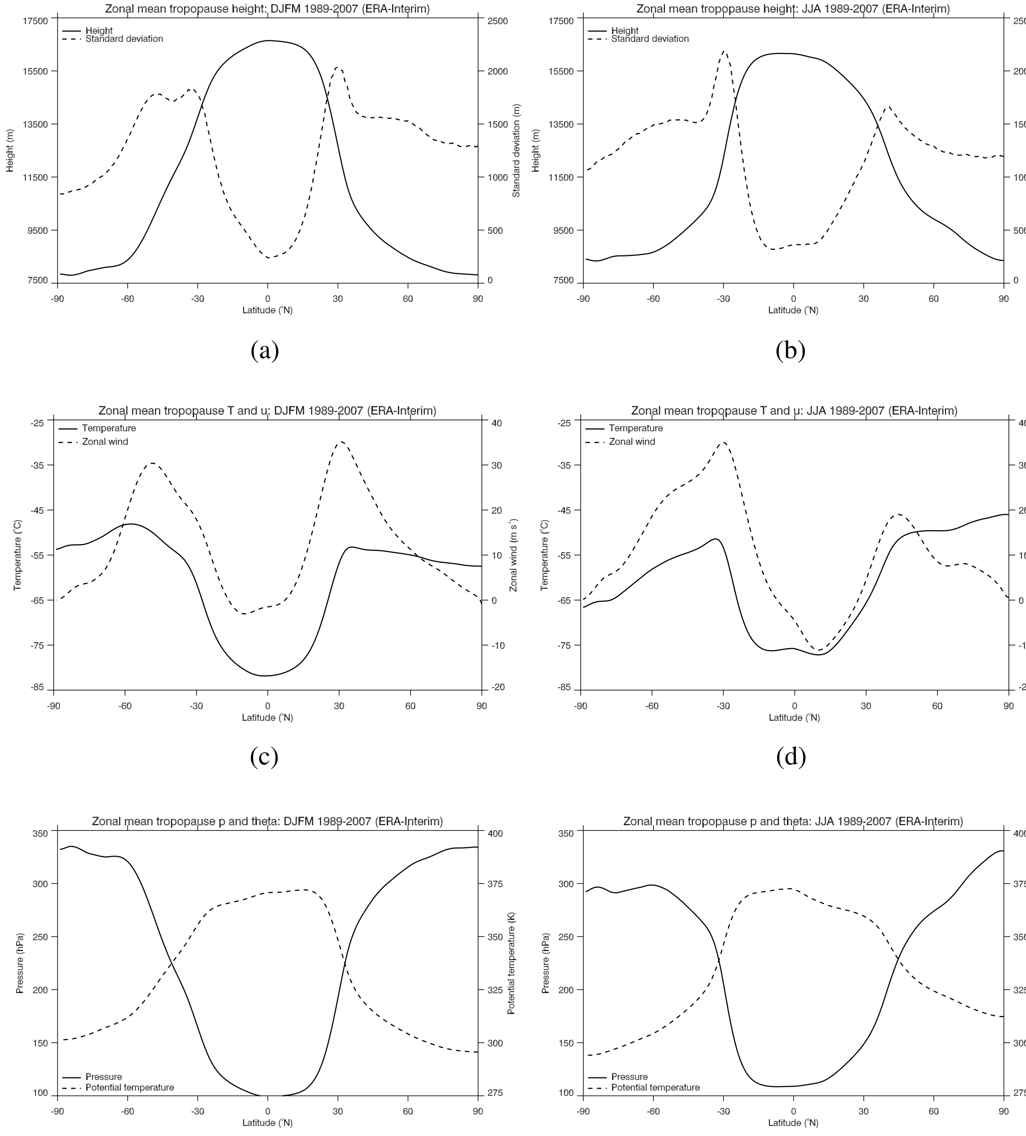
tropopause height increases gradually with distance from the poles in the extratropics, before increasing rapidly near 30° latitude (Figure 4.4(a)). In the SH there is a more pronounced transition near 60° latitude, but beyond this the height gradient is more uniform. The tropopause height has a zonal-mean of about 16500 m in the tropics, with a maximum near the Equator. The maxima in the zonal-mean six-hourly standard deviation of tropopause height (2000 m for the NH and 1750 m for the SH) are associated with variability of the steep sub-tropical transition between the tropical and mid-latitude tropopause heights (Figure 4.4(a)). There are also relatively large standard deviations in tropopause height associated with the upper level jets at 50 - 60° latitude. Standard deviation is larger in the SH extratropics, but it is smaller over the southern pole. The standard deviation of tropopause height has a deep minimum in the equatorial region (Figure 4.4(a)).

The DJFM zonal-mean tropopause temperature has a minimum of -82°C in the region of the high equatorial tropopause (Figure 4.4(c)), corresponding to the location of the height maximum. Tropopause temperature increases rapidly away from the equator to reach a maximum in the sub-tropics. In the NH this occurs at 30° latitude, which coincides with the maximum in zonal wind (Figure 4.4(c)). In the SH, the temperature is generally warmer than that in the NH, and the temperature maximum occurs some 10° poleward of the wind maximum. The tropopause level zonal wind has a minimum of -3 m s^{-1} in the tropics on the summer side of the Equator (Figure 4.4(b)), and maxima of 35 m s^{-1} at 30°N and 30 m s^{-1} at 45°S.

4.1.2 JJA climatologies

In JJA the global tropopause is generally higher than in DJFM (compare Figures 4.5(a) and 4.3(a)). The main tropical maximum is centred over the Indian Ocean and extends to the west Pacific, and across Africa (see Figure 4.5(a)). There is a polewards extension of the otherwise zonal structure over the Arabian peninsula, India and south-east Asia.

Like height, tropopause temperature is more zonally uniform in JJA compared to DJFM

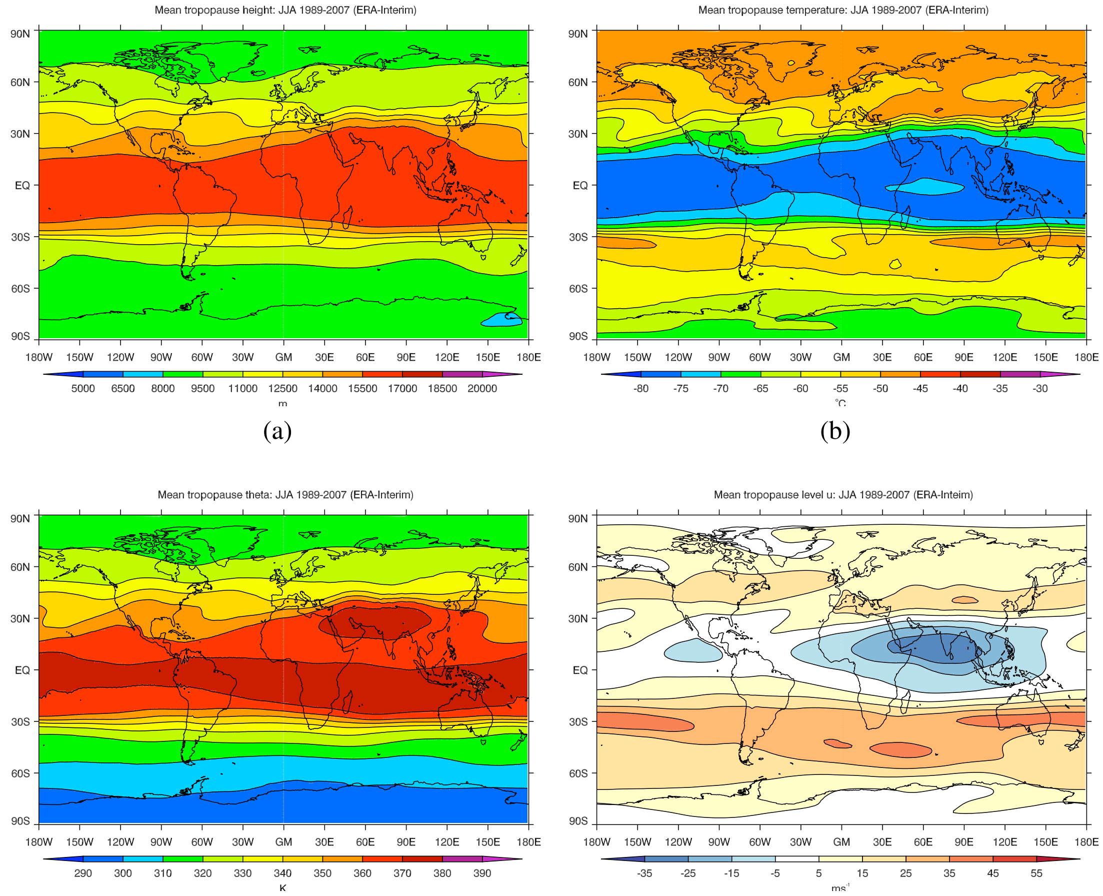


(a) (b)

(c) (d)

(e) (f)

Figure 4.4: *Zonal-mean sections for the ERA-Interim (1989-2007) seasonal climatologies. Blended tropopause height (m) and zonal-mean seasonal-mean 6-hourly standard deviation of height (m) for DJFM (a) and JJA (b), tropopause temperature (°C) and zonal-mean tropopause level zonal wind (m* s^{-1}*) for DJFM (c) and JJA (d), and tropopause pressure (hPa) and potential temperature (K) for DJFM (e) and JJA (f).*



(a) (b)

(c) (d)

Figure 4.5: *JJA climatologies at the blended tropopause from ERA-Interim (a): Tropopause height (m) (b): Tropopause temperature (°C) (c): Tropopause potential temperature (K) (d): Tropopause level zonal wind (m* s^{-1}*)*

(compare Figures 4.5(b) and 4.3(b)). The Arctic and extratropical tropopause in JJA is warmer compared to the Antarctic tropopause in DJFM, and the Antarctic tropopause in JJA is much colder than the DJFM Arctic tropopause. The tropical tropopause is warmer in JJA compared to DJFM. The tropical temperature minimum extends around the globe between ±15°latitude, with the minimum extending over the Asian subcontinent, like the height field. In the NH extratropics temperatures are still more zonally uniform than in DJFM, but they are generally warmer in the eastern hemisphere. Temperature maxima of up to -45°C are found over the NH high latitudes, and over Asia at 40°N. The temperature maxima at 40°N are located slightly to the west of maxima in zonal wind (Figure 4.5(b)).

The potential temperature maximum lies over the Asian subcontinent in JJA, and is associated with the Monsoon anticyclone. There is also a zonal band of high potential temperature values in the tropics, situated mostly to the south of the Equator. Potential temperature in the JJA SH high latitudes is comparable the that in the DJFM NH high latitudes. Potential temperature is larger in the JJA NH high latitudes compared to the DJFM SH high latitudes.

The NH jet is weaker in JJA than in DJFM, and more zonally symmetric (compare Figures 4.5(d) and 4.3(d)). In JJA the SH jet is stronger, and less zonally symmetric than in DJFM, with maximum zonal wind speeds in excess of 45 m s^{-1} over the South Pacific. The SH jet is split near 75°E. There are considerable differences in zonal wind in the tropics between DJFM and JJA. In JJA there are strong easterlies associated with the Asian monsoon over the northern Indian Ocean, and India, which are on the equatorward side of the height maximum (Figure 4.5(d)). There are also easterly winds over the eastern Pacific, in contrast to the westerlies found in DJFM. Zonal winds remain weak in the high latitudes.

Zonal means of JJA variables are shown in panels b,d, and f of Figure 4.4. Tropopause height is higher in the extratropics and high latitude in both hemispheres in JJA compared to DJFM, as was seen when comparing Figures 4.5(a) and 4.3(a). Height gradients in the SH in JJA are comparable to those in the NH in DJFM, while gradients in the NH in JJA are shallower than those in the SH in DJFM. Standard deviation of tropopause height is a maximum in the SH near 30°S in JJA, in the same location as the maximum in zonal wind. The double peaked standard deviation seen in the SH in DJFM, associated with a split jet, is not present in JJA. In the NH, the maximum standard deviation is shifted slightly poleward compared to DJFM, and is reduced in amplitude.

Zonal-mean zonal wind and temperature are shown in Figure 4.4(d). The zonal wind maximum is located near 30°S, and there is a greater difference between the SH and NH maxima compared to DJFM. Zonal-mean tropical easterlies are weaker in DJFM compared to JJA (a minimum of -3 m s^{-1} compared to -11 m s^{-1}). Tropopause temperature is warmer in JJA compared to DJFM, but the overall structure of zonal-mean temperature mirrors that seen in DJFM.

Tropopause pressure mirrors tropopause height. Pressure is higher in the tropics in JJA compared to DJFM. Tropopause pressure is comparable in the DJFM SH high latitudes compared to the JJA NH high latitudes. However, in JJA NH the pressure steadily increases with decreasing latitude, while in the DJFM SH it is roughly constant from 90 – 60°S. The pressure near the South Pole in JJA is less than that near the North Pole in DJFM, with steeper pressure gradients in the sub-tropics. In JJA the potential temperature maximum is on the south side of the Equator. Potential temperature is higher in the JJA NH compared to the DJFM SH, while potential temperatures are comparable in the JJA SH and the DJFM NH.

4.2 Regional analysis

A global analysis of the tropopause encompasses a large range of values, and features determined by a range of mechanisms. Therefore, it is useful for discussion to divide the globe into a number of regions. In this section the tropopause in the NH and SH high latitude regions, and the tropics, will be considered separately.

4.2.1 The Northern high latitudes

Monthly climatologies of the height and temperature of the ERA-Interim blended tropopause poleward of 40°N are shown in Figures 4.6 and 4.7 respectively. For most of the winter there is a minimum in height over north-east Canada and north-west Greenland (Figure 4.6). In November, a cutoff region of low tropopause height develops over the western North Pacific, which joins with the main region of low tropopause height from January until March (Figure 4.6). In April the height minimum becomes centred over the pole, and remains there until October (Figure 4.6). Tropopause height over the North Pacific is generally lower than that over the North Atlantic in the winter months (Figure 4.6).

From December to March there is a temperature minimum over Asia, 30°W of the height minimum (Figure 4.7). Temperature is less zonally uniform compared to height, especially from October to April when the eastern hemisphere is relatively cold compared to the western hemisphere (Figure 4.7). The storm track regions over the North Pacific and North Atlantic are warm compared to their surroundings throughout the year, with the temperature maxima located on the poleward side of the zonal wind maxima. There are also relatively warm temperatures over the pole from May to September (Figure 4.7).

4.2.2 The Southern high latitudes

Monthly climatologies of the height and temperature of the ERA-Interim blended tropopause poleward of 40°S are shown in Figures 4.8 and 4.9 respectively. The minimum in tropopause height is never centred over the pole, unlike the NH where the minimum is only away from the pole from December to March. Height is generally lowest in SH autumn (Figure 4.8), in contrast to the NH where the lowest tropopause heights occur in the winter (Figure 4.6). From June to November tropopause height is zonally asymmetrical, with lower heights in the eastern hemisphere compared to the western hemisphere (Figure 4.8). The region of low height forms a horseshoe shape off the coast of Antarctica. From July to September the minimum is centred near 90°E (Figure 4.8). It

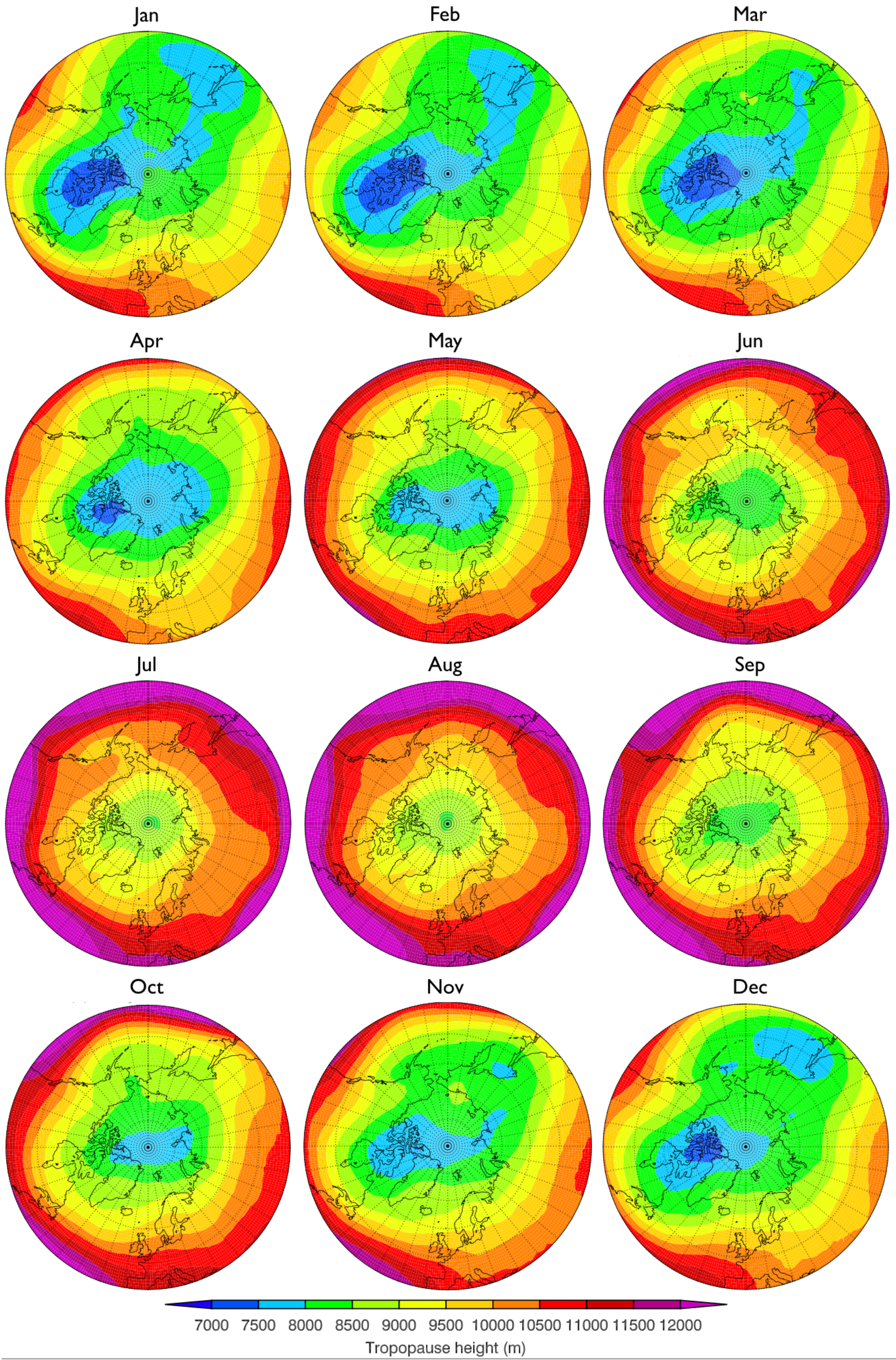


Figure 4.6: *Monthly climatologies of tropopause height (m) from 40-90°N from ERA-Interim (1989-2007) for January to December.*

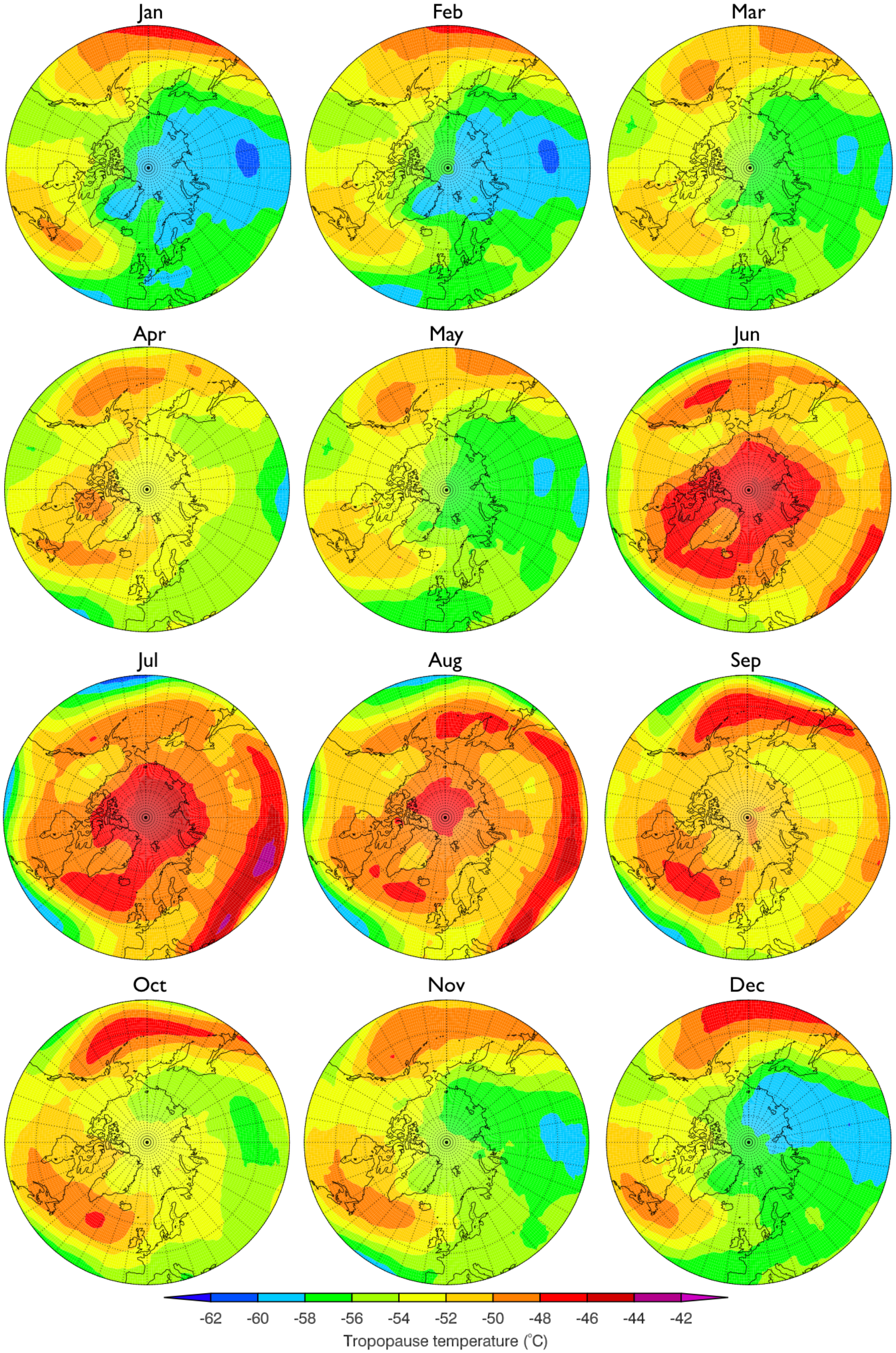


Figure 4.7: *Monthly climatologies of tropopause temperature (°C) 40-90°N from ERA-Interim (1989-2007) for January to December.*

moves over the Greenwich meridian in October to December, before fragmenting in January (Figure 4.8). The minimum then moves over the continent during the succeeding months (Figure 4.8).

There is a minimum in tropopause temperature located over the pole, or just off the pole at 90°E, for the whole year (Figure 4.9). The minimum covers a large area from June to October (Figure 4.9). During this time it remains quite circular, and centred on the pole, in contrast to its more asymmetrical appearance in other months. The area covered by the minimum expands in May and decays in November.

A region of high tropopause temperature develops in October, forming a horseshoe shape around the continent, centred on 55°S (Figure 4.9). This feature strengthens to a maximum in February, and weakens to disappear by June. There is no simple correspondence between the temperature and height of the tropopause in this region. However, as in the NH, there is a clear link between tropopause temperature and tropopause level zonal wind (not shown). Throughout the year strong westerlies form a horseshoe shape, centred on the equatorward side of the temperature anomaly at 40-50°S. The westerlies reach their maximum magnitude and zonal extent in January. The winds over Antarctica itself are weak throughout the year.

When considering the tropopause in the high Southern latitudes, it should be noted that the tropopause is not always easy to define in this region in the winter months. It has been shown in Chapter 3 of this thesis (e.g. Figures 3.4 and 3.5), and is generally accepted, that the application of the thermal tropopause definition can be problematic. However, it also appears that dynamic tropopause definition also needs to be used with care (see Figure 4.10).

In the January profile shown in Figure 4.10, roughly constant values of PV can be seen in the troposphere. The transition between tropospheric and stratospheric values of PV occurs between 1 and 10 PVU, and a steady increase in PV with pressure identifies the stratosphere. However, in the July profile (Figure 4.10) there is an constant increase in PV from 1 PVU. There is no constant PV with decreasing pressure indicative of the troposphere, and no clear transition region. It is as if there were no troposphere

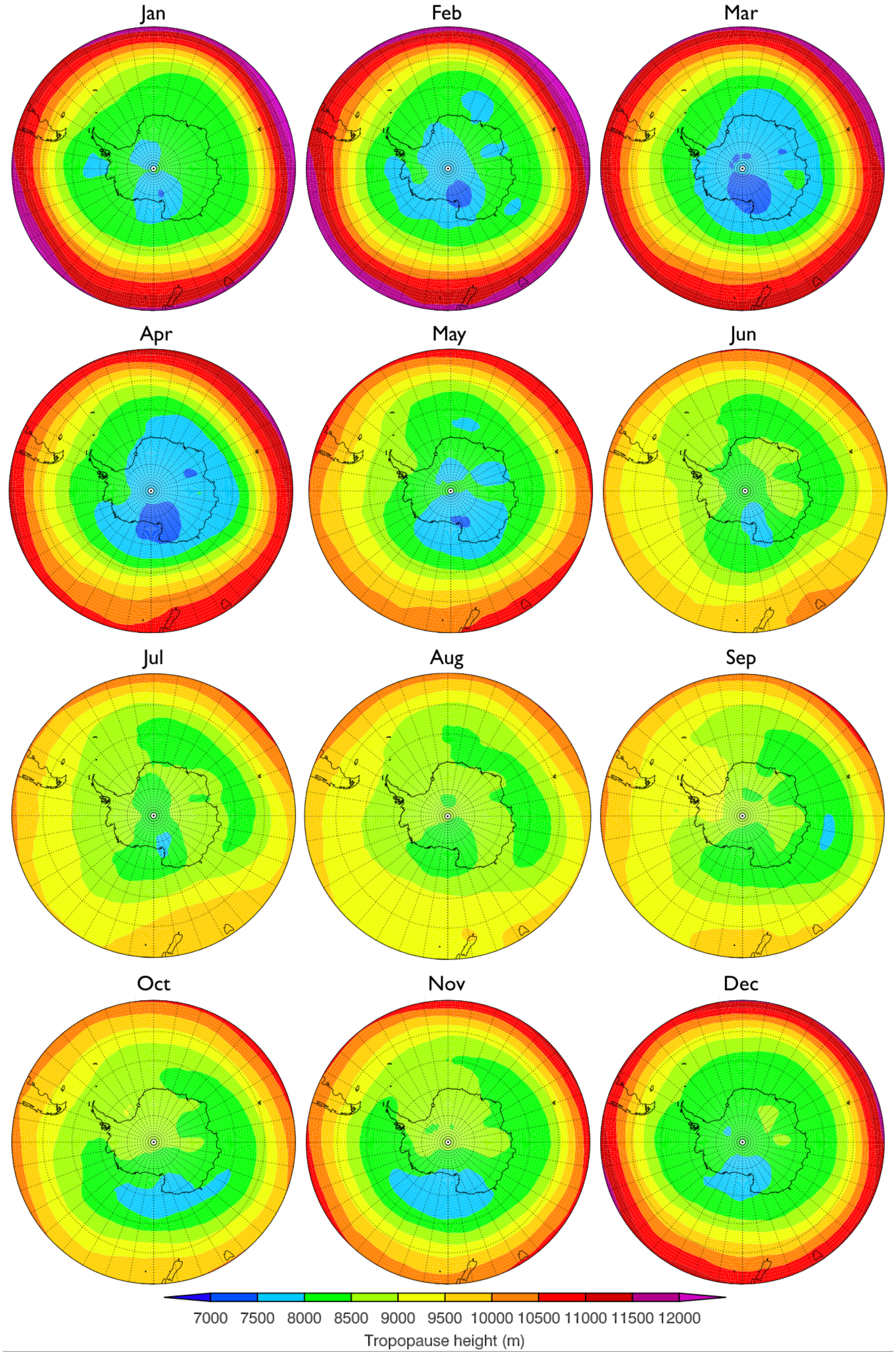


Figure 4.8: *Monthly climatologies of tropopause height (m) form 40-90°S from ERA-Interim (1989-2007) for January to December.*

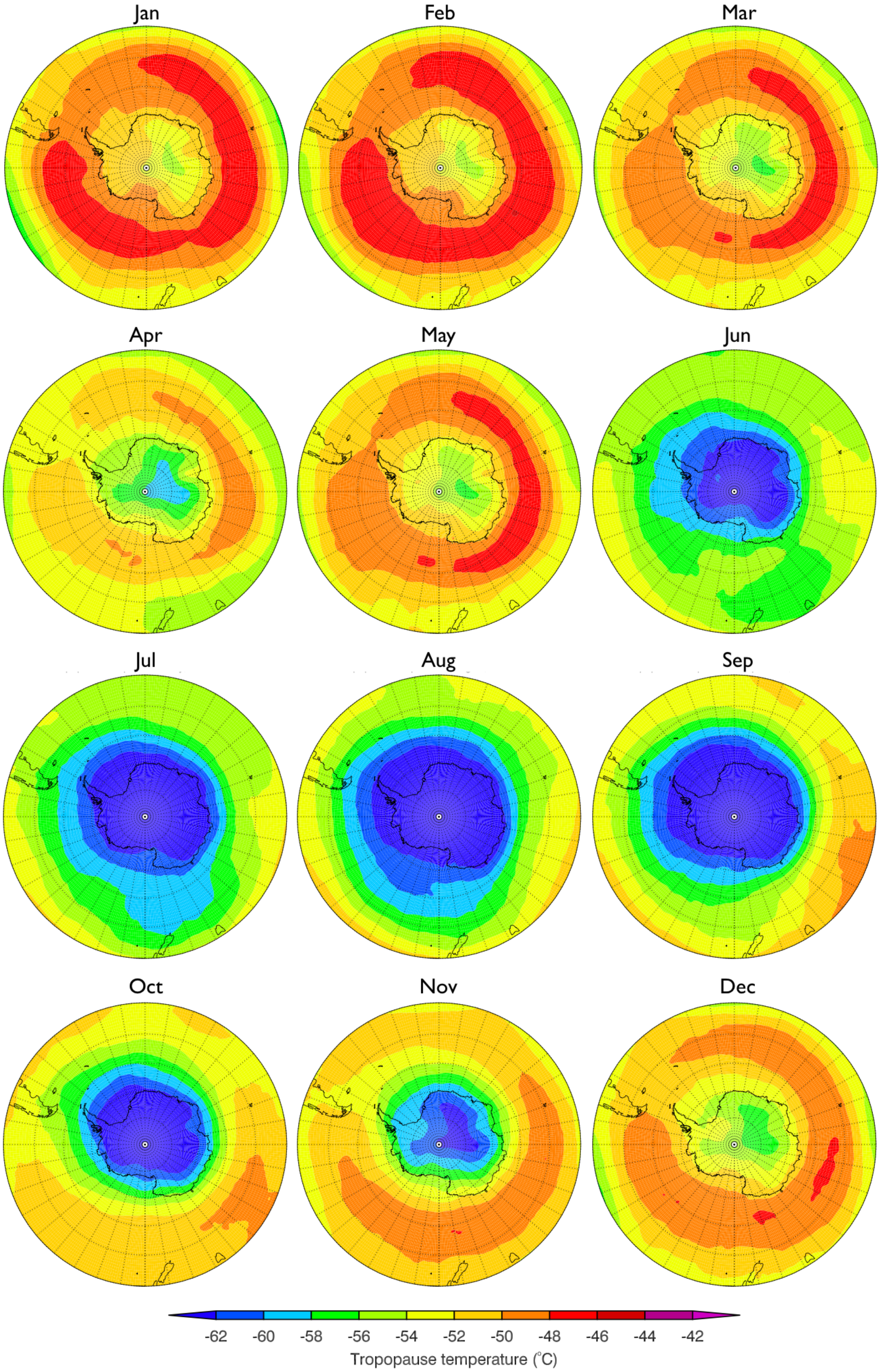


Figure 4.9: *Monthly climatologies of tropopause temperature (°C) 40-90°S from ERA-Interim (1989-2007) for January to December.*

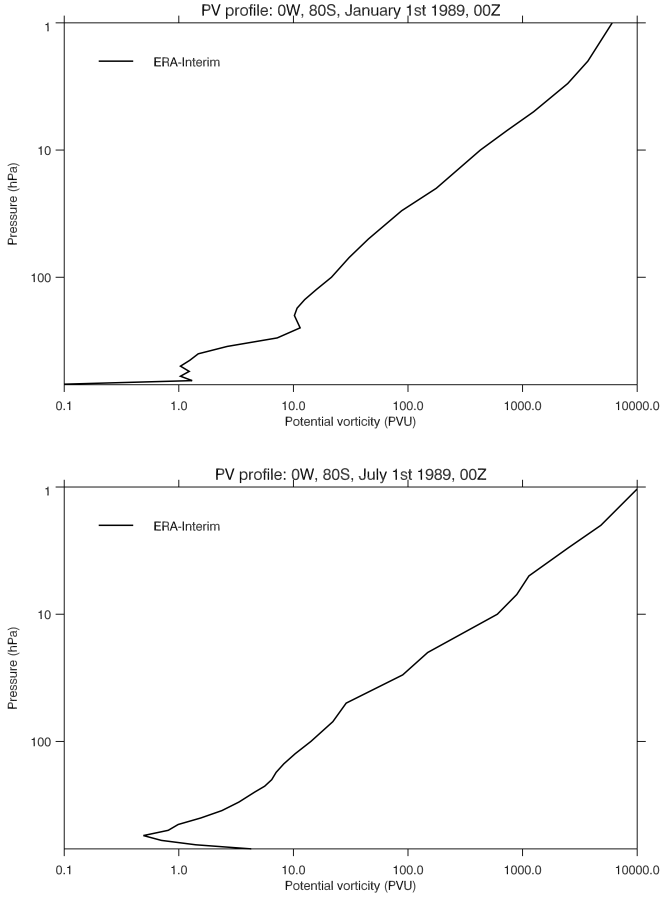


Figure 4.10: *PV profiles from 80°S, 0°W for 1 January 1989 and 1 July 1989.*

in July. Therefore, it is not necessarily appropriate to define a tropopause at all in the Antarctic winter. However, a position where |PV|=2 PVU can always be identified in these profiles, so the problems with spurious tropopause altitudes seen in the thermal case are not an issue here.

4.2.3 The Tropics

The tropical tropopause is generally high and cold, with only small variations in height and pressure. In the zonal mean, the tropical tropopause is ∼8 km higher than the extra-tropical tropopause, and the two regions are separated by large height gradients in the sub-tropics (Figure 4.1). Monthly climatologies of tropical tropopause temperature are shown in Figure 4.11.

It can be seen in Figure 4.11 that the tropical tropopause is a fairly constant feature for half of the year, from November to April (Figure 4.11). The temperature minimum is located, along with the height maximum, over the west Pacific. There are secondary temperature minima over the west coast of South America, and over equatorial Africa.

In May, the tropical tropopause warms. A region of low temperature extends northwards over the Asian subcontinent from the west Pacific. By July the temperature minimum lies over the Bay of Bengal. This is to the south-east of the height maximum, which is located over northern India. The temperature minimum remains over the Bay of Bengal until October, and returns to over the west Pacific in November. Tropical tropopause temperature decreases during this time.

From November to April, tropopause height maxima and temperature minima are co-located (height is not shown). However, when the region of high tropopause height extends northwards in May, there is only a small northwards extension in the region of cold tropopause temperatures. The tropical tropopause warms in summer, as the height decreases, with a temperature minimum remaining over south-east Asia, to the south-east of the height maximum.

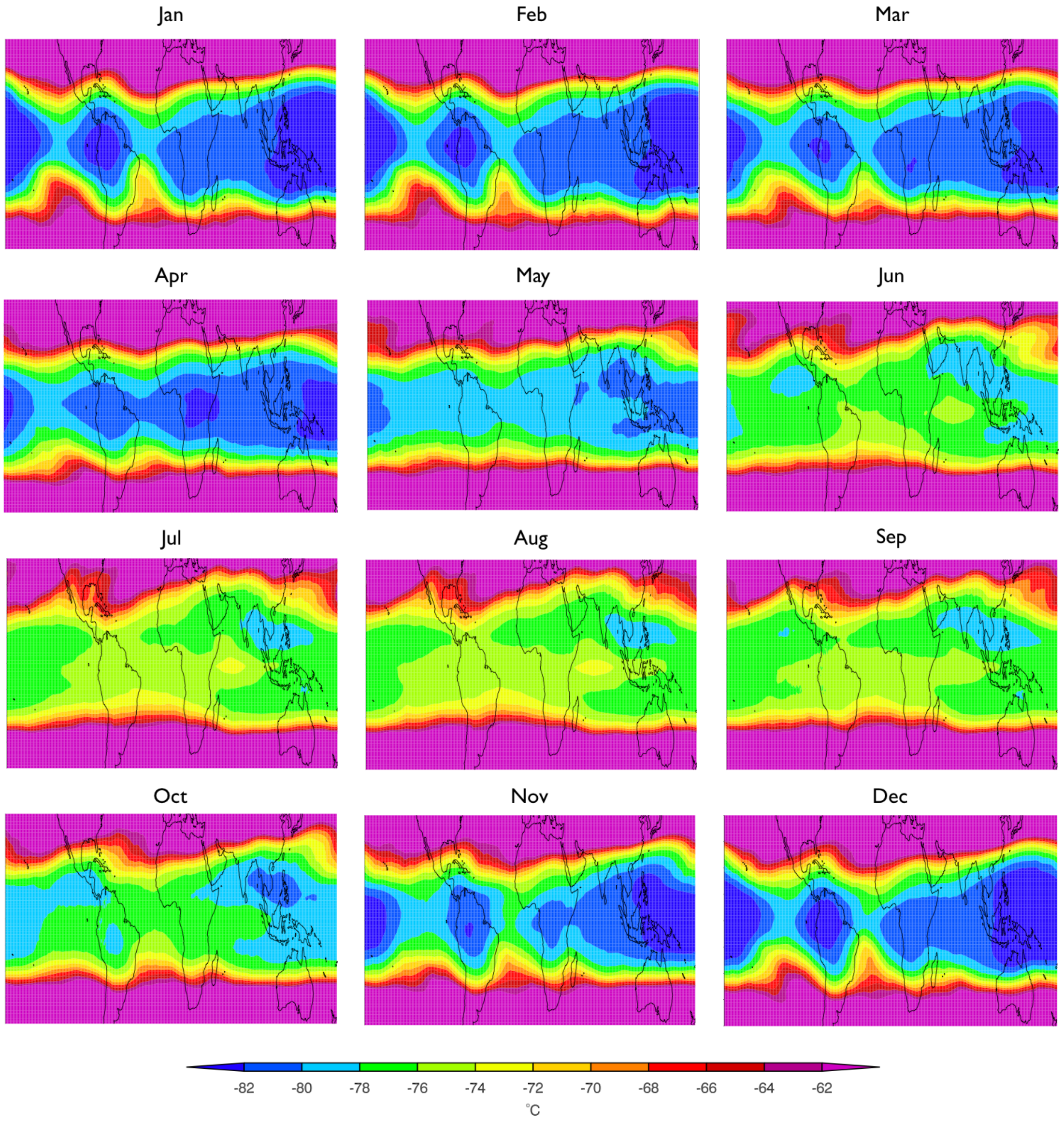


Figure 4.11: *Monthly climatologies of tropical (±40° latitude) tropopause temperature (°C) from ERA-Interim (1989-2007) for January to December.*

Tropopause level zonal wind is fairly constant in the tropics from October to March (not shown). The winds are easterly in most of the tropics, with a maximum over the Maritime continent. There are westerlies over the Pacific. In April, the strength of the easterlies decreases when the tropopause height is at its maximum. The strength of the easterlies increases through May and June, before reaching a maximum in July. The location of maximum wind speed moves north over the Indian Ocean in the summer, along with the tropopause height maximum and temperature minimum.

During El Niño sea surface temperatures (SSTs) in the east Pacific Ocean increase, causing a increase in convection in the east and central Pacific, and a decrease in the west Pacific. Increased convection causes a raised tropopause (decreased pressure), and conversely for reduced convection. This can clearly be seen in the regression of tropopause pressure onto the Niño 3.4 index shown in Figure 4.12(a). This Niño 3.4 index is the SST anomaly in the region bounded by 120°W, 170°W, 5°S, and 5°N (Trenberth, 1997).

The Pacific/North American pattern (PNA) is strongly influenced by ENSO. The positive PNA tends to be related to El Niño. The positive phase of the PNA is associated with increased tropopause height over Hawaii and the northern US and Canada, and decreased tropopause height over the southern US. This can be seen in the regression of tropopause temperature onto the Niño 3.4 index (Figure 4.12(b)). A Kelvin wave raised tropopause signature in the western hemisphere, and Rossby wave depressed tropopause signature in the eastern hemisphere are also apparent in this Figure.

4.3 Seasonal cycles

There are large annual cycles in all parameters at the tropopause in the high latitudes. If the averages poleward of 65° latitude are considered, Arctic tropopause height has a minimum of 7800 m in March, and a maximum of 9000 m in July. This differs slightly from Highwood et al. (2000), who found a maximum in Arctic tropopause pressure in April and a minimum in August. However, when the annual cycle of the thermal tropopause, as used by Highwood et al., is considered, the results are in agreement:

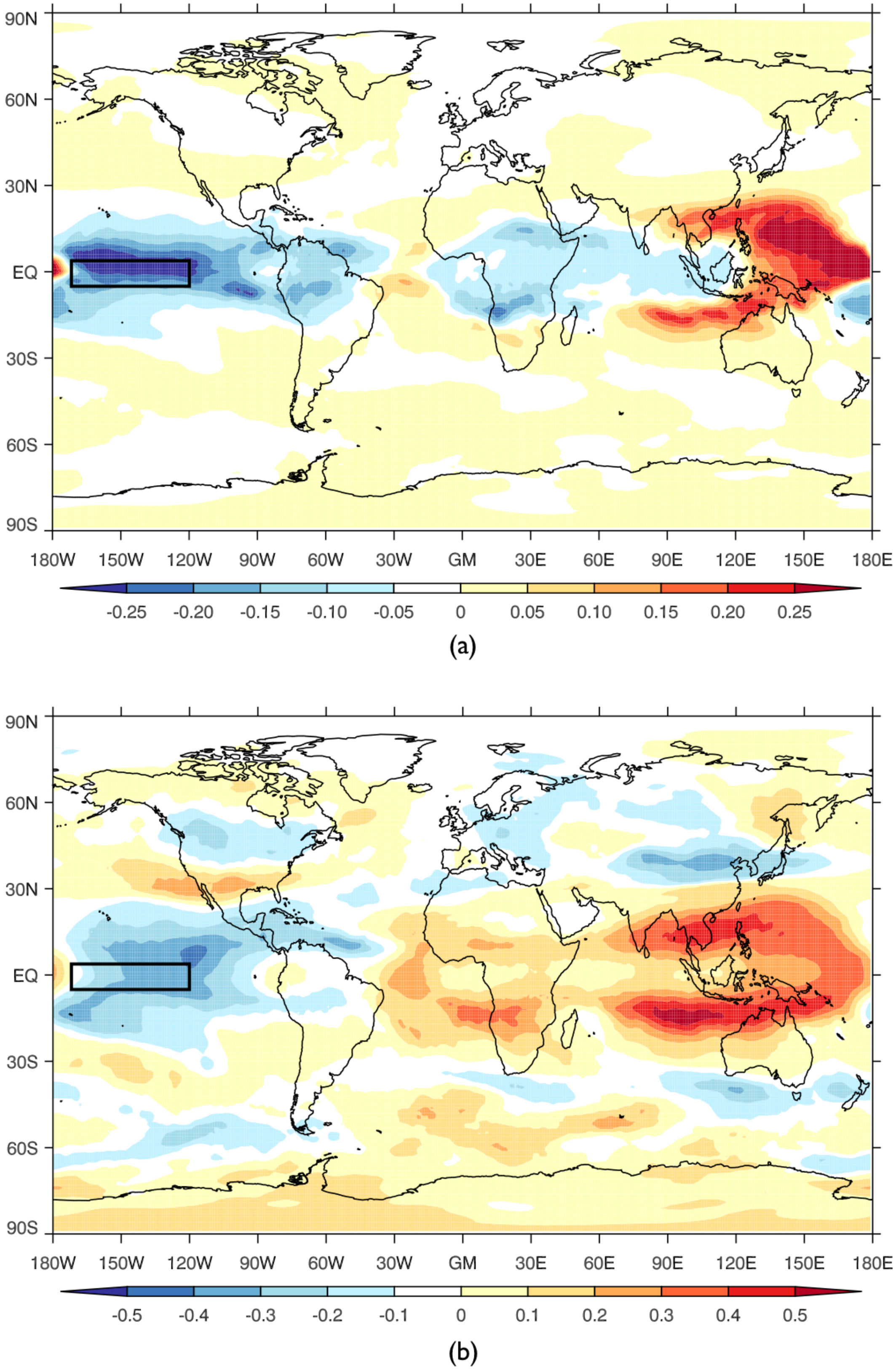


Figure 4.12: *Regression of tropopause pressure (a) and temperature (b) onto the Niño 3.4 SST index. The location of the Niño 3.4 region is indicated by the black boxes.*

Arctic temperature has a minimum in January and maximum in July.

Antarctic tropopause height has a minimum of 7700 m in March, and a maximum in September of 8600 m. Antarctic temperature and pressure both have a minimum in August and a maximum in February. Although the amplitude and magnitude of the annual cycle in tropopause height in the Arctic and Antarctic are similar, the Antarctic tropopause is colder. Temperatures range from -66°C and -51°C, compared to -58°C to -46°C in the Arctic. Such a difference in temperature is consistent with Hoinka (1999).

The Antarctic is the only region considered that does not have a height minimum in winter. The thermal tropopause definition is known to be difficult to apply in winter as the transition from a tropospheric lapse rate profile to a stratospheric profile is very gradual. Although it is always possible to identify a position where PV=-2PVU, the significance of the dynamic tropopause definition in Antarctic winter also has to be questioned as the transition from tropospheric to stratospheric potential vorticity is also gradual (Figure 4.10). As such, tropopause parameters in the Antarctic region in Austral winter should be approached with a degree of caution.

Zängl and Hoinka (2001) investigated the seasonal cycles of tropopause parameters in the polar regions. They characterised them into ‘single waves’ with one maximum, and ‘double waves’ with two distinct maxima. All the annual cycles identified here in the ERA-40 and ERA-Interim climatologies have a ‘single wave’ pattern. All the cycles in height have a maximum in NH summer, with the exception of those in the tropics.

Zängl and Hoinka (2001) identified a double wave pattern in the seasonal variation of Arctic tropopause pressure, with pressure maxima in summer and winter for most of the Arctic. There is some suggestion of this pattern in the Arctic cycles in the ERA-40 and ERA-Interim climatologies, but it is very weak. The summer height maximum is larger than the one found in winter, meaning that there is little to distinguish this pattern from the other regional cycles. Here the Arctic has been taken as the region polewards at 65°N. If a region is specified with a boundary at a higher latitude then the double wave identified by Zängl and Hoinka (2001) becomes more apparent.

The annual averages of tropopause parameters in the tropics all have clear annual cycles, although the amplitude is small compared to the cycles in the high latitudes. Tropical average (±20°latitude) height has a minimum of 15900 m in July, and a maximum of 16400 m in April. Temperature and specific humidity both have a minimum in January and a maximum in August. Tropical mean tropopause temperature ranges from -79.5°C to -75.5°C.

Area averages for the 'deep tropics' were also considered. This has been defined as the region equatorwards of 10° latitude, following Solomon et al. (2007). Deep tropical height has a minimum of 16000 m in August, and a maximum of 16650 m in April. Temperature has the same annual cycle in the deep tropics as in the tropics, but the temperatures are colder, ranging from -81.5°C to -76.0°C.

4.4 Difference between ERA-40 and ERA-Interim climatologies

The ERA-Interim and ERA-40 climatologies have been compared for the period January 1989- August 2002, where they overlap. The height differences (ERA-40 minus ERA-Interim) in DFJM and JJA are shown in Figure 4.13.

Globally, there are only small differences between the height of the ERA-40 and ERA-Interim tropopause heights. The two surfaces are generally within 500 m of each other, which is less than the difference between the thermal and dynamic tropopauses (Figure 4.1), and most ERA pressure levels. The only difference of more than 500 m in height between the two climatologies in DJFM is in the lee of the Himalayas. Here, the ERA-40 tropopause is 2 km higher than the ERA-Interim tropopause. The ERA-40 tropopause level westerlies here are also weaker than those from ERA-Interim, and the temperatures are cooler. The differences in this region are most pronounced in DJFM, and appear to be associated with large vertical, and, to a lesser extent, horizontal oscillations in the ERA-40 potential vorticity, potential temperature and zonal wind fields. The

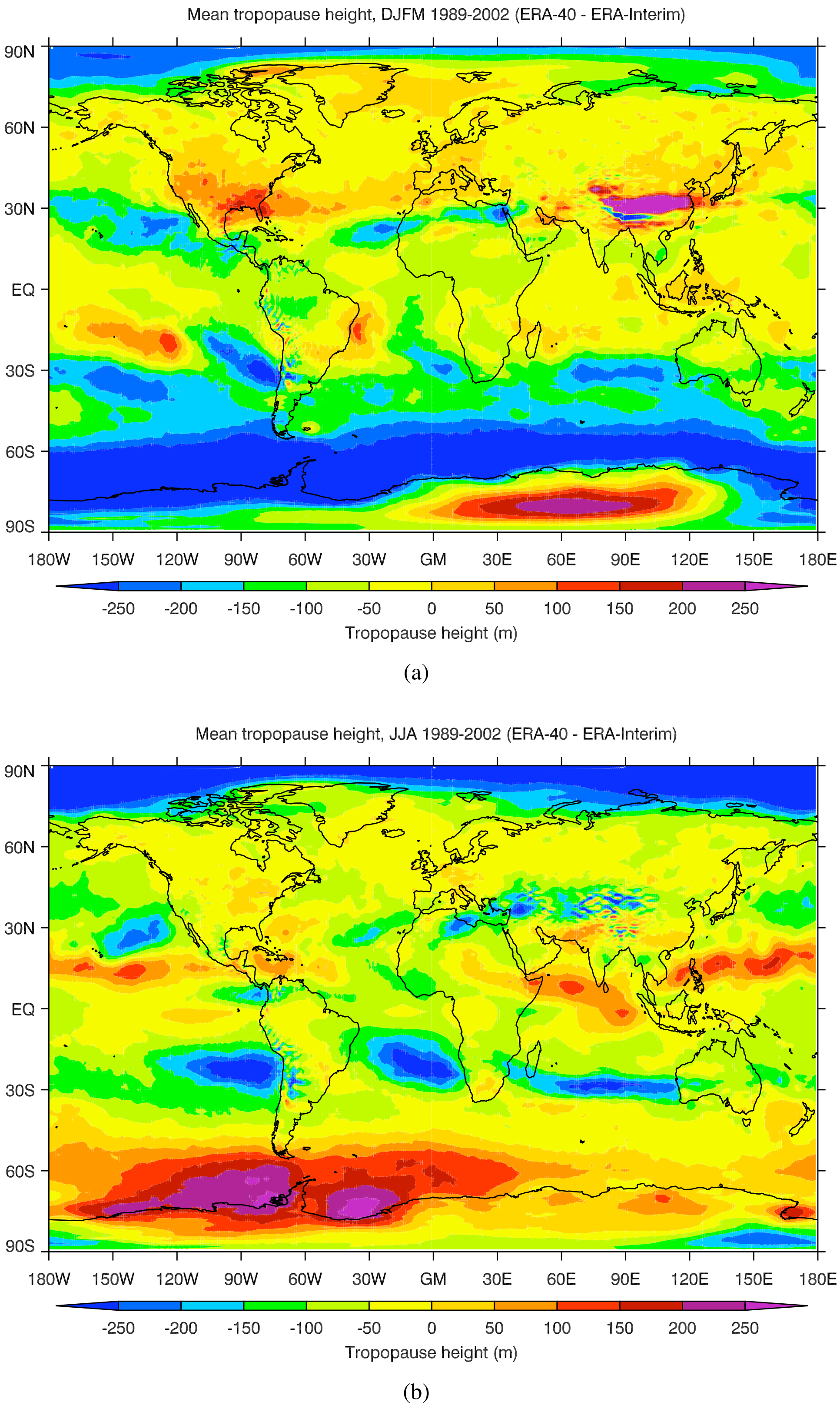


Figure 4.13: *Difference (ERA-40 - ERA-Interim) between the DJFM (a) and JJA (b) seasonal climatologies of tropopause height calculated using the ERA-40 and ERA-Interim data sets (1989-2002).*

origins of these oscillations are unknown. It is suspected that they may be the result of an overparameterisation of gravity wave drag in ERA-40.

In addition to the very large differences in the lee of the Himalayas, there are smaller differences between the tropopause heights from the two climatologies. The most notable of these are in the SH high latitudes. Here, the (DJFM) ERA-Interim tropopause is over 250 m higher than the ERA-40 tropopause around the Antarctic coast, but lower over the continent itself. This is probably the result of the change from 3-D to 4-D Var between ERA-40 and ERA-Interim, which is known to have a large impact on the SH (Andersson and Thépaut, 2008).

There are also differences in height of up to 100 m in the tropics. It is possible that this is due to the improved treatment of humidity in ERA-Interim, and the subsequent impacts on convection and latent heating. Differences between tropopause heights from the two climatologies are smaller in the NH, where there is more observational coverage and the data is less sensitive to changes in the forecasting system.

As in DJFM, tropopause heights from ERA-40 and ERA-Interim are always within 500 m of each other in JJA. Differences in height in the tropics are slightly larger in JJA compared to DJFM, but the tropopause heights from the two climatologies are within 200 m of each other there. The largest differences in the tropics are over the South Pacific and South Atlantic, where the ERA-Interim tropopause is higher, and over the Indian Ocean and the West Pacific, where ERA-40 is higher.

There are large differences between the surfaces in the SH high latitudes, but with a different structure than found in DJFM. In JJA the ERA-40 tropopause is generally higher (by up to 250 m) than the ERA-Interim tropopause over the whole region. However, it should be noted that the tropopause is not necessarily well defined in the SH high latitudes in JJA.

Tropopause temperature (not shown) is generally lower in the ERA-Interim climatology compared to the ERA-40 climatology in the the tropics and Arctic in both seasons. The ERA-40 tropopause is warmer by 1-2 K in the Arctic. In the tropics the difference can be as much as 6 K, with a maximum temperature difference over the central Pacific in JJA. The temperature difference over the Antarctic can be as large as 5 K, and changes sign throughout the year, having the opposite sign to the height difference.

Tropopause level zonal winds (not shown) are generally within 2 m s^{-1} in the two climatologies. However, in JJA the westerly winds over the Indian ocean are up to 6 m s^{-1} weaker in ERA-40 compared to ERA-Interim. There are also large differences in the lee of the Himalayas in DJFM, in the same location as the very large height differences.

4.5 Summary

A global tropopause climatology of tropopause height, pressure, temperature, potential temperature, and zonal-wind has been produced for 1989-2007 using ERA-Interim. Key features in this climatology have been discussed for the high latitudes and tropics, alongside seasonal cycles in tropopause variables.

In the NH tropopause height has a minimum over Canada and Greenland in the winter months, which moves over the pole in summer. Temperature has a minimum in the western hemisphere in winter, and maxima over the pole and near the position of the jet streams in summer. In the SH there is a height minimum and temperature maximum located on the Antarctic coast throughout the year. In Antarctic winter there is a strong temperature minimum over the pole. The tropical tropopause is a fairly constant feature throughout half of the year, with a temperature minimum over the Western Pacific. In the summer months the temperature of the tropical tropopause increases, and its structure is dominated by the Monsoon anticyclone. The regression of tropopause pressure onto the Niño 3.4 index reveals that the tropical tropopause is also influenced by ENSO.

Comparison of ERA-40 and ERA-Interim climatologies for the overlap period between the two reanalyses (1989-2002) has also been presented. It was found that the two tropopauses are generally within 500 m of each other. The largest difference was in the lee of the Himalayas in DJFM where the ERA-40 tropopause was 2 km higher than that from ERA-Interim. Generally there was more difference between the two tropopauses in the SH as there are fewer observations and hence the analysis is more susceptible to changes in the forecasting system. The ERA-40 tropical tropopause is warmer compared to ERA-Interim, with differences of up to 6°C over the central Pacific in JJA. Winds from the two reanalyses are generally within 2 m s^{-1}, but JJA westerlies over the Indian ocean are up to 6 m s^{-1} stronger in ERA-40 compared to ERA-Interim.

Chapter 5

Tropopause trends

Increases in tropopause height (decreases in tropopause pressure) are associated with a warming of the troposphere and a cooling of the stratosphere. There is a variety of factors that can influence the tropopause by altering the temperature profiles of the troposphere and stratosphere. These include increases in well-mixed greenhouse gases, which results in a warming of the troposphere, cooling of the stratosphere, and a subsequent increase in tropopause height; explosive volcanic eruptions that warm the stratosphere, causing a decrease in tropopause height; and increases in sulphate aerosols that cool the troposphere, again resulting in a decrease in tropopause height. Decreases in stratospheric ozone cause both the stratosphere and the troposphere to cool, but the stratospheric cooling has a dominant influence on the tropopause (Santer et al., 2004), causing an overall increase in tropopause height. Increases in solar irradiance cause both the troposphere and the stratosphere to warm. However, in this case it is the tropospheric warming that has the dominant influence on the tropopause, and the overall effect is an increase in tropopause height (Santer et al., 2003a).

The primary cause of observed increases in tropopause height is thought to be increases in well-mixed greenhouse gases. Santer et al. (2004) found that stratospheric cooling and tropospheric warming are robust signals of increasing atmospheric carbon dioxide, and that the pattern of increases in tropopause height is consistent with model predictions of the response to increases in carbon dioxide alone. Santer et al. (2003a) found that anthropogenic changes in ozone and well-mixed greenhouse gases accounted for 80% of the increase in tropopause height between 1979 and 1999. Sausen and Santer (2003) also found that the observed increases in global tropopause height cannot be explained by natural variability. It has been suggested that changes at the tropopause provide a sensitive indicator of climate change (e.g. Randel et al., 2000, Santer et al., 2003, 2004), as time-series of global-mean tropopause height contain much less natural variability than those of tropospheric or surface temperature. Sausen and Santer (2003) suggested that changes in tropopause height could be detected approximately 20 years in advance of changes in surface temperature.

Time series of global-mean annual-mean tropopause height calculated from ERA-40 and ERA-Interim are shown in Figure 5.1. It can be seen in Figure 5.1 that the global-mean annual-mean tropopause is higher in ERA-Interim compared to ERA-40, which is consistent with the differences between the seasonal climatologies shown in Figure 4.13. Although there is a difference in tropopause height between the two climatologies, there is a clear increasing trend in both time series. Where the two time series overlap, the rate of increase is slightly smaller in the ERA-Interim climatology. Global-mean annual-mean tropopause height calculated from ERA-Operational data is also shown in Figure 5.1 from 2002. It is very likely that the apparently large increasing trend in ERA-Operational tropopause height in the period 2002-2007 is due to changes in the forecasting system, and not to real changes in the tropopause.

Santer et al. (2003, 2004) identified volcanic signals in reanalysis data, and Seidel and Randel (2006) found significant changes following volcanic eruptions using radiosonde data. There are indications of decreases in tropopause height in ERA-40 following the Agung (1963) and El Chichón (1982) eruptions (Figure 5.1). However, they do not exceed the natural variability. Clear decreases in tropopause height can be seen following the Pinatubo (1991) eruption, in both ERA-40 and ERA-Interim, that appear to be outside the typical range of natural variability.

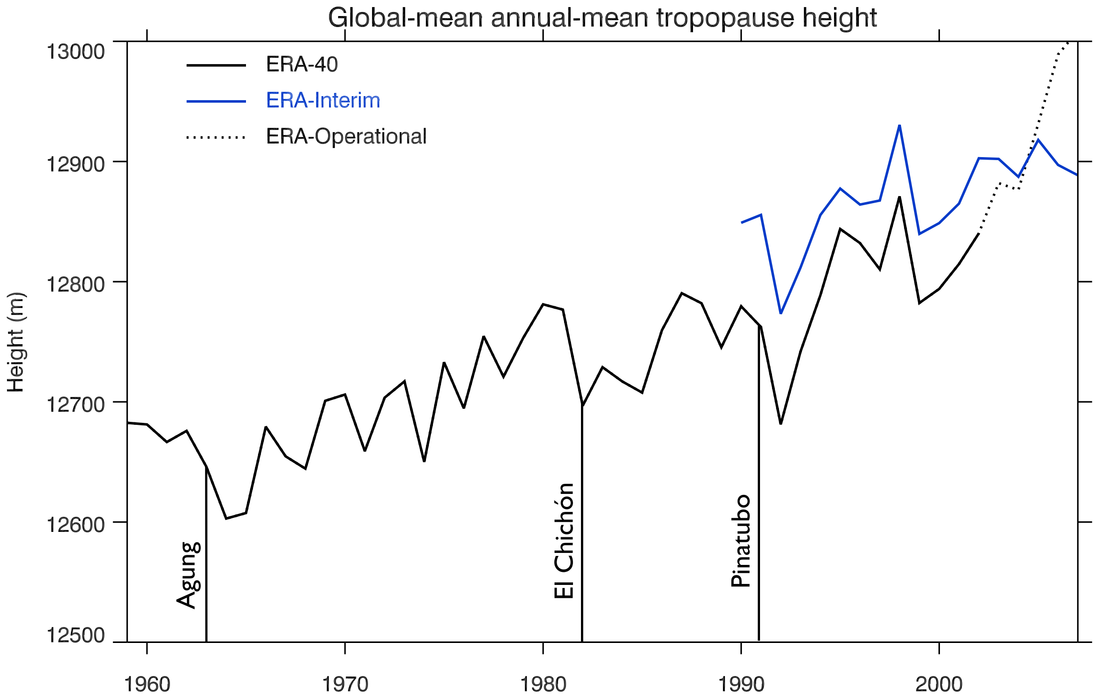


Figure 5.1: *Global-mean annual-mean tropopause height (m) from ERA-40 (black), ERA-Interim (blue), and ERA-Operational (dotted).*

5.1 Trends in ERA tropopause characteristics

Tropopause height, pressure, temperature and tropopause-level zonal wind have been calculated locally at the highest analysis resolution (Chapter 4). Trends in all variables, for all regional scales, were assumed to be linear, and were calculated using the least squares method. A 1-tailed t-test was used to identify when trends were significantly more positive or negative than zero. Trends were calculated at each grid-point and for area-mean tropopause variables, which were weighted according to the cosine of latitude where appropriate.

Positive trends in annual-mean global-mean tropopause height have been identified previously in several reanalysis and observational data sets (e.g. Santer et al. (2004); Seidel and Randel (2006)). Positive trends in global tropopause height can also be identified in both the ERA-Interim and ERA-40 time-series (see Tables 5.1 and 5.2).

ERA-Interim (1989-2007)	Annual	MAM	JJA	SON	DJF
Blended height	47.11**	49.02**	47.68**	37.14**	42.69*
Dynamic height	43.50**	47.12**	46.08**	32.74**	38.26*
Thermal height	46.08**	39.92*	49.74**	39.81**	41.57*
Blended pressure	-0.97**	1.10**	-1.06**	-0.78**	-0.74*
Thermal pressure	-0.61*	-0.44	-0.92	-0.60*	-0.27
Blended temperature	-0.18*	-0.17*	-0.15*	-0.12*	-0.24*

Table 5.1: *Trends in global-mean tropopause variables from ERA-Interim (1989-2007). Height trends are in m decade*$^{-1}$*, pressure trends are in hPa decade*$^{-1}$*, and temperature trends are in K decade*$^{-1}$*. Trends significantly different from 0 at the 1% level are indicated by **. Significance at the 5% level is indicated by *.*

ERA-40 (1979-2001)	Annual	JJA	DJF
Blended height	39.29**	39.16**	45.29*
Thermal height	103.90**	85.20**	132.63**
Blended pressure	-0.80*	-0.69	-0.83*
Thermal pressure	-2.61**	-2.01**	3.13**
ERA-40 (1958-2001)			
Blended height	39.42**	48.11**	30.32**
Thermal height	56.49**	26.96**	83.75**
Blended pressure	-0.87**	-1.37**	-0.33*
ERA-40 (1989-2001)			
Blended height	72.43*	32.24	91.23*
Thermal height	167.86**	136.56**	175.38*
Blended pressure	-1.02	-0.65	-1.80
Thermal pressure	-3.78**	-2.96**	-3.74*
ERA-Interim (1989-2001)			
Blended height	44.43	39.65	42.92
Thermal height	84.75*	71.59*	69.29
Blended pressure	-0.97	-0.74	-0.93
Thermal pressure	-1.90*	-1.46*	-1.19

Table 5.2: *Trends in global-mean tropopause variables from ERA-40. Height trends are in m decade*$^{-1}$*, pressure trends are in hPa decade*$^{-1}$*. Trends significantly different from 0 at the 1% level are indicated by **. Significance at the 5% level is indicated by *. ERA-Interim values for the overlap period (January 1989- August 2002) are also shown.*

5.1.1 Trends in the global-mean tropopause

Trends in global tropopause pressure previously identified in reanalyses range from +0.1 hPa decade^{-1} (1979-1993) (Hoinka, 1999) to -2.66 hPa decade^{-1} (1979-2001) (Santer et al., 2004), although most are in the range -1 to -2 hPa decade^{-1} (see Table 5.3 for a full summary). Seidel and Randel (2006) identified a global trend of -1.7 hPa decade^{-1} (1980-2004) using radiosonde data. They also used the same data to calculate trends for different time periods (see Table 5.3), which clearly demonstrated the sensitivity of calculated trends to the end points of the time series. There is a difference of more than a factor of two between their largest and smallest trend. This spread of results should be considered when trends are being compared. All of the trends shown in Table 5.3 are for a thermal tropopause, with the exception of those from Hoinka (1998, 1999). Hoinka used a latitude-dependent tropopause definition, based on that from Hoerling (1991), with a thermal tropopause in the tropics and 3.5 PVU in the extratropics, with interpolation between 19° and 36° latitude.

The trends in global-mean blended tropopause height identified here in both ERA data sets (Tables 5.1 and 5.2) are smaller than most of those given in previous work (Table 5.3), with the exception of those calculated by Hoinka (1998, 1999) who used a unique method to find the position of the thermal tropopause (Santer et al., 2003). They compare well with the Parallel Climate Model (PCM) results of Santer et al. (2003a), but are approximately half the magnitude of their trends calculated from the National Centres for Environmental Prediction (NCEP) reanalysis. Our ERA trends also lie at the low end of the range of values calculated by Seidel and Randel (2006) using radiosonde data for different time periods, and are comparable to the NCEP (1979-1993) trends from Santer et al. (2003). However, the trends found here are small compared to those calculated by Santer et al. (2004) using ERA-40. They calculated trends of -2.36 hPa decade^{-1} from data with 60 vertical levels, and -2.66 hPa decade^{-1} from data with 23 vertical levels for 1979-2001, two to three times larger than the values found here.

The comparatively small trends found here in ERA-40 are primarily the result of the tropopause definition used. The thermal tropopause height trend is +104 m decade^{-1}

Study	Data	Period	Global trend
Hoinka (1998)	ERA-15	1979-1993	-0.1 hPa decade^{-1}
Hoinka (1999)	ERA-15	1979-1993	+0.1 hPa decade^{-1}
Sausen and Santer (2003)	NCEP/NCAR	1979-1997	-1.82 hPa decade^{-1}
	GSOP	1979-1997	-1.47 hPa decade^{-1}
Santer et al. (2003)	NCEP/NCAR	1979-2000	-2.16 hPa decade^{-1}
	ECMWF	1979-2000	-1.13 hPa decade^{-1}
	NCEP	1979-1993	-1.05 hPa decade^{-1}
Santer et al. (2003a)	PCM	1979-1999	-0.95 hPa decade^{-1}
	NCEP	1979-1999	+95m decade^{-1}
Santer at al. (2004)	ERA-40 (60 lev)	1979-2001	-2.36±0.47 hPa decade^{-1}
	ERA-40 (23 lev)	1979-2001	-2.66±0.48 hPa decade^{-1}
	NCEP	1979-2001	-1.79 hPa decade^{-1}
	PCM	1979-2001	-1.13 hPa decade^{-1}
Seidel and Randel (2006)	Radiosonde	1980-2004	+64±21m decade^{-1}
		1980-2004	-1.7±0.6 hPa decade^{-1}
		1980-2004	-0.41±0.09K decade^{-1}
		1980-1993	-0.77±1.42 hPa decade^{-1}
		1980-1997	-1.60±1.00 hPa decade^{-1}
		1980-2000	-1.89±0.81 hPa decade^{-1}
		1980-2001	-1.83±0.75 hPa decade^{-1}
Son et al. (2009)	CCM	1979-1999	-1.53±0.42 hPa decade^{-1}
	NCEP/NCAR	1979-1999	-1.67 hPa decade^{-1}
	ERA-40	1979-1999	-2.2 hPa decade^{-1}
	AR4	1979-1999	-1.03±0.50 hPa decade^{-1}

Table 5.3: *Trends in global-mean tropopause height from previous studies. All studies use a thermal tropopause definition, with the exception of Hoinka (1998,1999) who uses a latitude-dependent blended tropopause based on the thermal tropopause and the 3.5 PVU surface. All acronyms are expanded in the list of acronyms at the end of this thesis.*

(-2.61 hPa decade^{-1}) in ERA-40 (1979-2001), much larger than that for the blended tropopause (40 m decade^{-1}, -0.80 hPa decade^{-1}). The thermal tropopause pressure trend is in excellent agreement with the ERA-40 trends found by Santer et al. (2004) using the same 23 level gridded data (-2.66 hPa decade^{-1}). The thermal tropopause trend is also comparable to the ERA-40 (1979-1999) trend from Son et al. (2009), although there is some difference due to the different periods considered. Using the thermal tropopause trends, rather than the blended tropopause trends, for comparison with previous studies actually places our trends at the upper end of the range of previous estimates (Table 5.3).

The difference between blended and thermal tropopause trends in ERA-Interim is smaller than in ERA-40 (see Table 5.1), which is to be expected as the thermal tropopause definition is less prone to spurious height anomalies in ERA-Interim. As the results of Seidel and Randel (2006) also demonstrate, the magnitude of the trends are also sensitive to the end points of the period for which they are calculated (see Tables 5.2 and 5.3).

Trends have also been calculated for January 1989 to August 2002 using ERA-40 and ERA-Interim (Table 5.2). Although this is a short time period for the calculation of trends, it allows a direct comparison between ERA-40 and ERA-Interim. There is reasonable agreement between the height trends for JJA, but ERA-40 height trends for DJF are more than double those from ERA-Interim. In all cases, the thermal tropopause trends from ERA-40 are much larger than those from ERA-Interim: more than double in the annual and DJF mean, and double in JJA. The ERA-Interim result is likely to be more reliable as the increased resolution resulted in an improved representation of the thermal tropopause.

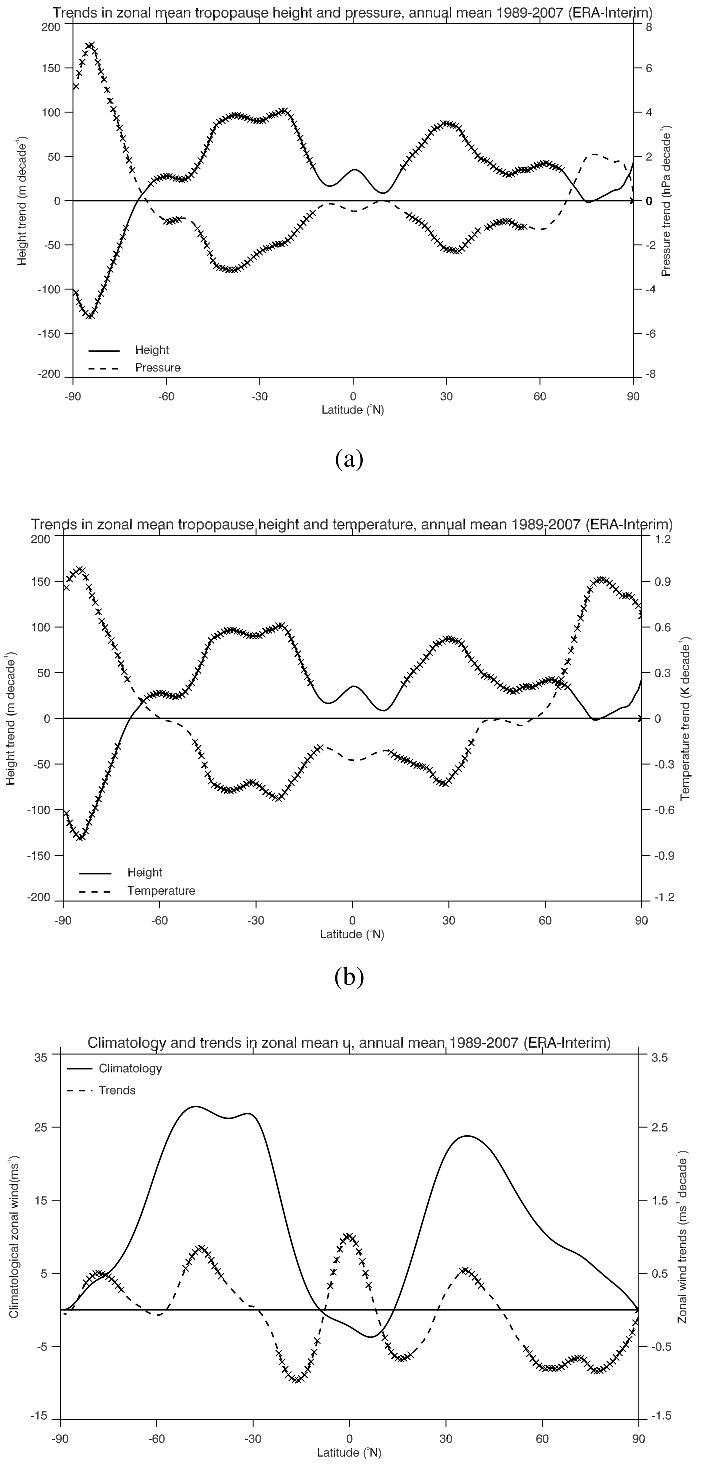
5.1.2 Trends in the zonal-mean tropopause

Considering the zonal-mean annual-mean trends, it can be seen in Figure 5.2(a) that the positive trend in tropopause height outside the tropics and high latitude regions is a 25 to 100 m decade^{-1} increase (corresponding to a 1 to 3 hPa decade^{-1} decrease in pressure). The zonal-mean annual-mean trends in the tropics are not significantly different from 0 at the 5% level. The largest positive height trends are found in the sub-tropics. These trends will be discussed in more detail in Section 5.3. In high Arctic latitudes there is a poor correspondence between the height and pressure trends. This deviation from hydrostatic balance, discussed later in this section, appears to be the result of the zonal and temporal averaging. Area weighting means that this only has a small influence on trends in regional average tropopause heights.

Zonal-mean annual-mean trends in tropopause height and temperature are shown in Figure 5.2(b). Temperature trends mirror the height trends, with negative trends in the subtropics, and positive trends at high latitudes. Trends in all of these regions are significantly different from 0 at the 5% level. There is also a smaller negative temperature trend in the tropics, associated with the small positive trend in height. Outside the polar regions, where there are large positive temperature trends of up to 1 K decade^{-1}, temperature trends are of the order of -0.4 K decade^{-1}. The ERA-Interim trend in the annual-mean globally-averaged tropopause temperature is -0.18 K decade^{-1}.

Trends in zonal-mean zonal wind are shown in Figure 5.2(c), alongside a climatology for reference. There are positive trends in zonal-mean zonal wind near 35°N (0.5 m s^{-1}decade^{-1}) and 45°S (0.8 m s^{-1}decade^{-1}). The maxima in these trends coincide with the latitudes of the climatological maximum zonal wind speed, which suggests an increase in the strength of the jets with time, as found by Lorenz and DeWeaver (2007). There are also positive trends of 1 m s^{-1}decade^{-1} and 0.5 m s^{-1}decade^{-1} at the Equator and 80°S respectively, which are significant at the 5% level. There are negative trends of -0.6 m s^{-1}decade^{-1} and -1.0 m s^{-1}decade^{-1} at ±15°latitude, on the equatorial side of the negative trends in tropopause temperature. The location of these negative trends in zonal wind suggests a broadening and weakening of the tropical easterlies, which will be discussed in Section 5.3. Additionally, there are negative trends of as much as -0.7 m s^{-1}decade^{-1} in the NH high latitudes, poleward of 50°N, that are significant at the 5% level.

Zonal-mean seasonal-mean trends for DJF and JJA are shown in Figure 5.3. The overall pattern of the zonal-mean trends in DJF and JJA are similar, but there is more zonal struc-



(a)

(b)

(c)

Figure 5.2: *(a) Annual-mean zonal-mean blended tropopause height and pressure trends (m decade^{-1} and hPa decade^{-1}) from ERA-Interim (1989-2007). (b) Tropopause height and temperature trends (m decade^{-1} and K decade^{-1}). (c) Tropopause level zonal-mean zonal wind climatology and trends (m s^{-1} and m s^{-1} decade^{-1}). Crosses indicate where trends are significant at or above the 5% level.*

ture in DJF. In DJF there are maximum positive trends in tropopause height of near 30°N and 25°S (+150 m decade^{-1}), and secondary peaks at 60°N, 45°S ($\sim$75 m decade^{-1}), and the Equator (+60 m decade^{-1}). The zonal-mean trends close to these latitudes are significant at or above the 5% level. There are also large negative height trends in the polar regions, particularly in the SH where the trends are statistically significant at the 5% level in both seasons. As in Figure 5.2, there is some discrepancy between the height and pressure trends in the NH high latitudes.

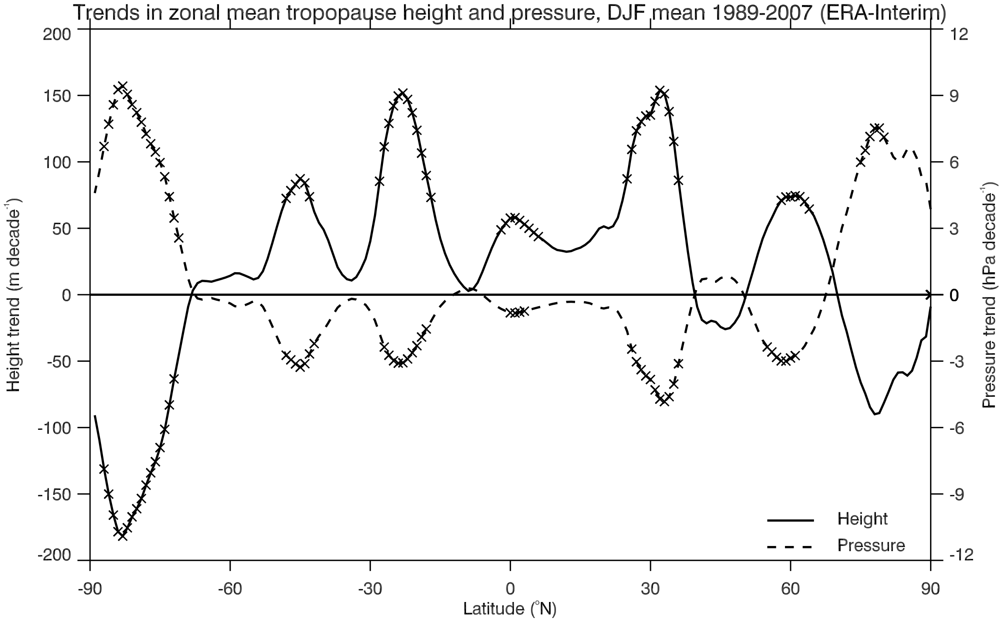
In JJA there are local peaks in the tropopause height trends at 80°N, 45°N, 25°S, and 50°S (+70, 100, 125, and 75 m decade^{-1}). The maximum zonal-mean trends have a smaller magnitude, but the trends are significant at or above the 5% level for more latitudes either side of these maxima compared to DJF. There is also a significant negative trend in the SH high latitudes, as in DJF. Comparison of Figures 5.2 and 5.3 shows that the annual-mean zonal-mean trends resemble the JJA trends more strongly than those for DJF, reflecting the similarity between the trends in JJA and the equinoctial seasons.

The apparent discrepancy between height and pressure trends in the NH high latitudes has been investigated, and appears to be the result of the temporal and spatial averaging used to calculate the zonal-mean tropopause. It is expected that tropopause height (z) and pressure (p) should scale according to the equation of hydrostatic balance:

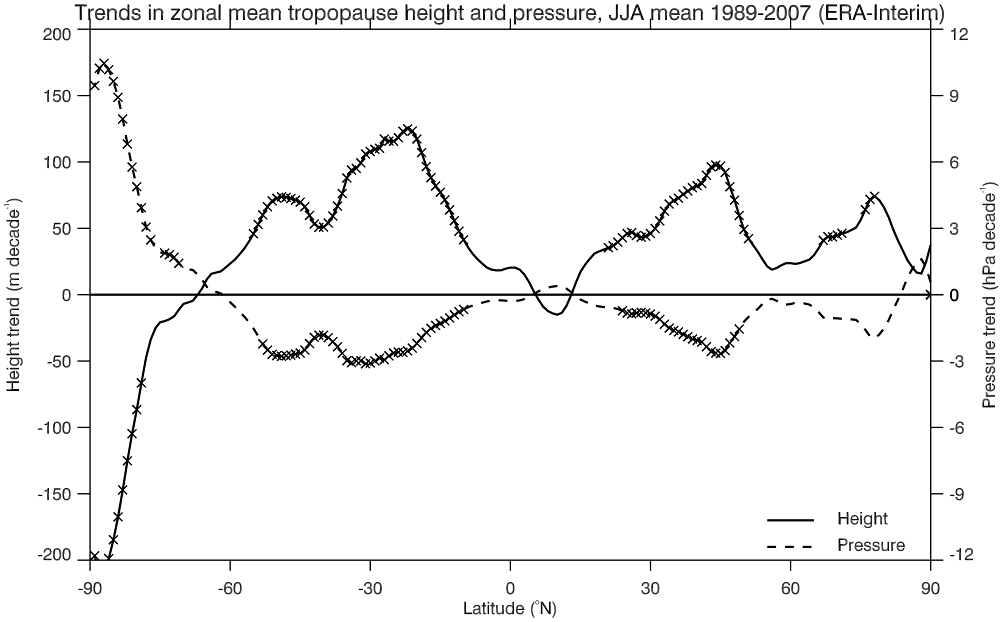
$$\frac{\Delta p}{\Delta z} = -\rho g \tag{5.1.1}$$

where ρ is density and g is the acceleration due to gravity.

Tropopause pressure was calculated at each latitude and longitude from the hydrostatic balance equation (p_{hydro}), using actual tropopause height and air density at the tropopause, and compared to actual tropopause pressure for intervals of 6 hours, 1 day, 1 month, and 3 months. Tropopause pressure was also calculated from Equation 5.1.1 for these periods using the average density for the period (p_{avrho}). It is the average density that is effectively used when comparing the zonal-mean seasonal trends in tropopause height and pressure.



(a)



(b)

Figure 5.3: *Seasonal-mean zonal-mean blended tropopause height and pressure trends (m decade*$^{-1}$ *and hPa decade*$^{-1}$*) from ERA-Interim (1989-2007). (a) December to February mean. (b) June to August mean. Crosses indicate where trends are significant at or above the 5% level.*

Discrepancies between actual pressure and p_{avrho} were largest in regions with large tropopause variability, and the NH high latitudes in particular. However, it should be noted that this is only an important factor to consider when looking at trends in the zonal-mean tropopause. The cosine latitude weighting applied in the calculation of global-mean trends means that NH high-latitude trends only make a small contribution to global mean trends, so there is generally good correspondence between height and pressure trends.

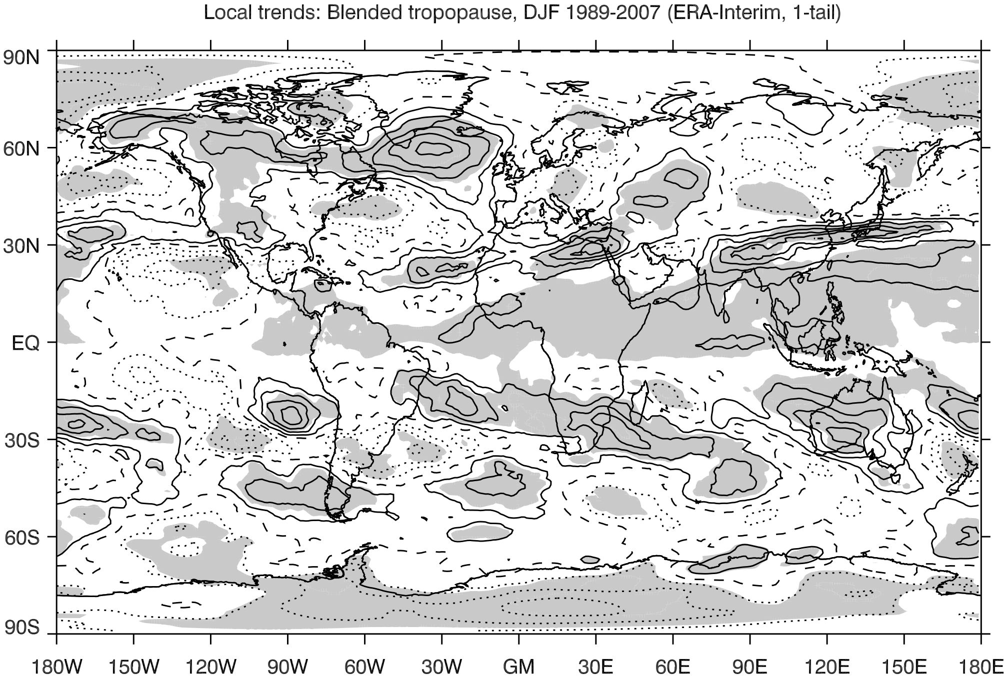
5.1.3 Geographical structure of trends

Grid-point trends for the two seasons reveal the spatial structure of the trends in further detail (see Figure 5.4). The dominant feature in Figure 5.4 is the positive height trends in the subtropics in both DJF and JJA, along with positive trends in the tropics in DJF. These features will be discussed further in Section 5.3. In addition to the subtropical trends, there are tripole structures in height over the North Pacific and North Atlantic in DJF (Figure 5.4(a)). Over the North Pacific an increase in height of $\sim$300 m decade^{-1} in the region near 35°N is flanked by decreases in height of $\sim$-300 m decade^{-1} to the south, and $\sim$-200 m decade^{-1}. Over the North Atlantic there is a decrease in height of $\sim$-200 m decade^{-1} in the mid-latitudes flanked by increases of over 500 m decade^{-1} in the north, and up to 300 m decade^{-1} to the south. This is suggestive of a relationship with the negative North Atlantic Oscillation (NAO) pattern. The NAO index[1] decreases over the period of the climatology, and the difference between the tropopause pressures in the north and south of the region correlates well with the NAO index, with r^2=0.6. The correlations between tropopause height over the North Pacific with the West Pacific Pattern and the North Pacific Index are weaker.

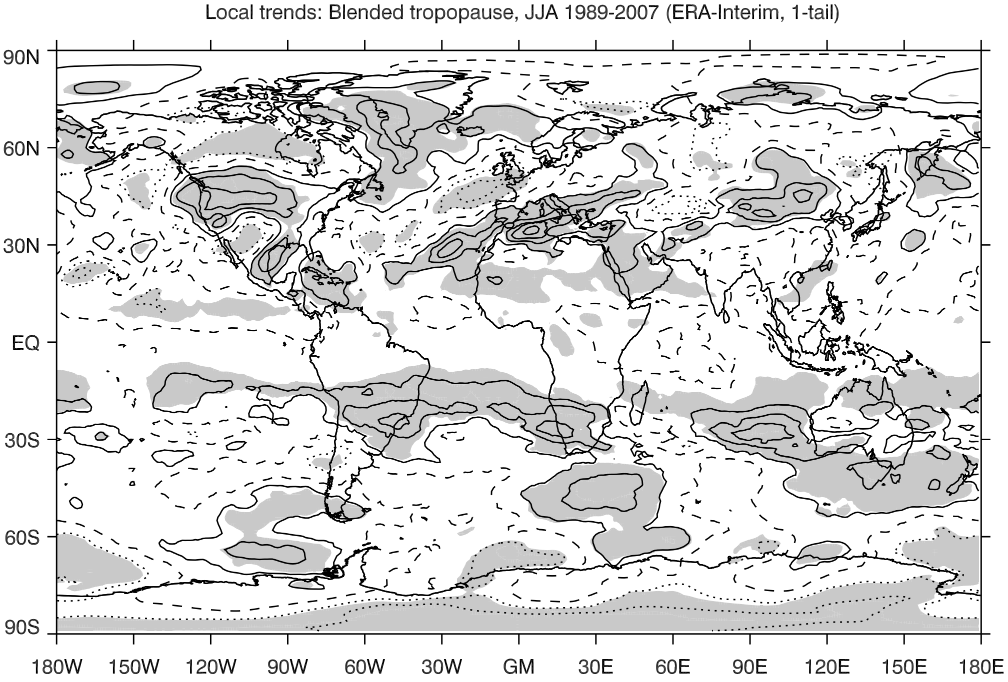
It can also be seen in Figure 5.4(a) that there are significant positive height trends across most of the tropical region in DJF, with the exception of negative height trends over the central Pacific. These trends have a similar structure to the La Niña anticyclones, which

---

[1] The NAO index is the difference between normalised sea level pressure in the Azores and Iceland. A time series is available at www.cgd.ucar.edu/cas/jhurrel/indices.html



(a)



(b)

Figure 5.4: *Trends in blended tropopause height DJF (panel a) and JJA (panel b) for ERA-Interim (1989-2007). Solid contours show positive trends, the zero line is dashed, and negative contours are dotted. The contour interval is 100 m decade*$^{-1}$ *for height trends. Shading highlights where the trends are significant at or above the 5% confidence interval.*

is a reflection of the large number of La Niña years in the climatology. It can be seen in Figure 5.3 that the zonal-mean height trends in the tropics in DJF are $\sim$50 m decade^{-1}. However, it can be seen in Figure 5.4(a) that isolated trends can be more than double this value in the eastern hemisphere.

Height trends in JJA (Figure 5.4(b)) are generally smaller than those in DJF, but there are still several regions where the trends are significant. There are increases in height of more than 300 m decade^{-1} in the SH subtropics and extratropics, as in DJF, although the trends are weaker over the South Pacific. The negative trends over the Antarctic continent are also $\sim$100 m decade^{-1} weaker compared to those in DJF, but still significant at the 5% level. There is a tripole structure in the trends over the North Atlantic. It is shifted polewards and eastwards compared to the DJF pattern, and the positive trend over Greenland is smaller in magnitude: only $\sim$200 m decade^{-1}. Over North America an increase in height of up to $\sim$300 m decade^{-1} in the region 40°N to 50°N is seen.

Trends in tropopause-level zonal wind (not shown) have a similar spatial pattern to the tropopause height trends, particularly in the NH, although they are shifted polewards compared to the height trends.

5.2 Antarctic trends

The Antarctic is the only substantial region with statistically significant negative height trends. These trends are present in all seasons, and they do not extend far beyond the Antarctic coastline, if at all. It can be seen in Figure 5.4 that the trends are slightly larger in DJF compared to JJA, with local maxima in excess of 300 m decade^{-1} compared to 200 m decade^{-1}. This is also true for the regional average trends for 75°S-90°S from ERA-Interim, which are shown in Table 5.4.

Trends from ERA-40 for different periods, and ERA-Interim for the overlap period, are also shown in Table 5.4. ERA-40 is known to be unreliable in the Antarctic (Bromwich and Fogt, 2004), so the ERA-40 trends in Table 5.4 should be treated with a degree of

ERA-Interim (1989-2007)	Annual	JJA	SON	DJF
Blended height	-88.90**	-82.82**	-84.48**	-147.12**
Blended pressure	5.04**	4.77*	4.93*	7.54*

ERA-40 (1958-2001)	Annual	JJA	DJF
Blended height	132.28**	129.68**	53.03*
Blended pressure	-4.95**	-4.72**	-1.61
ERA-40 (1979-2001)			
Blended height	196.97**	167.09**	134.82
Blended pressure	-7.70**	-5.40*	-6.22
ERA-40 (1989-2001)			
Blended height	-258.23	-233.86	-88.59
Blended pressure	12.86	10.54	4.40
ERA-Interim (1989-2001)			
Blended height	-131.39*	-101.82	-141.25
Blended pressure	6.74*	5.58	4.65

Table 5.4: *Trends in 75-90°S mean blended tropopause height and pressure from ERA-Interim 1989-2007, and ERA-40 for 1958-2001, 1979-2001, and 1989-2001. ERA-Interim trends are shown for 1989-2001 for comparison. Height trends are in m decade^{-1}, pressure trends are in hPa decade^{-1}. Trends significantly different from 0 at the 1% level are indicated by **. Significance at the 5% level is indicated by *.*

caution. There is also some question about the validity of defining a tropopause at all in the Antarctic winter (Section 4.2.2). However, there is a clear difference in the sign of the annual and solsticial seasonal trends for 1958-2001 and 1979-2001, and 1989-2001.

It is possible that the negative tropopause height trends for periods beginning after 1989 may be the result of ozone recovery. Ozone depletion causes the height of the tropopause to increase, and vice versa (Santer et al., 2004). Although ozone recovery did not begin until 2000, tropospheric halogen loading peaked in 1993 (Eyring et al., 2010). Hence, it is possible that trends for 1989-2007 are reflecting a period of small changes in lower stratospheric ozone, followed by a period of ozone recovery, which will result in negative trends in tropopause height. Negative trends in the decades immediately following 2000 can be seen in some model simulations in Gettelman et al. (2010) for 60-90°S and in the 1990s to $\sim$2010 in some model simulations in Son et al. (2009) for 45-90°S.

Time series of annual-mean Antarctic tropopause height from ERA-40 and ERA-Interim are shown in Figure 5.5, which clearly reflect the area mean trends given in Table 5.4.

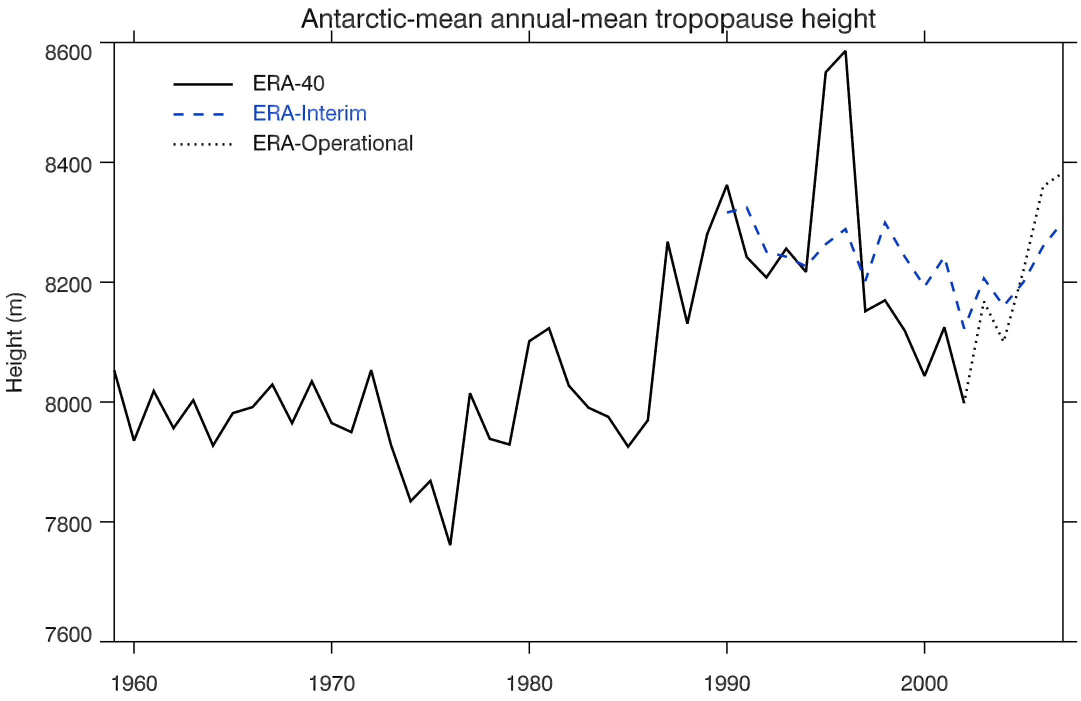


Figure 5.5: *60-90°S-mean annual-mean tropopause height (m) from ERA-40 (black), ERA-Interim (blue dashed), and ERA-Operational (dotted).*

An increasing trend in tropopause height can be seen in ERA-40 from the late 1970s to 1990, followed by a general decrease in height. This decrease in height is mirrored by ERA-Interim. There is a suggestion of a return to positive or near zero trends towards the end of ERA-Interim. The time series in Antarctic tropopause height correspond to the trends in ozone shown by Gettelman et al. (2010). It is also very clear from Figure 5.5 that the choice of end points is critical to the value of the trend.

The leading mode of variability in geopotential height and zonal wind in the SH is the SH Annular Mode (SAM), which is defined as the difference of zonal-mean sea-level pressure between 40°S and 65°S. A significant increase in the SAM has been observed since 1958 (Marshall, 2003) (pressures over Antarctica becoming lower relative to pressure in the mid-latitudes). This increase in the SAM is linked to a strengthening of the circumpolar vortex, and intensification of the westerlies around Antarctica. Such an intensification of the westerlies can be seen in Figure 5.2. The positive tropopause height trends seen in the earlier ERA-40 trends (Table 5.4) are also consistent with a positive trend in the SAM.

Antarctic trends have been found to be particularly sensitive to the tropopause definition used. This sensitivity can be seen in Table 5.5 where thermal and blended tropopause trends are compared for the Antarctic region (75-90°S). Thermal and blended tropopause trends are generally very different. Very often they have the opposite sign, and differ by orders of magnitude. In addition to the fact that trends are being compared for very different surfaces, part of the sensitivity to tropopause definition is due to the low reliability of the thermal tropopause definition in Antarctic winter. Often, it is not possible to identify the required lapse rate condition at typical tropopause levels, resulting the allocation of the tropopause to anomalously high altitudes (Section 3.3). These anomalies are then included in trend calculations where they can have a substantial impact. There is also some question about the validity of the dynamic tropopause definition in Antarctic winter, raising questions about whether a tropopause should be defined at all (see Section 4.2.2). However, it is always possible to identify the |PV|=2 PVU surface, so there is no problem with spurious extreme values when the blended definition is used.

The influence of thermal tropopause anomalies on calculated Antarctic trends was tested by neglecting any grid points where the tropopause altitude was above a threshold level. As the anomalies are not always apparent in seasonal mean data, six-hourly tropopause surfaces were used. Trends were calculated for the annual mean, JJA, and DJF. In order to set the threshold at the lowest possible altitude, different thresholds were used in each case: 12.7 km for annual means, 12.5 km for JJA, and 11 km for DJF. The results of this analysis are shown in Table 5.5. Trends calculated using all Antarctic grid points are shown in the columns ‘All’. When anomalous thermal tropopause altitudes were identified, affected grid points were removed from the both the thermal and blended tropopause calculations. The results from these calculations are shown in the columns ‘Capped’.

It can be seen in Table 5.5 that blended tropopause trends are insensitive to the removal of grid points with anomalous thermal tropopause heights from the trend calculations. However, the impact on the thermal tropopause trends is large, with some values chang-

	Annual		JJA		DJF	
ERA-Int. (1989-2007)	All	Capped	All	Capped	All	Capped
Blended height	-88.84**	-88.72**	-82.82*	-82.67*	-147.13**	-148.89**
Blended pressure	5.04**	5.04**	4.77 *	4.77*	7.54*	7.61**
Thermal height	-9.14	-27.46	259.93	76.26	-116.13	-118.16
Thermal pressure	1.51	1.69	-6.58*	-2.57	5.27*	5.33
ERA-40 (1979-2002)	All	Capped	All	Capped	All	Capped
Blended height	196.62**	196.30**	141.33	140.38	134.56*	134.43*
Blended pressure	-7.69**	-7.68**	-4.38	-4.35	-6.28*	-6.28*
Thermal height	252.94	329.35	-84.54	172.37	428.05	427.77
Thermal pressure	-9.55**	-11.24**	1.05	-4.66	-18.24**	-18.23**
ERA-Int. (1989-2002)	All	Capped	All	Capped	All	Capped
Blended height	-71.35	-71.33	-101.82	-102.03	-141.26	-143.23
Blended pressure	3.09	3.08	5.58	5.59	4.65	4.72
Thermal height	174.19	87.75	260.02	49.40	-1.92	-6.20
Thermal pressure	-6.47	-4.52	-6.60	-1.87	-2.53	-2.40
ERA-40 (1989-2002)	All	Capped	All	Capped	All	Capped
Blended height	-164.90	-165.02	-233.84	-235.89	-88.61	-88.73
Blended pressure	8.29	8.29	10.54	10.60	4.40	4.41
Thermal height	381.76	423.60	-220.51	42.26	448.73	448.26
Thermal pressure	-13.38*	-14.28*	5.73	-0.45	-18.52	-18.51

Table 5.5: *Trends in 75-90°S mean blended and thermal tropopause height and pressure from ERA-Interim 1989-2007 and 1989-2002 and ERA-40 for 1979-2001 and 1989-2002. Height trends are in m decade*$^{-1}$*, pressure trends are in hPa decade*$^{-1}$*. Trends significantly different from 0 at the 1% level are indicated by **. Significance at the 5% level is indicated by *. 'Capped' values are the result of trend calculations where grid points with a tropopause height in excess of the threshold value were neglected.*

ing by orders of magnitude, or even changing sign. There is generally better agreement between trends in the thermal and blended tropopause when grid points with an anomalous thermal tropopause are neglected, but in some cases it is considerably worse.

Comparison of the ERA-Interim and ERA-40 trends for 1989-2002 shown in Table 5.5 also highlight the sensitivity of Antarctic trends to the data set used. ERA-40 is known to be less reliable than ERA-Interim in the Antarctic (Bromwich and Fogt, 2004), and this is reflected in the large differences in the calculated trends, even for the blended tropopause. The number of anomalous thermal tropopause grid points is reduced in ERA-Interim compared to ERA-40. However, this has no clear affect on the agreement between the thermal tropopause trends with and without anomalous points included.

The lack of consistency between thermal tropopause trends calculated with and without anomalous points, particularly in JJA, suggests that thermal tropopause trends are unreliable in the Antarctic. However, although Antarctic trends are large they only make a small contribution to global trends. Global thermal tropopause trends are insensitive to the removal of anomalous Antarctic points.

5.3 Tropical broadening

One of the most prominent features in Figures 5.2 and 5.4 is the large positive trends in the sub-tropics. Tropopause height gradients are large here, and the large height trends are indicative of a latitudinal expansion of the region of high tropopause height characteristic of the tropics. This ‘tropical broadening’ has been identified in a number of previous studies (e.g. Seidel and Randel, 2007, Lu et al., 2007). The widening trend is systematic (Hu and Fu, 2007), with only small seasonal variation (Fu et al., 2006). Previous studies (summarised in Table 5.6) suggest that the Hadley cell has widened by 2-5° since 1979 (Johanson and Fu, 2009). However, the mechanisms behind the expansion are poorly understood, and the observed rate of expansion exceeds model projections (Seidel et al., 2008).

Study	Data	Period	Method	Rate
This study	ERA-Interim	1989-2007	Critical altitude	0.9° decade^{-1} (DJF)
				0.9° decade^{-1}(JJA)
		1989-2007	Height trends	1.3° decade^{-1} (DJF)
				1.1° decade^{-1} (JJA)
		1989-2007	Wind trends	1.0° decade^{-1} (DJF)
				2.2° decade^{-1} (JJA)
Birner (2010)	NCEP/NCAR, NCEP/DOE, JRA25, ERA-40	1979-2005	Seidel and Randel (2007)	$\sim$1–2° decade^{-1}
	As above	1979-2001	Objective width definition	$\sim$1° decade^{-1}
	NCEP/NCAR	1979-2009	Objective height definition	0.7±0.3° decade^{-1}
	NCEP/DOE	1979-2009		0.4±0.4° decade^{-1}
	JRA25	1979-2009		-0.5±0.4° decade^{-1}
	ERA-40	1979-2001		Not significant
	All above		Alternative height definition	0.6–1° decade^{-1}
Seidel and Randel (2007)	NCEP/NCAR	1979-2005	Critical tropopause altitude	3.1±1.4° decade^{-1}
				1.8±0.6° decade^{-1}
				1.7±0.6° decade^{-1}
	ERA-40	1979-2002	Critical tropopause altitude	comparable to above
Hu and Fu (2007)	ERA-40	1979-2002	Meridional mass streamfunction=0 kg s^{-1}	1.29° decade^{-1}
	NCEP/NCAR	1979-2005		1.44° decade^{-1}
	NCEP/DOE	1979-2005		2.04° decade^{-1}
	HIRS	1980-2003	OLR=250 W m^{-2}	1.88° decade^{-1}
	ISSCP	1984-2004		1.67° decade^{-1}
	GEWEX	1984-2004		0.95° decade^{-1}
Fu et al. (2006)	MSU	1979-2005	Temperature observations	0.74° decade^{-1}
Hudson et al. (2006)	TOMS	1979-2003	Satellite observations of ozone	1° decade^{-1} (NH)

Table 5.6: *Rates of tropical broadening from previous studies. All acronyms are expanded in the list of acronyms given at the end of this thesis.*

There is no consensus as to the primary cause(s) of tropical broadening, although several possible mechanisms have been suggested. Seidel and Randel (2007) suggested that the broadening is the result of anomalously low tropical lower stratospheric temperatures, and anomalously high tropical tropopause temperatures. Johanson and Fu (2009) suggest that these changes may have acted together to increase the phase speed of baroclinic eddies, widening the Hadley cell. Lu et al. (2007) found that increases in subtropical static stability forced the baroclinic instability zone poleward, along with the outer edge of the Hadley cell. They found that changes in stability alone accounted for 60% of the variance in the spread of broadening estimates in the IPCC Fourth Assessment Report (AR4) models. Additionally, they found that it could account for the consensus estimate of the total expansion. It has also been suggested that increases in sea surface temperature (Johanson and Fu, 2009, Seidel et al., 2008) and changes in ENSO (Seidel et al., 2008) may contribute to the broadening. However, Hu and Fu (2007) found that there was no correlation between ENSO and tropical broadening.

As the edge of the Hadley cell moves polewards, so too does the downwelling region associated with it. This results in a migration of the arid regions. It has been suggested in some studies (e.g. Chen et al., 2002, Wielicki et al., 2002) that in addition to widening, the Hadley cell is also strengthening, which will result in enhanced precipitation in existing moist regions, and less precipitation in current arid regions. However, it should be noted that Trenberth et al. (2002) raise concerns that the trends observed by Chen et al. (2002) and Wielicki et al (2002) were actually due to offsets and discontinuities in their satellite data. Although intensification has also been identified in model studies (e.g. Quan et al., 2004), Mitas and Clement (2005) have also expressed concern about the reliability of suggestions of Hadley cell strengthening, as the signal is not consistent across all the ‘commonly used’ data sets (NCEP-NCAR, NCEP-DOE, and ERA-40).

The zonal mean of tropical and subtropical tropopause heights, averaged over the first and last five years of ERA-Interim, is shown in Figure 5.6. In both DJF and JJA the average height for the last five years is greater than or equal to the average height for the first five years and the de-trended standard deviation of height for most latitudes. It is apparent from Figure 5.6 that expansion into the SH makes a greater contribution to

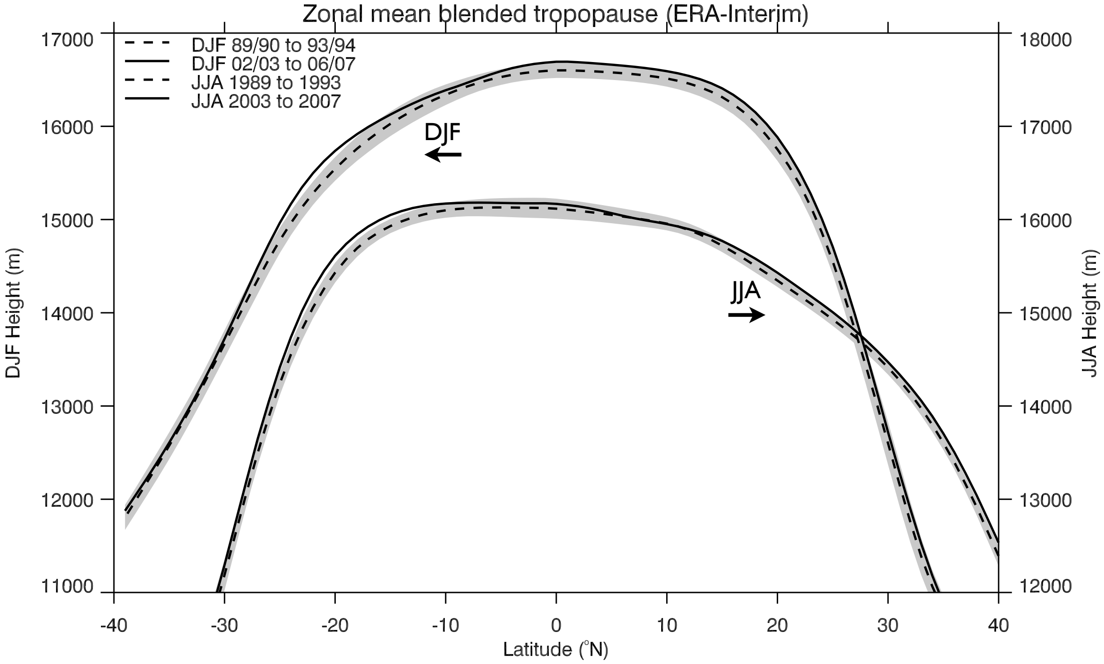


Figure 5.6: *Zonal mean blended tropopause height for ERA-Interim. The average for the first (dashed line) and last 5 years (solid line) of the climatology are shown for DJF and JJA. Shading shows the average of the first five years ± the de-trended standard deviation for the whole of ERA-Interim. The left and right hand axes refer to DJF and JJA respectively.*

the total broadening than that into the NH, a feature of this expansion that has also been observed by Hu and Fu (2007), and Seidel and Randel (2007). The rate of expansion into the SH also shows less seasonal variation than that into the NH.

Several different metrics have been used in previous studies to identify the edge of the 'tropical tropopause' in order to quantify the latitudinal expansion of the tropical belt. These include the location where the thermally-driven jet first becomes unstable (Lu et al., 2007), the number of 'high tropopause' days (Seidel and Randel, 2007), and specified values of outgoing long wave radiation (Johanson and Fu, 2009). Hudson et al. (2006) used critical values of ozone, and Johanson and Fu (2009) identified the latitude where meridional mass flux in the troposphere, and P-E, first become zero on the poleward side of the subtropical maximum (minimum for P-E).

The use of different diagnostics makes it difficult to compare the results of different studies, and can result in very different estimates of tropical broadening. However, most

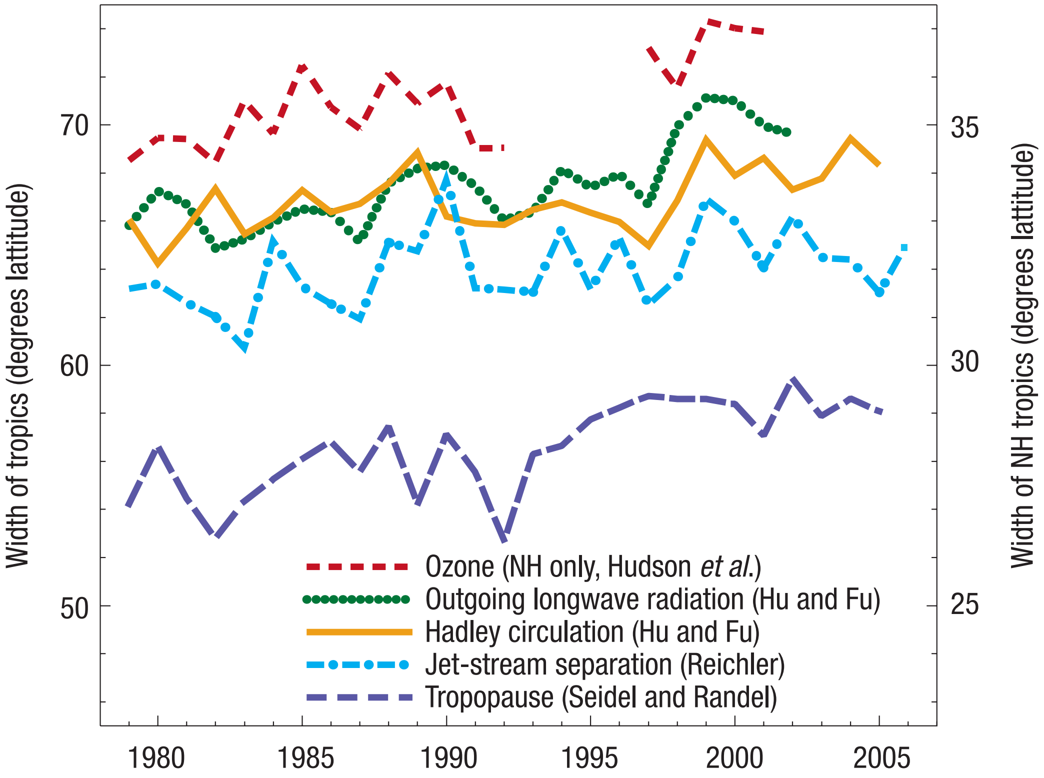


Figure 5.7: *The width of the tropics from different studies using different metrics, from Seidel et al. (2008). In each case an increase is seen since 1979, but the rates of expansion vary from 0.8° decade*$^{-1}$ *to 1.92° decade*$^{-1}$.

estimates of tropical broadening in previous studies lie in the range 0.5 to 3° decade^{-1} (Table 5.6), and the broadening signal appears to be robust in the majority of metrics used (see Figure 5.7). In this thesis, three metrics are used to quantify the rate of tropical broadening.

Seidel and Randel (2007) define the tropical belt as the region where the tropopause is higher than a critical altitude of 15 km for more than 100, 200, or 300 days per year. Their metric is adapted here, and the tropical blended tropopause is defined as having a seasonal-mean zonal-mean altitude of 15 km or more. The rate of expansion was evaluated at 15, 15.5, and 16 km by taking the difference in latitude between the means of the first and last five years of the climatology. The mean expansion over these altitudes was taken as the rate of tropical broadening. The total rate of expansion in ERA-Interim

(1989-2007) is 0.9° decade^{-1} in both DJF and JJA. In ERA-40 (1958-2001), the rate is smaller (0.7° decade^{-1} in both seasons). Both estimates are smaller than those of Seidel and Randel, who found trends of 1.6° latitude decade^{-1} in the SH, and 1.5° latitude decade^{-1} in the NH. However, the values presented in this thesis are in agreement with previous estimates, which range from -0.5° to 3.1° latitude decade^{-1} (see Table 5.6). Birner (2010) notes that Seidel and Randel use the NCEP tropopause product. This is based on the WMO thermal tropopause definition, but does not include the 2 km criterion. Birner found that this resulted in large trends compared to those calculated using the full WMO definition using NCEP reanalysis data, which may explain some of the difference between the estimates in this thesis, and those from Seidel and Randel (2007).

It can be seen in Figure 5.4 that the large positive trends in zonal-mean height in the NH sub-tropics seen in Figure 5.3 are due mostly to positive height trends in the eastern hemisphere. This is especially the case in DJF when there are height increases of up to +400 m decade^{-1} over the east Mediterranean, south-east Asia, and Japan (Figure 5.4(a)). There are only small or negative trends in the sub-tropics in the western half of the NH. This is in contrast to the findings of Seidel and Randel (2007) who found that the increases in subtropical tropopause height were restricted to the western hemisphere. Trends in the SH are more zonally uniform, although they are still larger in the eastern hemisphere, with positive trends of up to +400 m decade^{-1} at 20-30°S. In JJA the trends in the NH subtropics are weaker, with only the Mediterranean and the central Atlantic having significant positive height trends (Figure 5.4(b)). In the SH, the picture is very similar to DJF, although slightly smaller in magnitude.

The tropical broadening shown in Figure 5.6 can also be seen in Figure 5.3. The large positive height trends in the sub-tropics in both seasons are at similar latitudes to the steep gradients in tropopause height. The average height trend and the latitudinal gradient in tropopause height can be combined to give an alternative measure of the expansion rate. Here the rate of expansion is defined as the ratio of the average height trends and average height gradients for a specified latitude range, a second metric used here.

	3°		5°		9°	
Blended	Rate	% SH	Rate	% SH	Rate	% SH
DJF	2.0	72	1.7	72	1.4	71
JJA	1.6	54	1.3	53	1.1	53
Dynamic						
DJF	2.2	70	1.7	70	1.5	68
JJA	1.7	56	1.4	55	1.2	56

Table 5.7: Rates of tropical broadening [° decade^{-1}] for an ERA-Interim (1989-2007) blended and dynamic tropopause, for different widths of latitudinal averaging centred on the latitude of maximum height trend.

The rate of tropical broadening in ERA-Interim (1989-2007), using the blended and dynamic tropopause definitions, is shown in Table 5.7. Results from the use of averages over a 3°, 5° and 9° latitude range, centred on the latitude with the maximum sub-tropical trend, are shown. The rate of expansion is sensitive to the number of latitudes over which the calculation is performed, ranging from 2.0° decade^{-1} (1.6° decade^{-1}) for the 3° latitude range to 1.4° decade^{-1} (1.1° decade^{-1}) for the 9° range for DJF (JJA). However, the contribution from each hemisphere to the total expansion is not sensitive to this choice. Expansion into the SH accounts for ∼72% of the total expansion in DJF, and ∼55% in JJA.

The rate of tropical broadening is again smaller in ERA-40 (1979-2001) compared to ERA-Interim (1989-2007): 1.2 °decade^{-1} in DJF and 0.6 °decade^{-1} in JJA when a 3° latitude range is considered. In contrast to the ERA-Interim result, in ERA-40 the largest contribution to the total expansion always comes from the summer hemisphere.

The latitudes where large sub-tropical tropopause height trends occur coincide with the latitudes where the blended tropopause definition begins to deviate from the dynamic tropopause definition. To explore the possibility of a sensitivity of the rate of broadening to the blending process, the trends have been determined separately for the two alternative data sets based on the dynamic and thermal tropopause definitions. It can be seen in Table 5.7 that using the dynamic definition alone results in very similar estimates for tropical broadening to those using the blended definition. The values are slightly larger, which is consistent with the generally slightly smaller values found using the thermal definition. It can be concluded that the blending process that often occurs in the sub-tropical region is not critical in the estimates obtained for tropical broadening.

A third metric for tropical broadening has been obtained here through the broadening tendency of the tropical easterlies noted in Section 5.1.2, especially Figure 5.2. The length of time that it would take for the largest negative trends in zonal wind to reduce the climatological westerly at that latitude to zero was calculated. This could then be compared to the climatological latitude where u=0 to give a rate of broadening i.e.:

$$Rate = \frac{latB - latA}{(u/\frac{du}{dt})_B} \tag{5.3.1}$$

where $latA$ is where the zonal-mean climatological zonal-wind (u) is 0, and $latB$ is the latitude with the largest negative zonal-wind trends. For ERA-Interim, the broadening trends were found to be 1.0° decade^{-1} in DJF and 2.2° decade^{-1} in JJA. In both cases, 70% of the total broadening is due to expansion into the SH.

The rates of tropical broadening calculated using the three metrics are comparable in both seasons, with the exception of the broadening calculated using easterly wind tendencies in JJA, which is approximately double that from other methods. All estimates lie within the range of estimates from previous studies (Table 5.6). However, while the zonal-mean tropopause height and wind calculations indicate that expansion into the SH always makes a larger contribution to the total broadening, the metric using zonal-mean height trends indicates that it is expansion into the summer hemisphere that makes the largest contribution. The different magnitudes of the JJA trends, and the discrepancies between the hemispheric contributions to the expansion, demonstrate the sensitivity of the expansion to the metric used.

5.4 Summary

Trends in tropopause variables have been calculated for ERA-Interim (1989-2007) and ERA-40 (1958-2001 and 1979-2001). Trends have also been calculated for January 1989 to August 2002, when the reanalyses overlap. Although the global-mean annual-mean tropopause is higher in ERA-Interim compared to ERA-40, there are similar trends in blended tropopause height in both reanalyses for most of the periods considered. For ERA-Interim (1989-2007) the trend in annual-mean global-mean tropopause height is +47 m decade^{-1} (-0.97 hPa decade^{-1}). Both the height and pressure trend are significant at the 1% level. There is also a corresponding trend of -0.18 K decade^{-1} in annual-mean global-mean tropopause temperatures.

Trends in annual-mean global-mean pressure and height, from both ERA-Interim and ERA-40, are small compared to trends given in the literature. Direct comparison with the ERA-40 results of Santer et al. (2004) demonstrates that this difference is due mostly to the use of the blended tropopause definition, instead of the thermal definition. The thermal tropopause trends from ERA-40 (+104 m decade^{-1}, -2.61 hPa decade^{-1}) are in excellent agreement with the results of Santer et al. (2004), and are at the upper end of the range of previous estimates. Thermal tropopause trends from ERA-Interim (+46 m decade^{-1}, -0.61 hPa decade^{-1}) are in good agreement with those for the blended tropopause.

It can be seen from annual-mean zonal-mean tropopause height trends that the largest positive trends, of up to 100 m decade^{-1}, are in the sub-tropics. Smaller positive trends are found in the extratropics. Significant negative height trends are found in the SH high latitudes. Pressure and temperature trends mirror height trends. Additionally, there are large negative trends in zonal-wind in the sub-tropics, and large positive trends in the tropics, suggesting a broadening and weakening of the tropical easterlies.

When the geographical structure of the trends in tropopause height was examined at a grid-point scale it was revealed that most of the large trends in the subtropics are in the eastern hemisphere, especially in DJF. There are also tripole structures in the trends over the North Atlantic and North Pacific oceans. It can also be seen that the large negative trends in the SH high latitudes are restricted to over the Antarctic continent.

The Antarctic is the only substantial region with statistically significant negative tropopause height trends that persist in all seasons. The positive trends in circumpolar winds are consistent with positive trends in the SAM, while the negative trends in tropopause height after 1989 are consistent with an increase in lower stratospheric ozone. Comparing ERA-Interim and ERA-40 trends since 1989 with trends in ERA-40 with earlier trends in ERA-40 reveals a change in sign in Antarctic trends. Trends for periods starting before 1989 are positive, while trends for periods starting in 1989 are negative. This is true for all seasonal and annual averages. The changing signs of the Antarctic height trends can clearly be seen in a time series of Antarctic tropopause height, and closely follow the trends in stratospheric ozone.

Antarctic tropopause trends were found to be sensitive to the tropopause definition used. The Antarctic tropopause can be difficult to identify using the thermal definition, which results in an anomalously high tropopause being recorded. Removing all the points where a very high thermal tropopause was recorded can change the thermal tropopause trends by orders of magnitude, and even change their sign. The same procedure has little impact on blended tropopause trends, which are more robust in the Antarctic region.

A broadening of the tropical region has been identified, and quantified using three metrics based on a critical tropopause altitude, the product of sub-tropical height trends and gradients, and the poleward extent of easterly winds. If the tropics are defined as the region where the seasonal-mean zonal-mean tropopause has an altitude of 15 km or more, the total rate of expansion of the tropics in ERA-Interim is found to be 0.9° decade^{-1} in both DJF and JJA. Estimates from the product of sub-tropical height trends and gradients gives an expansion rate of 2.0° decade^{-1} in DJF and 1.6° decade^{-1} in JJA. When easterly wind tendencies are used, trends are calculated to be 1.0° decade^{-1} in DJF and 2.2° decade^{-1} in JJA.

When critical altitudes and wind tendencies are used to quantify the rate of tropical broadening $\sim$70% of the total broadening is due to expansion into the SH. However, when height tendencies are used, 72% of the total broadening is due to expansion into the SH in DJF, compared to 55% in JJA. Most of the widening of the zonal-mean tropical belt is due to large trends in the eastern hemisphere.

Chapter 6

Stratospheric deposition and lifetime of aviation water vapour emissions

Background water vapour concentrations in the atmosphere change markedly with height. In the stratosphere concentrations tend to be between 2 ppmv and 5 ppmv. Previous studies have shown that concentrations due to aviation emissions are approximately one order of magnitude smaller than the background concentration in the stratosphere, and several orders of magnitude less in the troposphere (Hoinka et al., 1993). However, Hoinka et al. also found that the lifetime of water vapour in the stratosphere is significantly larger than that in the troposphere, resulting in greater potential for the accumulation of emissions.

In this Chapter, the amount of direct deposition of emissions into the stratosphere will be quantified, along with its sensitivity to the tropopause definition and emissions inventory used (Section 6.1). A literature review of transport across the tropopause is presented in Section 6.2, before the fully Lagrangian trajectory model used in this thesis to model transport is introduced in Section 6.3. The choice of removal criteria to determine when emissions leave the stratosphere is discussed in Section 6.4.1. E-folding lifetimes calculated from the trajectories are presented in Section 6.4.2, and compared to the large range of estimates from previous studies.

6.1 Direct emission into the stratosphere

Previous work has focused on the percentage of both global (Gettelman, 1998, Grewe and Stenke, 2008) and NAFC emissions (Hoinka et al., 1993, Forster et al., 2003) that enter the stratosphere directly. Unfortunately, comparison of the conclusions from these works is very difficult, even for the global studies. They each use different emission and meteorological data sets for different years. They also use different tropopause definitions, which has previously been shown to influence the amount of emissions calculated to enter into the stratosphere, and consequently the radiative impact of the emissions (Gettelman, 1998, Forster et al., 2003). The lack of an accepted definition of the boundaries of the NAFC introduces further complications into the comparison of regional case studies.

Unsurprisingly, there is a large range in previous calculations of the percentage of emissions directly into the stratosphere. Previous studies suggest that anything between 13% and 44% of global column total emissions enter the stratosphere directly (Fortuin et al., 1995, Gettelman and Baughcum, 1999). For the NAFC the estimates lie between 33% (Hoinka et al., 1993) and 62-67% (Forster et al., 2003). In this case, however, it should be noted that Forster et al. only consider emissions at pressures less than 700 hPa. This means that, in the region they define as the NAFC, they are neglecting up to 20% of the column total water vapour emissions. Hence, their results suggest that 50-54% of NAFC column total emissions enter the stratosphere directly.

In the NAFC the tropopause tends to lie at 9 to 10 km, within the range of typical cruise altitudes. Therefore, it is expected that the percentage of column total emissions entering the stratosphere directly will be sensitive to the position of the tropopause. Changes in tropopause position can be the result of seasonal or inter-annual variations in the height of the tropopause, or be caused by changes in the tropopause definition or changes in the meteorological data set used (see Chapter 3). The differences in the vertical distribution of emissions within the different inventories (see Chapter 2) means that the percentage of emissions entering the stratosphere directly is also expected to be sensitive to the emissions inventory used.

The percentage of column total emissions in the NAFC entering the stratosphere directly is shown in Figure 6.1. The percentages shown in the upper panel are calculated using the blended tropopause surface every 6 hours for February 2004, and monthly total emissions from the three inventories, so that all the time-variation shown in the plot is due to changes in the height of the tropopause. Using the AEDT and QUANTIFY inventories results in a higher percentage of the emissions entering the stratosphere than when the AERO2k inventory is used. On average 61% of AEDT and QUANTIFY emissions enter the stratosphere directly, compared to 52% of AERO2k emissions. This is due to the higher cruise altitude (AEDT), and higher proportion of time spent at cruise altitude (QUANTIFY), given in those inventories (Chapter 2). The same calculation is performed for the plot shown in the lower panel, but in this case 6 hourly totals are used from the AEDT inventory. The large amplitude of the diurnal cycle in North Atlantic traffic is clearly reflected in the deposition of emissions into the stratosphere. In times of peak traffic, in excess of 90% of emissions enter the stratosphere directly, with several February maxima in excess of 80% of column total emissions. Including the diurnal variation in traffic and, to some extent, the effects of flight routing, increases the percentage of AEDT emissions entering the stratosphere directly from 61% to 62%.

Gettelman (1998) also investigated the sensitivity of the percentage of emissions entering the stratosphere directly to the emissions inventory used. Comparing the NASA and ANCAT inventories on a global basis, he concluded that the percentage was ‘not substantially affected’ by the inventory used. He found that using the two different inventories lead to a difference of 5% of column total emissions entering the stratosphere directly. This is roughly half of the difference between using the AEDT inventory and the AERO2k inventory found here. Although 10% may only represent a small difference in the mass of the emissions deposited directly into the stratosphere, it may not be an insignificant difference when the height distribution of these emissions in the stratosphere, and the influence of emissions height on emissions lifetime and therefore accumulation, is considered. The percentage of AEDT emissions entering the stratosphere directly is higher than that for the AERO2k inventory because of the higher cruise altitudes found in the AEDT inventory. It is therefore reasonable to expect that the AEDT emissions

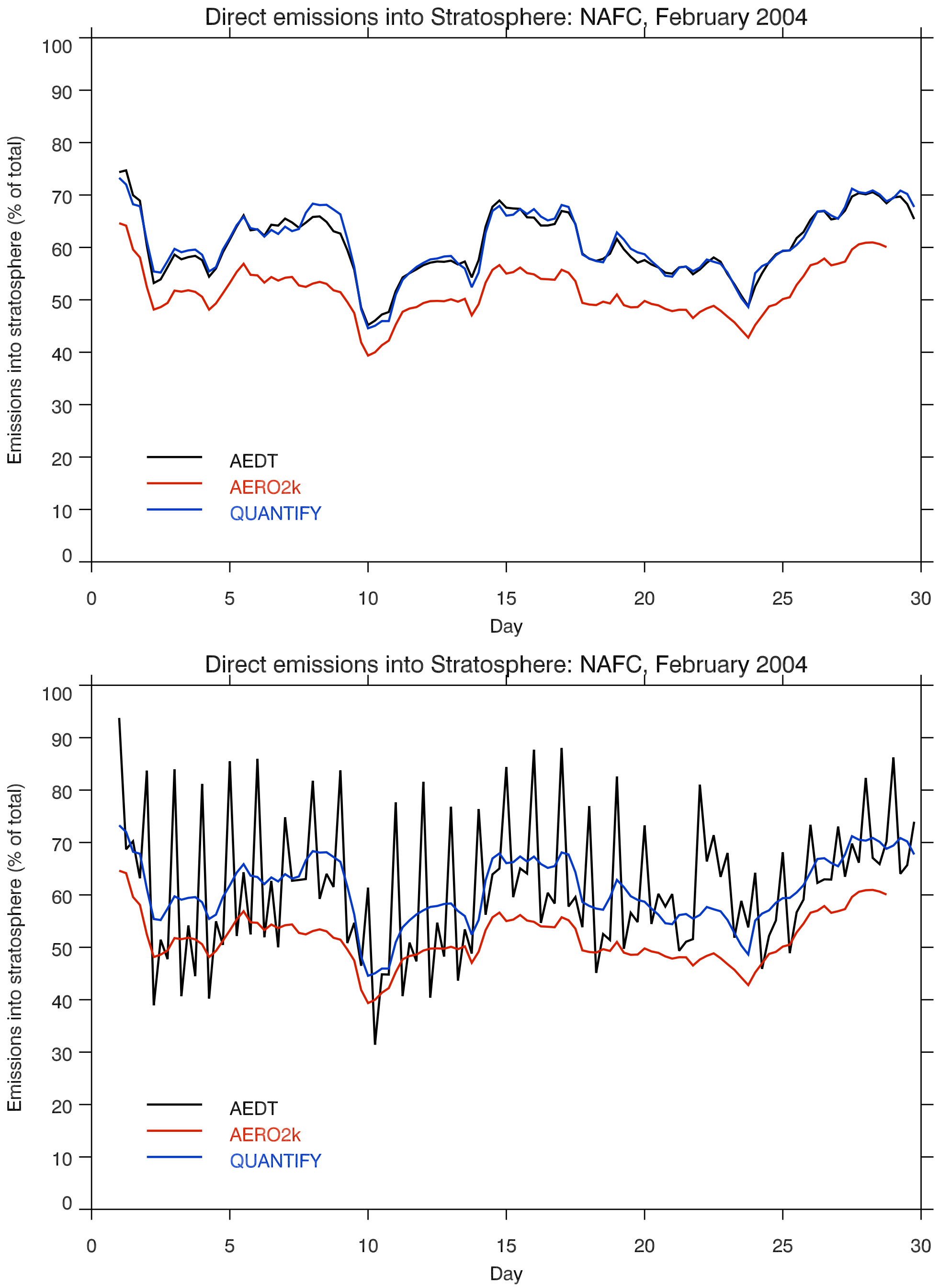


Figure 6.1: *The percentage of emissions entering the stratosphere directly for the NAFC with February 2004 meteorology, calculated using the blended tropopause and the AEDT, AERO2k, and QUANTIFY inventories. The results using monthly mean emissions and a 6 hourly tropopause are shown in the upper panel. In the lower panel 6 hourly AEDT emissions and 6 hourly meteorology are used.*

will have a longer lifetime in the stratosphere than the AERO2k emissions, leading to greater accumulation of AEDT emissions. Different accumulations will in turn lead to the calculation of different climate impacts when the two inventories are used, which may in the end be more significant than the differences in the percentage of stratospheric emissions.

No information is available about the spatial distribution of emissions within an inventory grid box. This means that assumptions have to be made about the distribution of emissions when calculating the percentage of emissions that enter the stratosphere. When the vertical resolution of the inventory is coarse, this assumption can have a substantial impact on stratospheric deposition. This can be seen in Figure 6.2 where the blended tropopause has been used to calculate the percentage of AEDT emissions entering the stratosphere directly. The AEDT inventory has been used with a vertical resolution of 1 km and 152 m, with the two extreme assumptions of placing all the emissions at either the upper bound or the lower bound of each grid box. When the low vertical resolution inventory is used the difference between the upper and lower bound calculations can be as much as 20% of column total emissions. The difference is no more than 2% of the column total for the high resolution data, making a strong case for the use of high vertical resolution inventories.

A comparison was made between the results from calculations with the emissions located at the upper and lower bounds of each grid box, and in the centre. This demonstrated that the percentage of column total emissions calculated using the assumption that emissions were located at the centre of each grid box did not necessarily give the same result as the mean of the upper and lower bound cases. As such, it has been assumed in all calculations of direct deposition that the emissions are located at the upper bound of each grid box. This gives a definite upper limit for each calculation, while it would be difficult to determine where an estimate using emissions placed at the centre of each grid box lay within the range of uncertainty. As most calculations in this thesis are performed using high vertical resolution AEDT data, this is unlikely to have a noticeable impact on the final result. However, it becomes more important as the vertical resolution of the emissions inventory decreases.

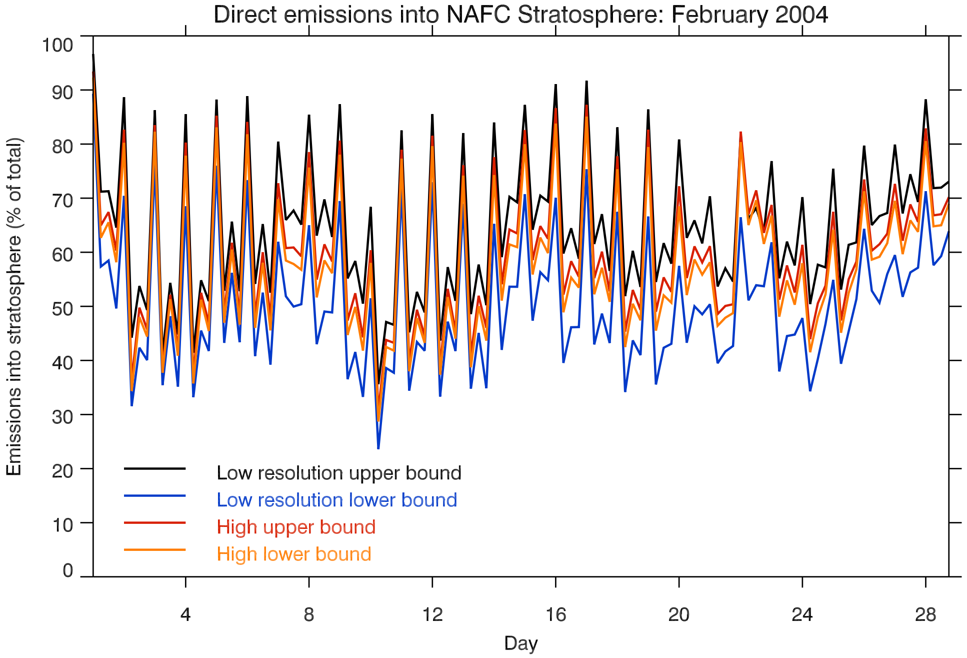


Figure 6.2: *The percentage of emissions entering directly into the stratosphere in the NAFC in February 2004, calculated using the blended tropopause and the AEDT inventory at low (1 km) and high (152m) vertical resolution. 'Upper bound' and 'lower bound' refer to results where it has been assumed that all emissions within a grid box have been placed at the highest possible and lowest possible altitudes respectively.*

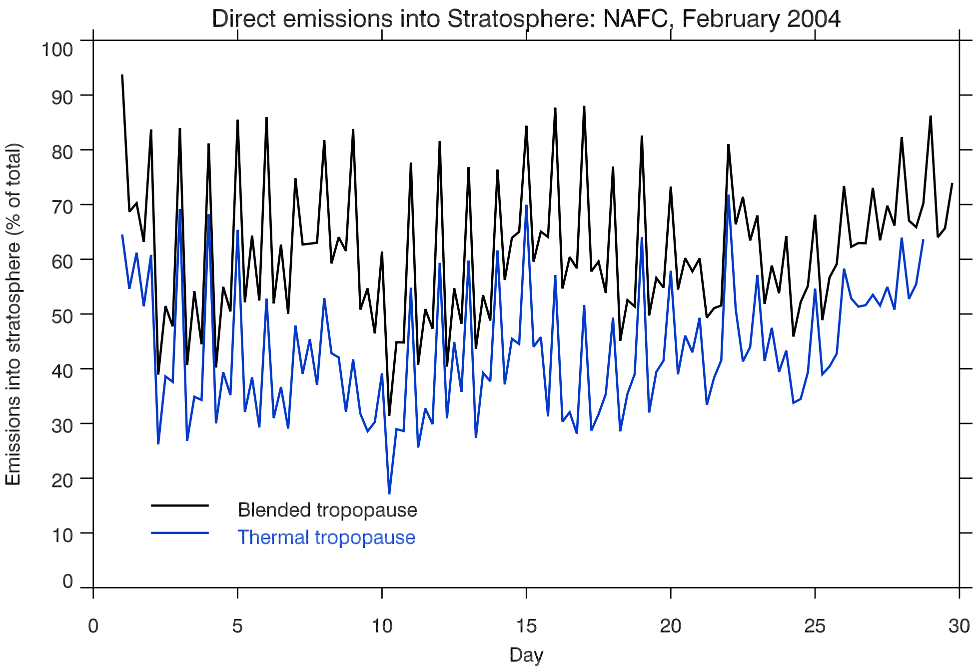


Figure 6.3: *The percentage of emissions entering the stratosphere directly in the NAFC with February 2004 meteorology, calculated using 6 hourly AEDT emissions, and a 6 hourly thermal and blended tropopause.*

The percentage of AEDT emissions entering the NAFC stratosphere, calculated using a blended tropopause and thermal tropopause definition for February 2004 is shown in Figure 6.3. There is a clear sensitivity of the percentage of direct emission in the stratosphere to the tropopause definition used. When the blended tropopause definition is used, an average of 62% of AEDT emissions enter the stratosphere. When the thermal tropopause definition is used the NAFC tropopause is, on average, 800m higher, and the average percentage of emissions entering the stratosphere directly is only 44%.

Gettelman (1998) and Gettelman and Baughcum (1999) calculated the percentage of emissions entering the stratosphere directly using the NASA 1992 inventory, and a 2, 3, and 5 PVU tropopause, and 1, 2, 3, 4 PVU and thermal tropopause respectively. In both studies it was found that stratospheric deposition was sensitive to tropopause definition. Gettelman (1998) found that 17-35% of global emissions entered the stratosphere directly in January, and 10-24% entered the stratosphere directly in July, depending on the tropopause definition used. Gettelman and Baughcum (1999) found that 18-44% of global emissions entered the stratosphere directly, in the annual average, depending on the tropopause definition used.

For the |PV|=2 PVU tropopause, Gettelman (1998) found that 35% of global emissions were emitted directly into the stratosphere in January 1992, and 24% in July 1992. Using the ERA-Operational tropopause and Gettelman's NASA inventory gives 37% and 23% for January and July respectively (not shown). This is slightly larger than the values found here using the global AEDT inventory: 32% and 19%. There is a slight difference between the NASA values found here, and those presented by Gettelman and Baughcum, which is due to the different meteorological data used. Gettelman and Baughcum concluded that the percentage of emissions entering the stratosphere directly is not sensitive to the meteorological data set used. Comparison of the values found here with their result suggests that there is a small sensitivity (4-5%), but that it is not as large as the sensitivity to the emissions inventory used.

The higher percentage of column total NASA emissions entering the stratosphere compared to those from AEDT is due to the combination of the higher cruise altitude and proportion of time spent at cruise altitude in the NASA inventory compared to AEDT (see Figure 6.4). However, the NASA vertical profile is the most similar to those from AEDT out of all the inventories considered (Figure 6.4). The influence of the differences in the vertical profiles are more pronounced when only the NAFC is considered, with a monthly mean of 70% of NASA emissions entering the stratosphere compared to 61% for AEDT (Figure 6.4).

The variation of the percentage of emissions entering the stratosphere with seasonal and annual variations in tropopause height has also been calculated, by allowing the tropopause to vary but keeping emissions fixed. The percentage of emissions entering the stratosphere directly in the NAFC was calculated using AEDT total emissions for February 2006, and a 6 hourly blended tropopause for the whole of 2006 (see Figure 6.5). When emissions are kept fixed, there is a maximum of 64% of emissions entering the stratosphere directly in March, and a minimum of 38% in June.

When AEDT monthly total emissions are also allowed to vary the seasonal cycle of stratospheric emissions is almost identical (Figure 6.5). Although there is a seasonal cy-

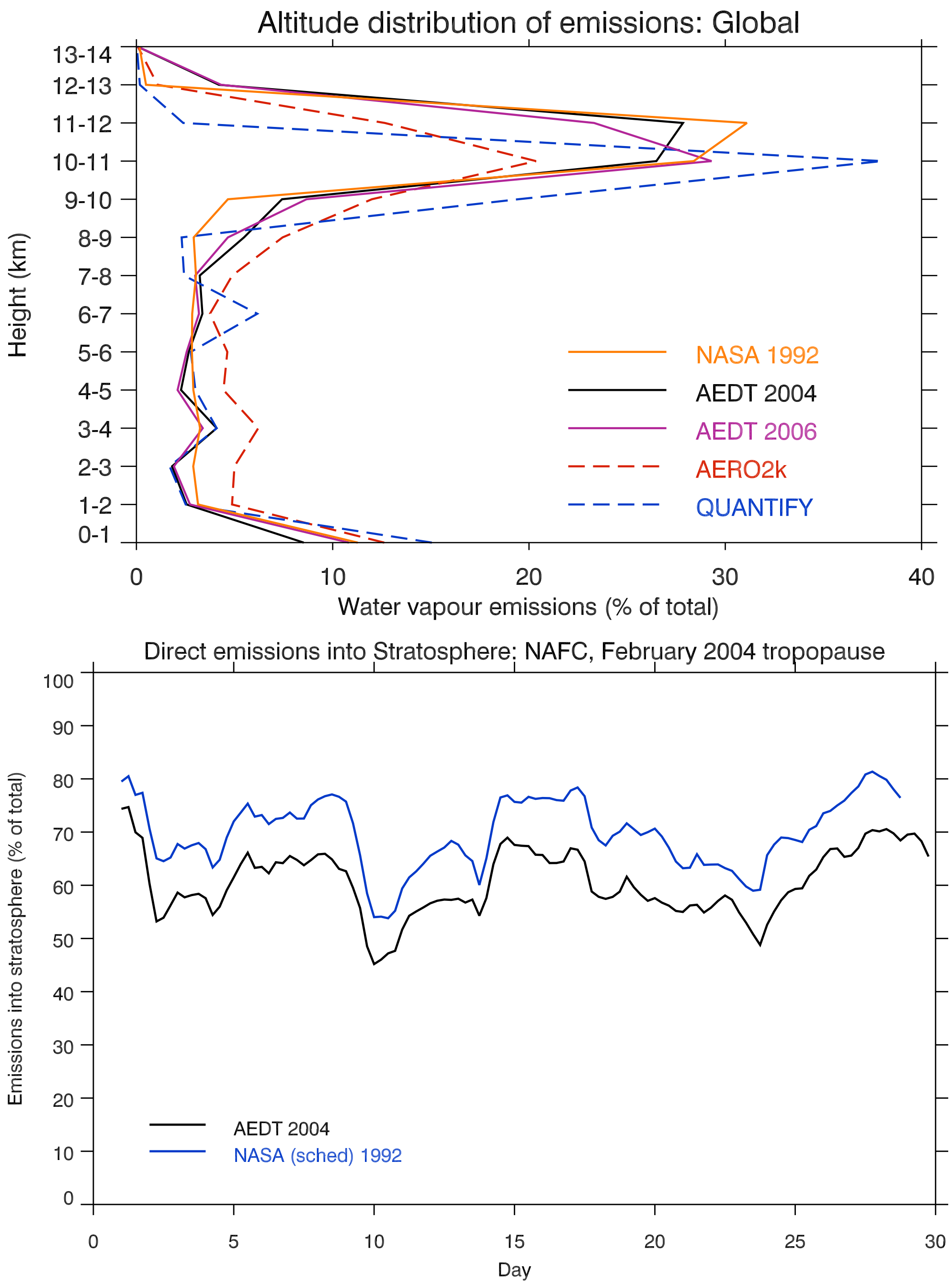


Figure 6.4: *The vertical profile of global emissions from the NASA inventory, compared to the inventories introduced in Chapter 2 (top panel). The percentage of column total NAFC emissions entering the stratosphere directly from the NASA and AEDT inventories (monthly total emissions), using the blended tropopause for February 2004 (lower panel).*

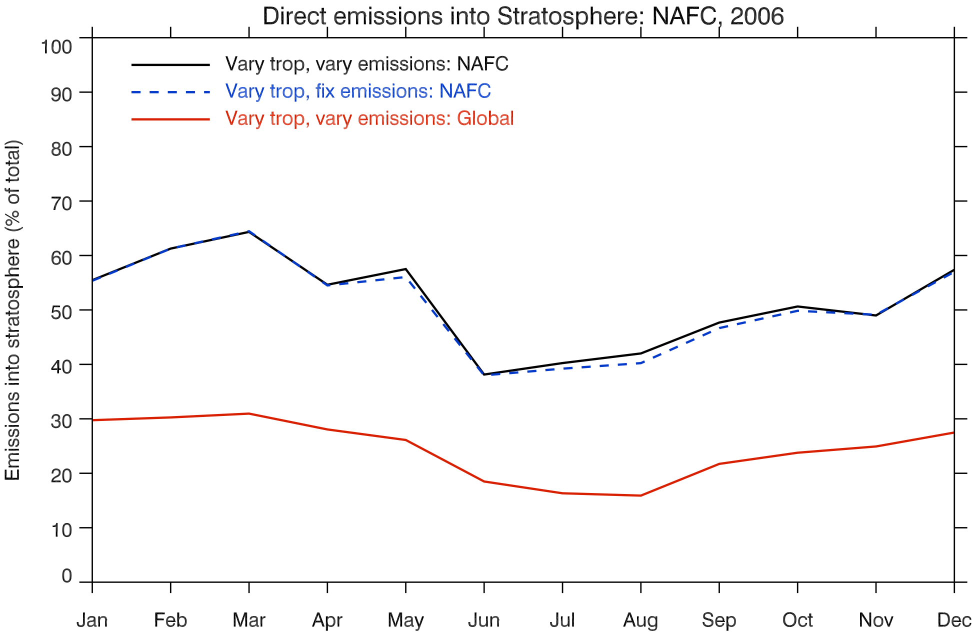


Figure 6.5: *Monthly mean emissions into the stratosphere in 2006. The black line shows the NAFC deposition calculated using 6 hourly emissions and a 6 hourly tropopause. The dashed blue line shows the same, but using February emissions. The red line shows the global deposition calculated using 6 hourly emissions and a 6 hourly tropopause. In each case the emissions are from AEDT, and the blended tropopause is used.*

cle in the total mass of emissions, the vertical distribution of the emissions is relatively invariant with season, hence the seasonal cycle in the percentage of column total emissions entering the stratosphere. Repeating the calculation for the globe reveals a similar seasonal cycle, albeit with a smaller amplitude, as the majority of global emissions enter the atmosphere in the NH extratropics (Figure 6.5). For global emissions, a maximum of 31% of total emissions enter the stratosphere directly in March, with a minimum of 16% in August. The amplitude of this cycle is close to the 10% identified by Gettelman and Baughcum (1999), and both the maximum and minimum values lie within the range of values presented in previous studies.

Hoinka et al. (1993) found no correlation between the percentage of NAFC emissions entering the stratosphere directly and the NAO index. As the phase of the NAO influences the gradient of the tropopause across the NAFC, the relationship between the NAO index and the stratospheric deposition of NAFC emissions was also investigated in this thesis. Inter-annual variations of up to 20% of column total emissions were identified in winter (Figure 6.6). However, there was no significant correlation between the NAO index and stratospheric deposition when AEDT and ERA data were used. Even when the three winters with the highest NAO indices since 1989 were compared with the three winters with the lowest indices, there was no significant correlation (see Figure 6.6).

No strong correlations were found between the NAO index and stratospheric deposition in the NAFC. Although the NAO influences the gradient of the tropopause across the NAFC, it does not necessarily influence the average tropopause height. The monthly-mean percentage of column total emissions emitted into the stratosphere directly was found to be correlated with the monthly-mean tropopause height, with r^2=0.88.

Over 20% of global emissions occur in the NAFC, and it was found that global deposition is also correlated with the average height of the NAFC tropopause, with r^2=0.94. This relationship is stronger than that between global deposition and global mean tropopause, which has r^2=0.56.

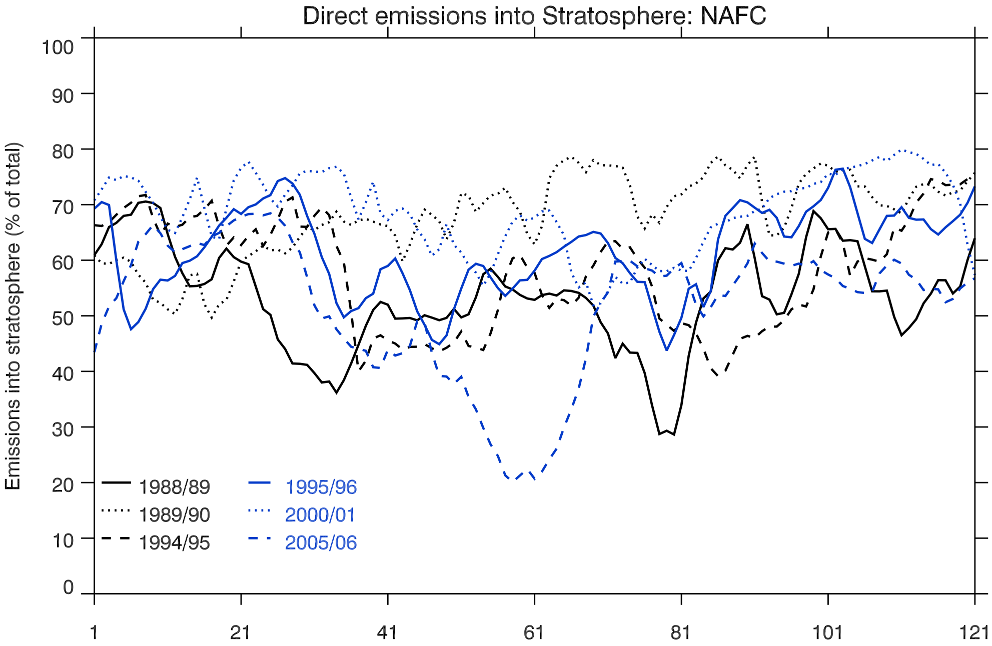


Figure 6.6: *Direct emission into the NAFC stratosphere using monthly mean AEDT emissions and a 6 hourly blended tropopause for DJF for selected years. Winters with a strongly positive NAO index are shown in black, and those with a strongly negative NAO index are shown in blue.*

6.2 Transport of emissions in the stratosphere and across the tropopause: a review

The circulation in the upper troposphere and stratosphere generally consists of upwelling from the troposphere to the stratosphere in the tropics, transport to the extratropics in the stratosphere and downward mass flux from the stratosphere to troposphere in the middle and high latitudes. The zonal average of this circulation is the Brewer-Dobson circulation. The Brewer-Dobson circulation is driven by Rossby wave and gravity wave breaking in what is often referred to as the ‘extratropical pump’ (Plumb and Eluszkiewicz, 1999).

When considering transport from the stratosphere to the troposphere it is useful to divide the lower stratosphere into two regions, overworld and middleworld (Hoskins, 1991). Overworld is the region at potential temperatures above 380 K. It is in the stratosphere at all latitudes, meaning that transport from overworld to the troposphere is slow as isentropic surfaces have to be crossed, which requires diabatic heating. Middleworld is the region between the tropopause and 380K. Isentropic surfaces in this region are in the stratosphere at high latitudes, but in the troposphere at low latitudes. This allows fast transport between the stratosphere and troposphere to take place along isentropic surfaces.

The descending branch of the Brewer-Dobson circulation means that most emissions into the extratropical stratosphere will be transported into the vicinity of middleworld and the tropopause. Once within the tropopause region, various processes contribute the the transportation of emissions from the stratosphere to the troposphere. In the extratropics, stratosphere to troposphere exchange (STE) is associated with synoptic- and mesoscale-processes. Transport from the stratosphere has been observed in tropopause folds (near both the polar- and sub-polar-jets), cut-off lows, mesoscale convective complexes, thunderstorms, and due to breaking gravity waves and the filamentation and fragmentation of stratospheric streamers (Stohl et al., 2003, Sprenger et al., 2007). Cut-off lows and tropopause folds in particular result in the exchange of large volumes of air as they are associated with large latitudinal displacements on isentropic surfaces (Stohl et al., 2003).

Extrusions from the stratosphere into the troposphere tend to be organised within small-scale filamentary structures that are produced by shear in the large scale wind fields. The lifetime of such filaments is typically a few days (Forster and Wirth, 2000). The link between STE and PV structures, such as filaments and streamers, is particularly strong in the extratropics, with 60-80% of STE events occurring in the vicinity of a PV structure, irrespective of season (Sprenger et al., 2007). Larger structures are more likely to be associated with STE.

The structures that tend to be associated with STE have preferred geographical regions. Sprenger and Wernli (2003) found that maximum stratosphere to troposphere transport (STT) in the Northern Hemisphere occurred between 40°N and 60°N. Wernli and Bourqui (2002) also found that maximum NH STE was within the storm tracks over the Atlantic and Pacific oceans. Stohl et al. (2003) are more specific, and link STT to the cyclogenesis regions: the North American west coast, the beginning of the North Atlantic storm track and the Mediterranean. Stohl et al. (2003-STE) note that there is less geographical variation associated with shallow exchange events.

Wernli and Bourqui (2002) and Stohl et al. (2003) have both identified an annual cycle in cross tropopause mass fluxes, with a maximum in winter and early spring and a minimum in late summer. STT is often divided into shallow and deep exchanges, with most STT events being shallow and short lived (Stohl et al., 2003). The amplitude of the seasonal cycle for deep STT events is much more pronounced than the amplitude of the cycle of all events (Stohl et al., 2003-STE). There is very little annual variation in shallow STT events (James et al., 2003).

In addition to an annual cycle in the amplitude of STT events, Sprenger and Wernli (2003) identified a geographical cycle in STT. They found that STT was strongest over the storm tracks in winter, and strongest over the continents in summer. Wernli and Bourqui (2002) also identified an annual cycle in the range of potential temperatures where STE was most likely to occur. They found that in winter STE tended to occur in the $285 < \theta < 320$ K isentropic layer, with almost no exchange occurring above 340 K. In summer, most STE occurred at potential temperatures between 310 K and 340 K. This difference is due to the annual cycle in the location of isentropes, rather than a change in the physical location of STE events.

Inter-annual variations in STE have also been inferred. Stohl et al. (2003) found that tropopause crossings in the North Atlantic region occurred at higher latitudes and altitudes in winters with a positive NAO index compared to other years. They also identified an intensification and poleward shift of STE in the eastern tropical Pacific associated with El Nino. Sprenger and Wernli (2003) also identified differences in STT in agreement with the differences in cyclone frequency associated with different phases of the NAO. However, they found that troposphere to stratosphere exchange was only weakly correlated with the NAO index. They note that there was only small inter-annual variability in the seasonal cycles of hemispherically integrated exchange, and that there do not appear to have been any trends in transport across the ERA-15 period (1979-1993).

6.3 Trajectory modelling

It is possible to model STE using either Eulerian or Lagrangian methods. Most General Circulation Models (GCMs) and Chemistry Transport Models (CTMs) are Eulerian models. In such models the flow of air through a fixed grid is studied. This means that they may not give a very accurate simulation of conservative properties and cannot be used to answer questions about a particular air mass. This is possible using Lagrangian models, where labelled parcels of air are followed in time. It is this exact computation of transport that is the main advantage of using Lagrangian models. However, this is closely linked to their main disadvantage: the need to follow a very large number of parcels in order to resolve important features of the circulation. The different methods are known to give different estimates of net fluxes and distribution (Wernli and Bourqui, 2002).

Stohl et al. (2003) compared trajectory models, Lagrangian particle models, Eulerian transport models and GCMs. They found that all models were able to capture an STT event, but that the range of magnitudes of tracer fluxes was large. Tracer fluxes in the GCMs were around four times larger than those in the Lagrangian models. They found that the use of Lagrangian models resulted in an underestimate of STT as the models sometimes failed to capture the presence of stratospheric air in the troposphere due to the absence of convection and turbulence schemes. Conversely, they found that GCMs overestimated the flux as they dispersed the tracer too deeply into the troposphere, and over geographical regions that were too large when compared to observations. They concluded that this is likely to be an artefact of numerical diffusion, compounded by low model resolution and strong tracer gradients across the tropopause. Stohl et al. (2003-STE) note that Eulerian estimates are highly sensitive to intrinsic parameters, resulting in poor agreement between different estimates. Eulerian methods also treat short- and long-lived excursions in the same way, which can result in high noise levels (Stohl et al., 2003-STE).

Although there are advantages and disadvantages to the use of both types of models, the advantages of using a Lagrangian model far outweigh those of using a Eulerian model.

There is no numerical diffusion in a Lagrangian model, and they are often less computationally expensive to achieve results an an equivalent resolution to a Eulerian model. They are strictly mass conserving, and they the strong gradients of trace species found in the vicinity of the tropopause, unlike semi-Lagrangian schemes, which smooth such gradients. However, the underestimation of STT in Lagrangian models identified by Stohl et al. (2003) motivates the use of high resolution wind data, and a high resolution release grid.

6.3.1 The UGAMP trajectory model

The Universities Global Atmospheric Modelling Project (UGAMP) trajectory model is a fully Lagrangian model, written by John Methven (Methven, 1997, Methven et al., 2003). Details of the model's operation and accuracy are given in Methven (Methven, 1997).

Using ECMWF spectral wind data, the trajectory model can be used to produce either forward or backward trajectories by solving the initial value problem,

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{u}(\boldsymbol{x}, t) \tag{6.3.1}$$

using a 4th order Runge-Kutta method:

$$k_1 = u(x_n, t_n)\delta t \tag{6.3.2}$$

$$k_2 = u(x_n + \frac{k_1}{2}, t_n + \frac{\delta t}{2})\delta t \tag{6.3.3}$$

$$k_3 = u(x_n + \frac{k_2}{2}, t_n + \frac{\delta t}{2})\delta t \tag{6.3.4}$$

$$k_4 = u(x_n + k_3, t_n + \delta t)\delta t \tag{6.3.5}$$

$$x_{n+1} = x_n + \frac{k_1}{6} + \frac{k_2}{3} + \frac{k_3}{3} + \frac{k_4}{6} \tag{6.3.6}$$

where **u** is velocity, **x** is position, t is time, and δt is a constant integrator time-step that is much smaller than the interval between the wind records. Methven (1999) found this method to be more accurate, and stable, than using the simpler Euler method.

Trajectories are integrated on a sphere in spherical co-ordinates, using the shallow atmosphere approximation. Horizontal wind components are taken directly from ECMWF wind data. Vertical wind is found from mass conservation via the continuity equation. The value of a given field is found at the position of each trajectory particle by interpolating the gridded data. Bilinear interpolation is used in the horizontal. In the vertical, Lagrange interpolation is used to capture the large gradients at the tropopause.

A step size of one hour has been used in the Runge-Kutta scheme in this work. This is sufficiently small compared to the 6 hours between wind records that the integration errors in the model will be negligible compared to field truncation errors. Methven et al. (2003) found that the model was able to accurately simulate photochemically produced ozone compared to observations in thermodynamic co-ordinates. They found that typical displacement errors associated with the trajectories were $\sim$30 km.

6.3.2 Example trajectory output

It is possible, using the trajectory model, to interpolate various parameters to particle position, and store them as ‘attributes’. For this work, specific humidity, potential temperature, and potential vorticity were recorded, in addition to particle pressure, latitude, and longitude. This allows the position of each trajectory particle relative to the tropopause to be calculated. It also gives an indication of background humidity of particle position.

Examples of trajectory model output are shown in Figures 6.7 and 6.8. In these examples, 25 particles were released on a regular latitude-longitude grid, bounded by 50°N, 48°N, 30°W and 28°W. It can be seen in Figure 6.7 that on this occasion the particles initially follow very similar horizontal paths, although there is a large spread in their initial PV values. There are two distinct groups of trajectories, those with lower values of PV (warm colours), and those with higher initial values (cool colours). Most trajectories with lower initial values of PV at their release descend steadily and enter the troposphere, or the immediate vicinity of the tropopause. The trajectories with higher initial values of PV remain at higher levels, after an initial fluctuation in this example where PV values dip to tropospheric values, and then rapidly revert to stratospheric values. It can be seen in the upper panel of Figure 6.7 that there are ‘kinks’ in some of the warmer coloured trajectories, particularly over Asia. These kinks are associated with the tropopause crossings that can be seen in the lower panel of Figure 6.7, and in Figure 6.8.

Similar releases of small groups of particles were made at different times and pressures, and from different locations (not shown). The results of these preliminary experiments illustrated that the trajectory of a particle is dependent on the pressure from which it is released, and on local meteorology. Generally, increased initial distance above the tropopause resulted in a tighter clustering of trajectories in both the horizontal and PV. Particles released at greater altitudes tended to travel further in a given time, and follow more zonal paths.

6.4 Lifetime estimates

While many previous studies have produced estimates of the stratospheric lifetime of water vapour emissions, there is little agreement between studies (Table 6.1). Generally an e-folding lifetime is calculated, and e-folding lifetimes have been presented in Section 6.4.2 for comparison with other work. However, analysis of the trajectories used to calculate these lifetimes has demonstrated that an e-folding lifetime is not always appropriate (Section 6.4.3). Following this analysis, a methodology was developed to allow the calculation of the accumulated mass of aviation water vapour emissions that does not rely on assumptions about the stratospheric lifetime of emissions (Chapter 7). Before any lifetimes could be calculated, it was necessary to define when a particle released into the stratosphere could no longer be considered to have stratospheric properties.

6.4.1 Defining removal criteria

As aviation water vapour emissions are negligible when compared to the background water vapour concentration in the troposphere, this thesis focuses only on stratospheric

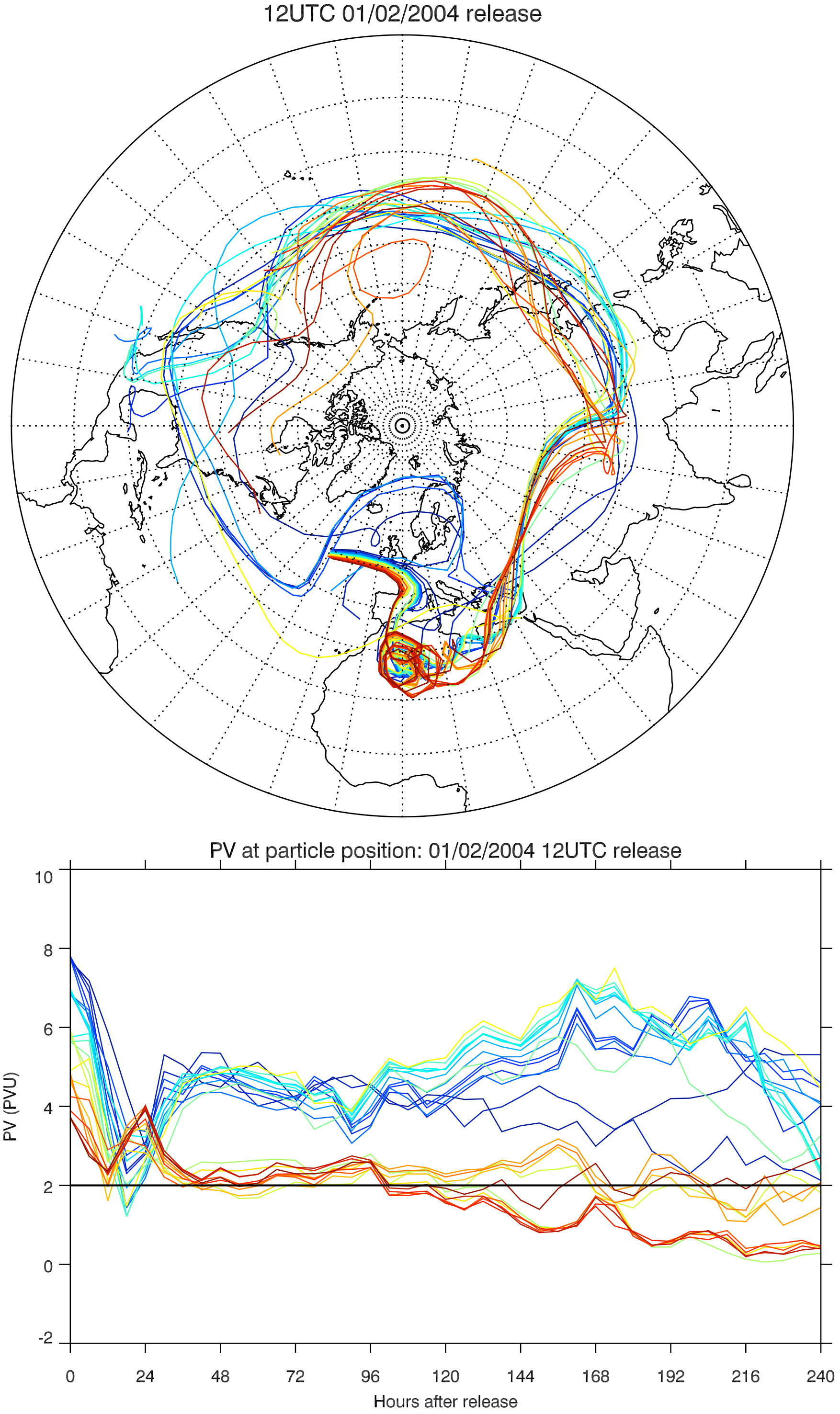


Figure 6.7: *The paths followed by particles in the first 10 days following their release at 200 hPa in the region bounded by 50°N, 48°N, 30°W and 28°W on 1 February 2004, 12Z. Each coloured line tracks the path of an individual particle. Their horizontal position is shown in the upper panel, while their PV is shown in the lower panel. Here, the approximate tropopause position is indicated by the heavy black line.*

12UTC 01/02/2004 release

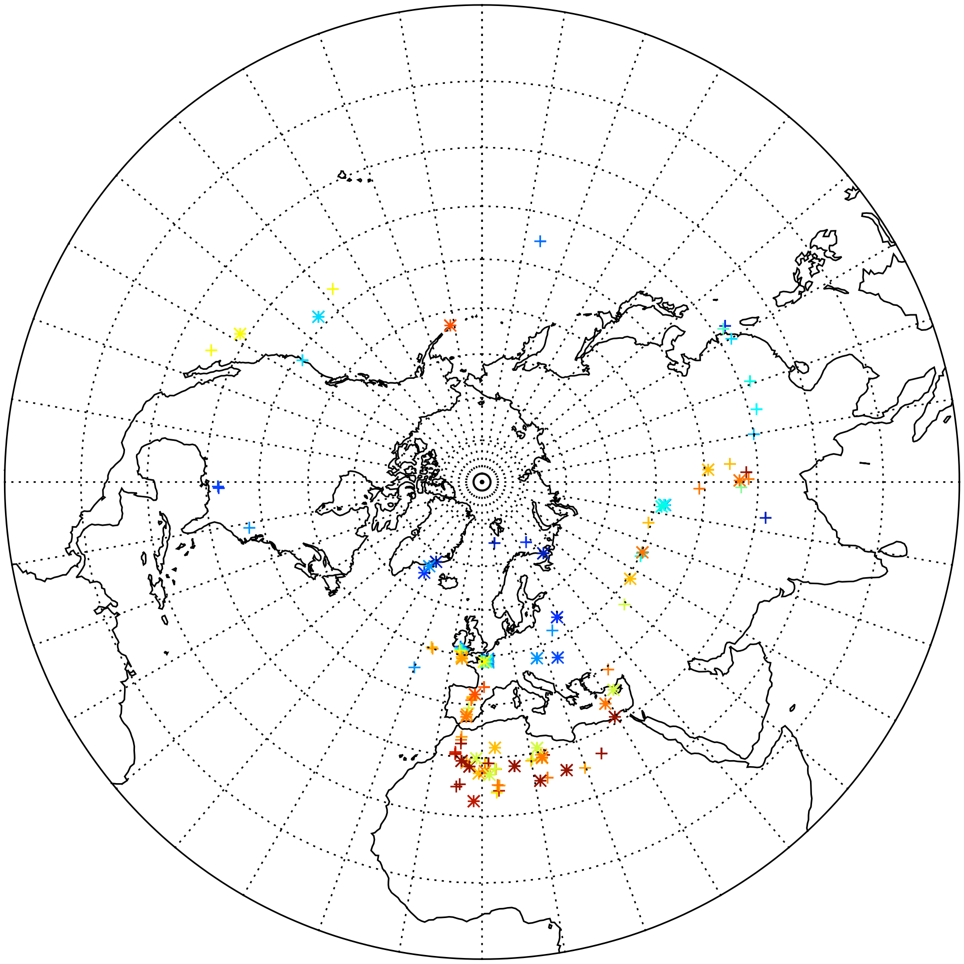


Figure 6.8: *The points where tropopause crossings occur. The colours correspond to the individual trajectory colours used in Figure 6.7. Crosses indicate downwards crossings, and stars indicate upwards crossings.*

Study	Release location	Lifetime
Hoinka et al., 1993	Stratosphere	6 months
Fortuin et al., 1995	Stratosphere	6 to 30 months
Ponater et al., 1996	Lower stratosphere	2 to 6 months
Schoeberl et al., 1998	11 km	∼0.5 months
	13, 14, 16, 19 km	∼12 months
Gettelman, 1998	Lowermost stratosphere	1.3 to 1.9 months depending on tropopause definition and season
Schoeberl and Morris, 2000	Middleworld	Order 3 months
	Overworld	Order 12 months
Forster et al., 2003	Stratosphere	0.7 months
Grewe and Stenke, 2008	52 hPa (∼20 km)	13 months
	89 hPa (∼17 km)	9 months
	132 hPa (∼14.5 km)	4.3 months
	198 hPa (∼12 km)	1 month
This thesis, ERA-Operational	300 hPa (9.2 km)	0.5 months
	225 hPa (11.0 km)	1.3 months
	200 hPa (11.8 km)	2.1 months
	150 hPa (13.6 km)	8.2 months
This thesis, ERA-Interim	300 hPa (9.2 km)	0.4 months
	225 hPa (11.0 km)	1.8 months
	200 hPa (11.8 km)	3.2 months
	150 hPa (13.6 km)	22.1 months

Table 6.1: *Estimates of the stratospheric lifetime of water vapour emissions from previous studies. The majority of the studies use different assumptions and methodologies, which influences the calculated lifetime.*

water vapour emissions. As such, it is necessary to define when a particle trajectory has left the stratosphere. Preliminary trajectory experiments revealed that most particles make multiple crossings of the tropopause before they are permanently removed from the stratosphere (see Figure 6.7)

There is no consensus amongst previous studies as to how it is best to define when a particle has been removed from stratosphere. Many definitions have been used, and the majority of studies use either temporal criteria, or a critical distance below the tropopause. Wernli and Bourqui (2002) investigated the use of 24, 48, 72 and 96 hour removal criteria, finding little difference between 48 to 96 hour criteria. Sprenger et al. (2003) defined an exchange event if a trajectory spent 24 hours in the stratosphere, followed by 24 hours in the tropopause. Sprenger et al. (2007) used a similar criteria, but with a critical value of 96 hours. Schoeberl and Morris (2000), and Morris et al. (2003), defined water vapour as being removed from the stratosphere when it reached 0.5 km below the tropopause.

It was decided to use a temporal removal criterion in this work, as this allows for some representation of the mixing that occurs when stratospheric air enters the troposphere. A wide range of temporal removal criteria have been used in previous work, ranging from instant removal (Forster et al., 2003) to removal after a particle spends 96 consecutive hours in the troposphere (Wernli and Bourqui, 2002). The final choice of removal criterion in this thesis will be based on the consideration of observations of mixing timescales at the tropopause, and the typical length of time spent in the troposphere by a particle making a temporary excursion below the tropopause. However, the choice of criterion is rather subjective.

A case study of the motion of particles near the tropopause after release in the NAFC was used to investigate the transition from the stratosphere to the troposphere. 50 day long trajectories were released on single pressure levels every 24 hours for January 2004 from a 2°×2° grid bounded by 40°N, 65°N, 70°W, and 0°W. Releases were made every 10 hPa between 300 hPa and 150 hPa ($\sim$ 9.2 km and 13.6 km), and also at 225 hPa ($\sim$11

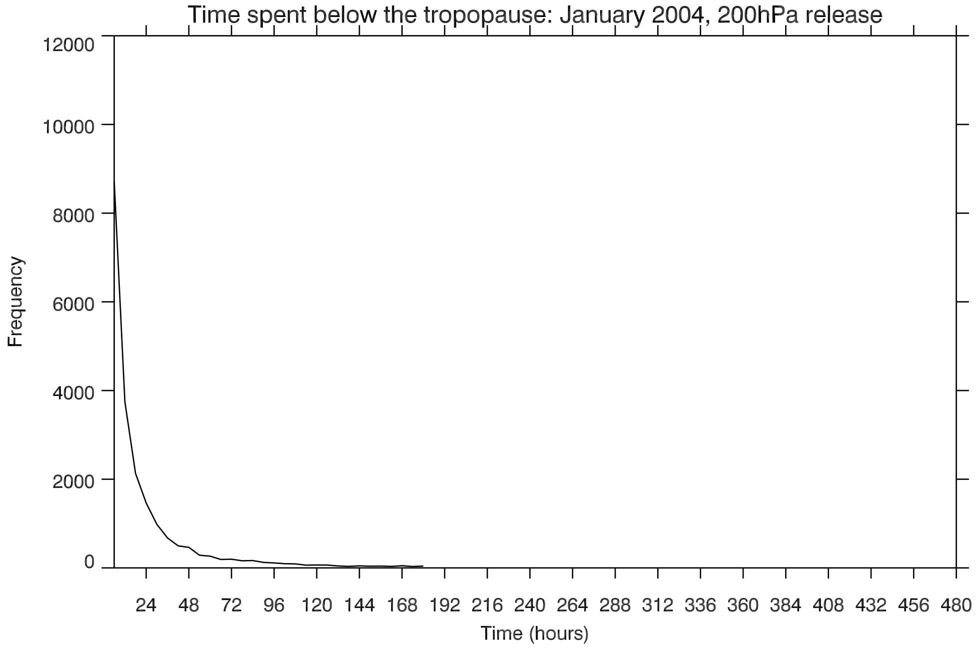


Figure 6.9: *The length of time spent below the tropopause by particles making short excursions into the troposphere from the stratosphere. Particles were released at 200 hPa on a regular grid covering the NAFC in January 2004.*

km), which is the median cruising pressure.[1] Both the dynamic and blended tropopause definitions were used to determine when each individual particle crossed the tropopause. The direction and location of each crossing was recorded, along with the time spent below the tropopause by particles that subsequently returned to the stratosphere.

It was found that most particles made multiple crossings of the tropopause following their release into the NAFC stratosphere. The number of excursions into the troposphere of a given length of time are shown in Figure 6.9. It can be seen from Figure 6.9 that most particles that make repeated crossings of the tropopause remain in the troposphere for 48 hours or less, until their final crossing. This result is true for all the release pressures tested, and for both the blended and dynamic tropopause definitions. A similar result was presented by Stohl et al. (2003) who found that in excess of 90% of STT particles in the troposphere stayed there for less than 6 hours. Therefore, a removal criterion of longer than 48 hours is not relevant.

---

[1] Trajectories are released on pressure levels. The heights given in brackets are the corresponding pressure-altitudes, as used in gridded AEDT data. All altitudes given in this Chapter are pressure-altitudes.

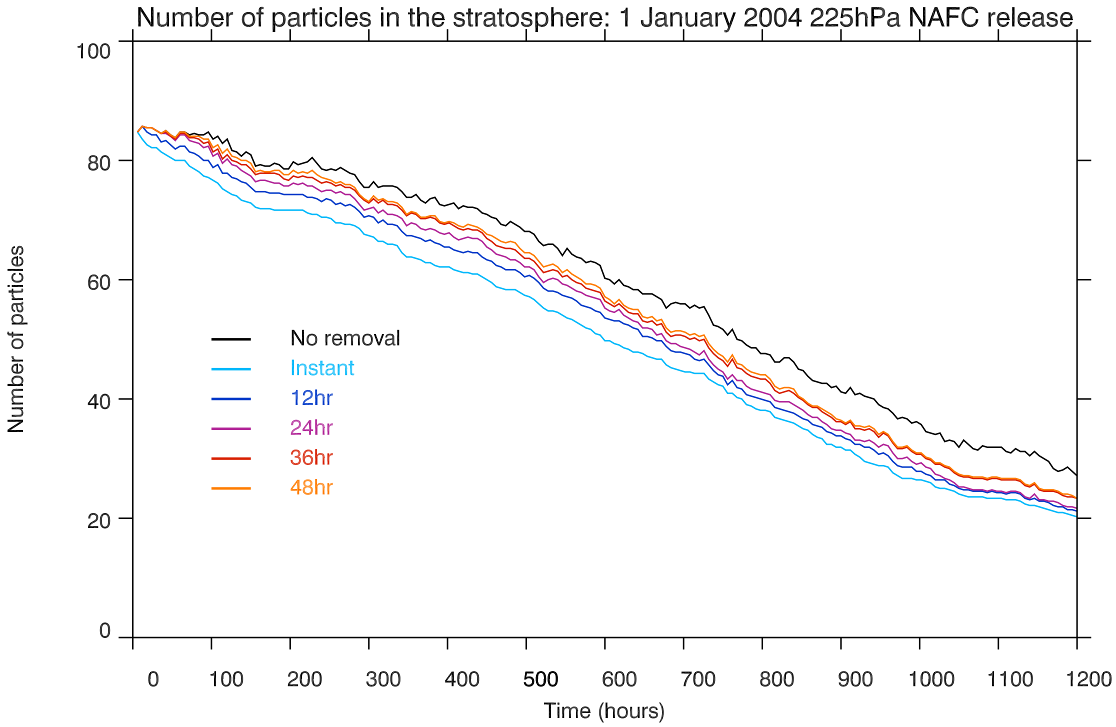


Figure 6.10: *The number of particles in the stratosphere with time after their release at 225 hPa in the NAFC in January 2004. Colours indicate the temporal removal criteria used. The black line, 'no removal', indicates the instantaneous total number of particles located above the tropopause at each time when no temporal criterion has been applied. The light blue line indicates the number of particles in the stratosphere when an instantaneous removal criterion is applied. The blue, purple, red, and orange lines show the result when a 12, 24, 26, and 48 hour removal criterion was used.*

The sensitivity of the number of particles remaining in the stratosphere at a given time after their release to different temporal removal criteria was also tested. Instant removal, and removal after 12, 24, 36, 48 and 96 hours of being in the troposphere were considered. There were large differences in the number of particles remaining in the stratosphere when instant removal, and removal after 12 hours were used (see Figure 6.10). The difference in the results found from using the different removal criteria decreased with increased length of time in the troposphere. The results from the 36 and 48 hour criteria are almost identical. This result suggests that a temporal criterion of 36 hours or less should be used. The use of a criteria with a tropospheric residence time of more than 36 hours is unlikely to give a different result to that found using the 36 hour criteria.

In his studies of tropopause folds, Shapiro (1980) found that concentration changes due to mixing in a 12 hour period were of comparable magnitude to the instantaneous concentrations within the fold. This suggests that mixing is of first order importance as a mechanism for chemical exchange, and that a particle within a tropopause fold is likely to have lost its stratospheric properties within 12 hours. However, not all stratosphere to troposphere exchange of particles will be due to tropopause folding events, so it is possible that a 12 hour removal criterion is too severe for use in this work. Additionally, the mixing that changes the properties of a trajectory from stratospheric to tropospheric occurs on scales much smaller than those resolved in the data used here (e.g. Shapiro), so it is necessary to make an estimate of the mixing timescale for gridscale data.

Specific humidity was also considered as an additional removal criterion. Trajectories can be labelled with the specific humidity at their location at each timestep. When a particle crosses the tropopause the background humidity, and therefore the humidity attributed to the particle, will increase, which allows a critical humidity to be used to determine whether a particle is in the stratosphere or the troposphere. The specific humidity of individual particles in the NAFC case study was considered. In every case tested, the particle was always labelled as tropospheric by the temporal/tropopause criteria before there was an increase in attributed humidity (see Figure 6.11 for an example). As no further information was gained from the use of the specific humidity attribute, and there are known issues with ERA humidity fields (e.g. Hagemann et al., 2005), a critical humidity criterion will not be used to determine when a particle has been permanently removed from the stratosphere.

Analysis of the results from the NAFC case study presented here, and the conclusion that a 12 hour removal criterion is likely to be too severe, indicates that particles should be considered to be permanently removed from the stratosphere after they have remained below the tropopause for either 24 or 36 hours. The decision between the two criteria will be confirmed based on the results of case studies of e-folding lifetimes (Section 6.4.2).

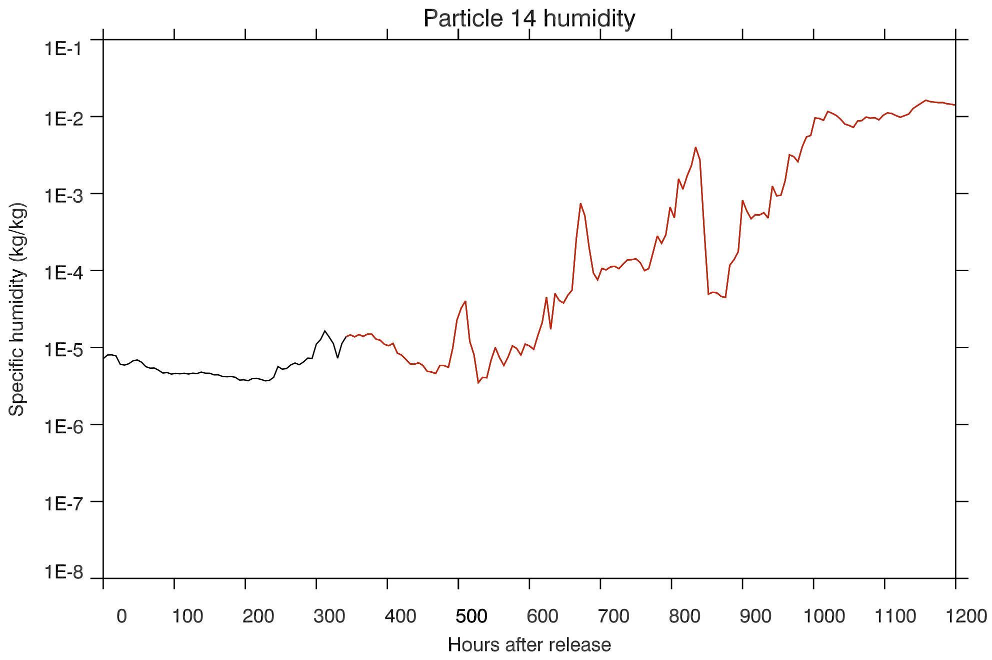


Figure 6.11: *The specific humidity of a randomly selected trajectory, released in the NAFC in January 2004. The line is black when the trajectory is labelled as stratospheric using the tropopause removal criteria, and red when it is labelled as tropospheric. The humidity of the trajectory increases when it crosses the hygropause.*

6.4.2 E-folding lifetime estimates

The e-folding lifetime of stratospheric emissions is the length of time taken for the number of particles initially released into the stratosphere to decrease by a factor of e. e-folding lifetimes have been calculated for three different case studies: release in the NAFC stratosphere in January 2004 and June 2004, and release in the stratosphere over the USA in January 2004.

50-day trajectories have been used to calculate e-folding lifetimes. Following Gettelman (1998), the e-folding lifetime of particles in the stratosphere may be defined:

$$lifetime = \left(-\frac{1}{t}\ln\frac{X_1}{X_0}\right)^{-1} \tag{6.4.1}$$

where X_0 is the number of particles released directly into the stratosphere, X_1 is the number of particles remaining in the stratosphere after 50 days, and t is the time between

X_0 and X_1 (50 days in the case studies).

Ideally, trajectories would have been released on a $1^\circ \times 1^\circ$ horizontal gird to match the resolution of the emissions inventories. Unfortunately, for 50 day long trajectories, this was not possible due to computer memory restrictions. The memory requirement of a trajectory job is determined by the number of attributes assigned to each particle (e.g. pressure, potential temperature, etc.), the length of the trajectories, and the number of particles released. The minimum number of attributes is fixed at three, so the number of particles had to be restricted to make it possible for 50 day trajectories to be calculated.

In each of the three case studies, groups of particles were released on a regular $2^\circ \times 2^\circ$ grid on constant pressure levels every 10 hPa between 300 hPa and 150 hPa ($\sim$9.2 km and $\sim$13.6 km), and at 225 hPa ($\sim$11 km). An e-folding lifetime was then calculated for each release pressure using a 0, 12, 24, and 36 hour removal criteria. Unless otherwise stated, trajectories and tropopause height were calculated using ERA-40 or ERA-Operational data (for post August 2002 calculations).

The calculated e-folding lifetimes for particles released into the NAFC stratosphere in January 2004 are shown in Figure 6.12. It can be seen in Figure 6.12 that e-folding lifetime increases rapidly with decreasing release pressure. The sensitivity of the e-folding lifetimes to the temporal removal criteria also increases with decreasing release pressure. For particles released at 200 hPa ($\sim$11.8 km) there are only small differences between the e-folding lifetimes calculated using the 24 and 36 hour removal criteria, approximately half of the difference between the lifetimes calculated using 12 and 24 hour criteria. As the e-folding lifetimes are relatively insensitive to the choice of a 24 or 36 hour removal criteria, the 24 hour criterion was chosen as this presents less opportunity for spurious mixing. The 24 hour removal criterion will be used for the remainder of the thesis.

e-folding lifetimes for particles released into the NAFC stratosphere in January 2004 range from 12 days for those released at 300 hPa ($\sim$9.2 km) to 64 days for those released at 200 hPa ($\sim$11.8 km) (Figure 6.12). At the median cruise pressure of 225 hPa ($\sim$11 km) the lifetime is 40 days.

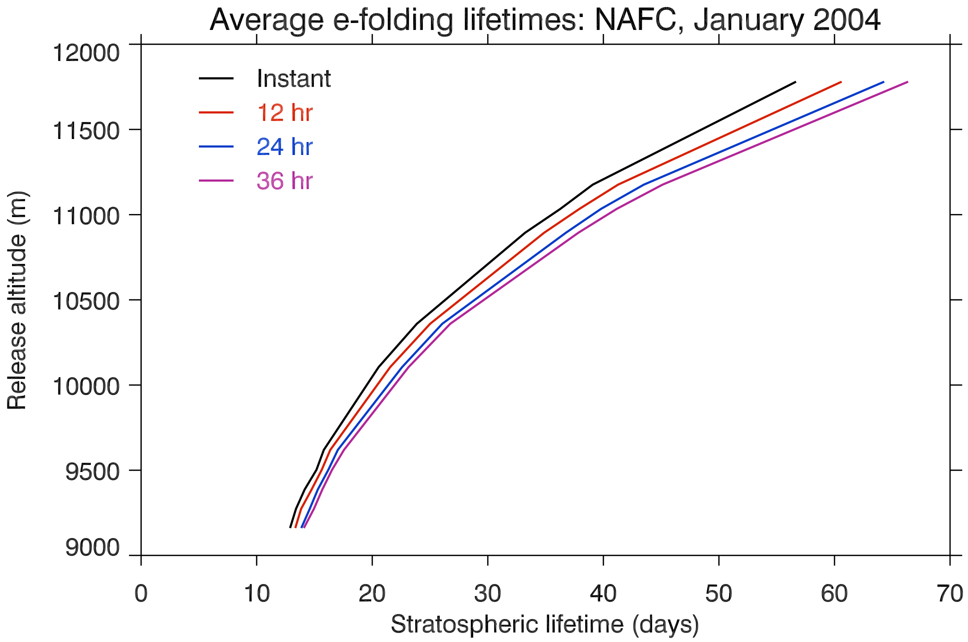


Figure 6.12: *The e-folding lifetime of particles released at different pressures in the NAFC in January 2004, calculated using different removal criteria.*

Sets of particles were released from the same initial grid in the NAFC in June 2004. The calculated e-folding lifetimes are show in Figure 6.13. It can be seen in Figure 6.13 that the stratospheric lifetime of emissions increases rapidly with decreasing release pressure in June as it does in January, but that the lifetimes are shorter for a given release pressure. This demonstrates that the stratospheric lifetime of emissions is dependent on the initial height above the tropopause at which deposition occurs, rather than the release pressure alone.

For particles released in the NAFC in June 2004, the stratospheric e-folding lifetimes range from 11 days for emissions released at 300 hPa (∼9.2 km), to 38 days for those released at 200 hPa (∼11.8 km) (Figure 6.13). At 225 hPa (∼11 km) the emissions have an e-folding lifetime of 25 days.

e-folding lifetimes were also calculated for emissions released over the continental United States in January 2004, between 30°N, 50°N, 130°W, and 70°W. This region is closer to the Equator than the NAFC, so it was expected that the e-folding lifetimes would be shorter due to the higher tropopause . This was found to be the case, with lifetimes rang-

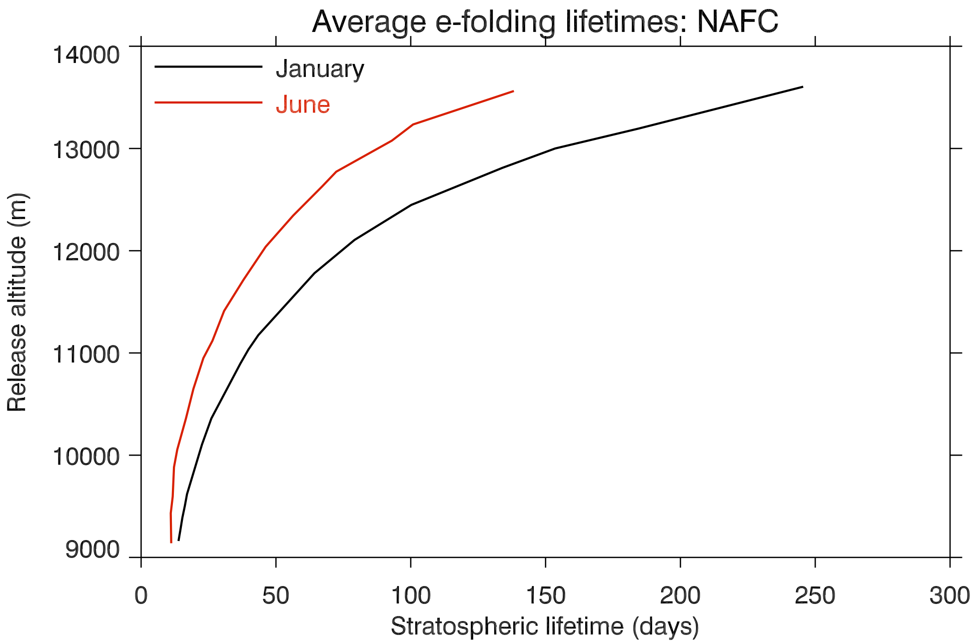


Figure 6.13: *The e-folding lifetime of particles released at different pressures in the NAFC in January and June 2004, assuming a 24 hour removal criterion.*

ing from 11 days at 300 hPa (~9.2 km) to 40 days at 200 hPa (~11.8 km) (not shown). At median cruise altitude the e-folding lifetimes were calculated to be 29 days.

When ERA-Interim data became available the NAFC case studies were repeated using trajectories and tropopause heights calculated from ERA-Interim. It was found that e-folding lifetimes were longer than those calculated using ERA-Operational data (see Figure 6.14), with a lifetime of 54 days for particles released at 225 hPa in January 2004, compared to 40 days using ERA-Operational. This result was expected following Liu (2009), who found that there are problems with over-dispersion in vertical transport calculated using the kinematic version of the UGAMP trajectory model and ERA-40 (Figure 6.15).

Liu (2009) used the UGAMP trajectory model in the form in which it was used in this work, alongside a modified version that produces ‘diabatic trajectories’. In the UGAMP model, vertical velocities are derived from the continuity equation, and horizontal transport is along pressure-based surfaces. Liu modified the model so that particles were transported along potential temperature surfaces. Vertical velocity was calculated di-

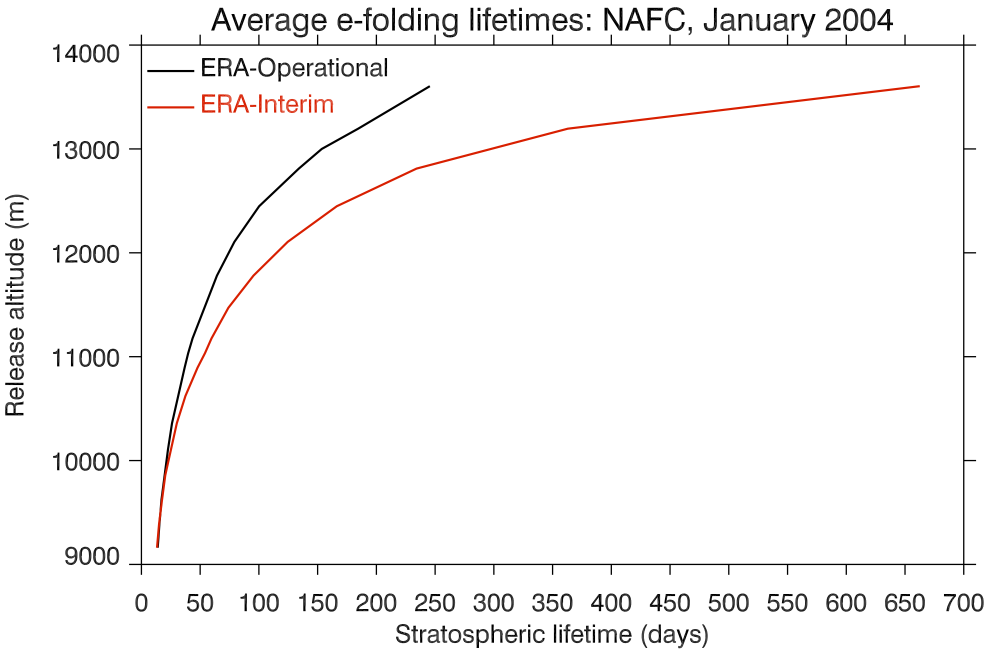


Figure 6.14: *The e-folding lifetime of emissions released in the NAFC in January 2004, calculated using ERA-Operational and ERA-Interim trajectories, with the assumption of a 24 hour removal criterion.*

rectly from the diabatic heating rates provided in ERA. If the heating rate is zero then the trajectories move under adiabatic conditions. Liu notes that as large-scale motion in sizeable parts of the atmosphere is approximately adiabatic, diabatic trajectories better represent the physics compared to kinematic trajectories (Liu, 2009). Liu compared trajectories calculated from ERA-40 and ERA-Interim using the original kinematic method and her diabatic method and found that the problem of over-dispersion was particularly acute when ERA-40 data was used with the kinematic method of calculating trajectories (Figure 6.15). Trajectories calculated from ERA-Interim using the two methods, and from ERA-40 using the diabatic method gave similar results, although there is still a small amount of over-dispersion in the kinematic trajectories calculated using ERA-Interim. In light of Liu's results, and the large differences between ERA-Interim and ERA-Operational e-folding lifetimes, all high-resolution accumulation calculations in Chapter 7 will be made using ERA-Interim.

There is a wide range of e-folding lifetime estimates for stratospheric water vapour emissions presented in previous literature, which are summarised in Table 6.1. The

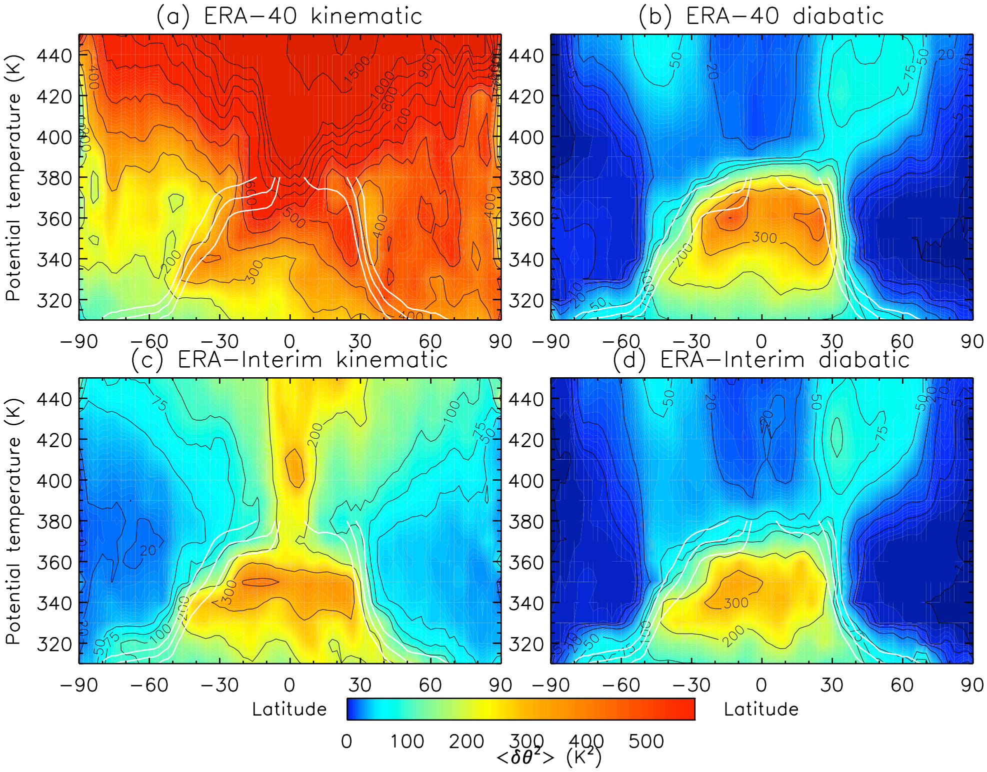


Figure 6.15: *Vertical dispersion from trajectories calculated using ERA-40 (top) and ERA-Interim (bottom), using the kinematic (left) and diabatic (right) version of the UGAMP trajectory model. Potential temperature variance is used as a measure of vertical dispersion:* $\langle \delta\theta(t)^2 \rangle = \langle [\theta(t) - \langle \theta(t) \rangle]^2 \rangle$ *where* θ *is potential temperature,* t *is time, and* $\langle\rangle$ *denotes the ensemble mean.*

lifetimes calculated here from ERA-Operational trajectories are roughly double the estimate from Grewe and Stenke (2008), approximately three times larger than the estimate from Schoeberl et al. (1998), and are comparable with the range of lifetimes presented by Gettelman (1998). The longer ERA-Interim lifetimes are roughly three times larger than the estimate from Grewe and Stenke (2008) and approximately four times larger than the estimate from Schoeberl et al. (1998). At 11 km they are still comparable with the range of lifetimes from Gettelman (1998), but at higher cruise altitudes, they exceed the upper limit of this range (see Table 6.1).

It has been shown in this Chapter that there are several factors that influence the e-folding lifetime of stratospheric emissions. They go some way towards explaining the large range of estimates in previous work, whilst also making it impossible to perform direct comparisons between the results. The e-folding lifetime of emissions is sensitive to their initial height above the tropopause. Therefore, it is sensitive to the definition of the tropopause, and the emissions inventory used, as both can change the relative positions of the tropopause and the emissions. It is also sensitive to the release region in consideration, due to the regional variations in tropopause height. e-folding lifetimes are also sensitive to the meteorological data set used in the calculation. This sensitivity is demonstrated by the comparison between ERA-Operational and ERA-Interim lifetimes (Figure 6.14).

6.4.3 The validity of e-folding lifetimes

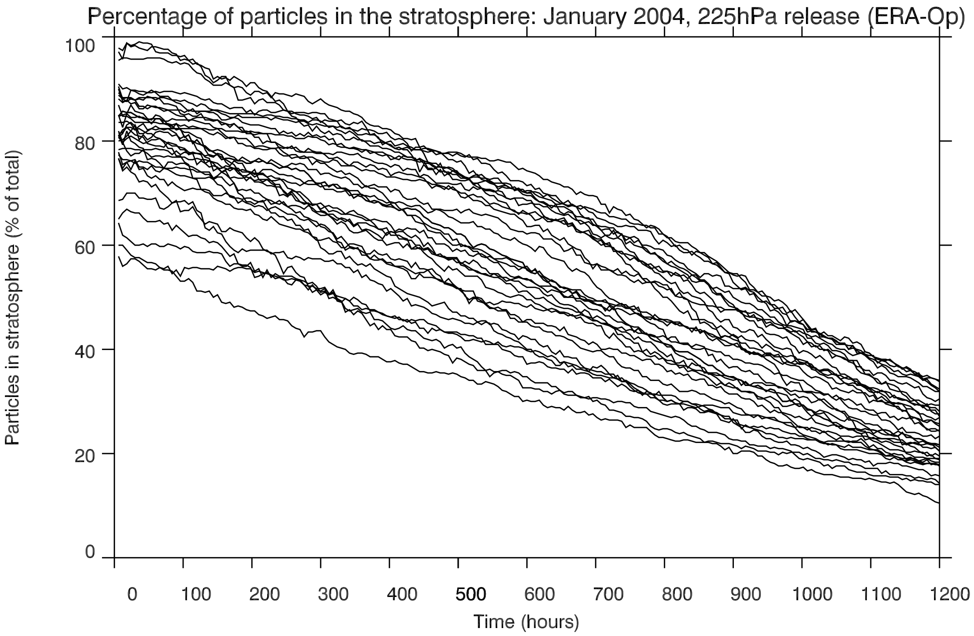
e-folding lifetimes are ubiquitous in the literature (e.g. Gettelman, 1998, Schoeberl et al., 1998, IPCC, 1999, Grewe and Stenke, 2008), and they provide a useful point of comparison with this work. However, analysis of the trajectories used to calculate the e-folding lifetimes in Section 6.4.2 produced results that suggested that the use of e-folding lifetimes is often not appropriate.

When an e-folding lifetime is calculated for stratospheric particles it is assumed that the number of particles in the stratosphere decreases exponentially with time. During analysis of the 50-day trajectories used to calculated e-folding lifetimes, it was found that stratospheric particle numbers did not decay exponentially when the particles were released far above the tropopause. A comparison between the number of particles remaining in the stratosphere for sets of particles released far above the extratropical tropopause at 225 hPa (∼11 km), and close to the tropopause at 300 hPa (∼9.2 km), is shown in Figure 6.16. The number of particles remaining in the stratosphere decreases exponentially with time for particles released at 300 hPa, but for particles released at 225 hPa the decrease is almost linear with time.

When the physical processes behind the vertical transport of emissions in the stratosphere, and their removal across the tropopause, are considered, it is not surprising to find a combination of linear and exponential decay of particle numbers. When particles are released far above the tropopause, their rate of descent is determined by diabatic heating as they have to cross isentropic surfaces. When they are released closer to the tropopause they can be removed from the stratosphere along isentropic surfaces, or via mixing events like tropopause folding. Particles are then removed almost at random, resulting in an exponential decrease in numbers.

When the decay in stratospheric particle numbers for particles released very far above the tropopause in the NAFC, at 150 hPa, was considered it was found that there was very little decrease in the number of particles in the stratosphere for the first few days after their release. This is because all of the particles released were so far above the tropopause that none of them could easily be removed. After a few days the decrease in stratospheric particle numbers eventually became linear. It was expected that this decrease would eventually become exponential with time as the particles' average distance above the tropopause became smaller.

Schoeberl et al. (1998) also found that the decrease in particle numbers is not always exponential. However, they only identified deviations from exponential decay at altitudes above the cruise altitudes of most subsonic aircraft. Schoeberl et al. (1998) examined the decay in stratospheric particle numbers for releases at 19, 16, 14, 13, and 11 km using a Lagrangian diabatic trajectory model. They found that stratospheric particle numbers



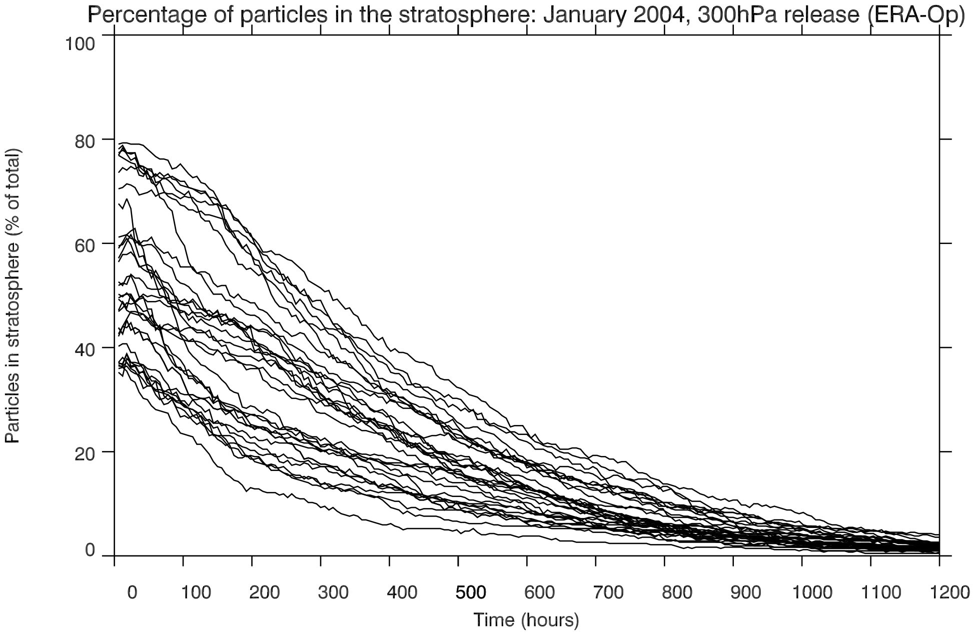


Figure 6.16: *The percentage of particles in the stratosphere with time after their release at 225 hPa (top) and 300 hPa (bottom) in the NAFC in January 2004, calculated from ERA-Operational trajectories.*

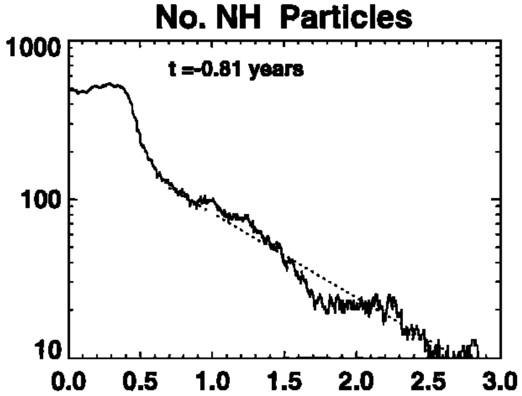


Figure 6.17: *Total number of stratospheric particles versus year for a release at 19 km between 30°N and the pole, from Schoeberl et al. (1998). The dotted line indicates the fit to the line slope, which corresponds to the timescale 't'.*

decreased exponentially for their 11 km release, and that the stratospheric lifetime of these emissions was short as they were rapidly flushed from middleworld. For releases at other altitudes they identified three clear time regimes: a short period with no change, a rapid drop in parcel count (almost like a delta function), and a slow exponential decay of the remaining parcels (see Figure 6.17, note log scale).

The decay in stratospheric particle numbers was examined here using trajectories calculated from ERA-Operational and ERA-Interim, in order to examine the sensitivity of e-folding lifetimes to the reanalysis used to calculate the trajectories. The results from the two data sets were qualitatively similar: a linear decrease in particle numbers for releases far above the tropopause, and exponential decrease in particle numbers for particles released close to the tropopause, and no decrease in numbers, followed by a linear decrease for particles released very far above the tropopause (Figure 6.16). As can be seen in Figure 6.18, there are slight differences in the rates of decrease in particle numbers when the two different reanalyses are used. The rate of decrease is slower in ERA-Interim, which is consistent with the longer lifetimes and reduced vertical dispersion found by Liu (2009) using ERA-Interim and ERA-40 and the UGAMP model.

In light of the results described above, e-folding lifetimes must be used with caution. The assumption that the decay in particle numbers is exponential for releases at all pres-

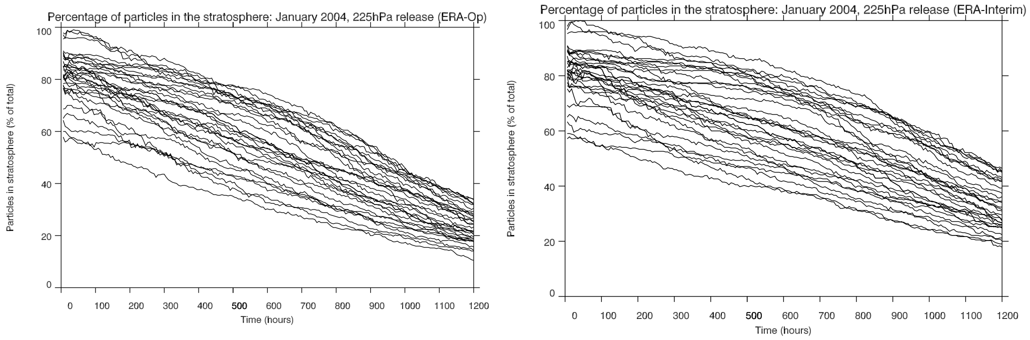


Figure 6.18: *The percentage of particles in the stratosphere with time after their release at 225 hPa in the NAFC in January 2004, calculated using ERA-Operational (left) and ERA-Interim (right) trajectories. The decrease in numbers is faster when ERA-Operational data is used.*

sures, when it may in fact be linear, means that an e-folding lifetime is likely to be an underestimate of the true stratospheric lifetime of emissions. This will not have a negative impact on work reported in later chapters of this thesis as the calculations of the stratospheric accumulations will be calculated directly from trajectories, without any assumptions about lifetimes.

6.5 Summary

The direct deposition of emissions into the stratosphere has been shown to be sensitive to tropopause definition, emissions inventory, and inventory vertical resolution. There is inter-annual variability in the mass of NAFC emissions deposited into the stratosphere directly, but there is not a strong correlation with the NAO index. Stratospheric deposition was not found to be particularly sensitive to the meteorological data set used to calculate tropopause position.

For the February 2004 case study, 61% of monthly total AEDT and QUANTIFY emissions, and 52% of AERO2k emissions enter the stratosphere directly. When 6 hourly AEDT emissions data is used, there can be in excess of 90% of column total NAFC emissions deposited directly into the stratosphere in individual 6 hour periods.

Using the 2006 tropopause and AEDT data, a seasonal cycle in direct deposition was identified with a maximum of 64% of column total NAFC emissions entering the stratosphere directly in March. Direct deposition has a minimum of 38% in June. For global column total emissions this is reduced to a maximum of 31% in March, and a minimum of 16% in August. Monthly mean percentages were found to be anti-correlated with monthly mean NAFC tropopause height, with r^2=0.88 for NAFC emissions and the NAFC tropopause, r^2=0.94 for global emissions and the NAFC tropopause, and r^2=0.56 for global emissions and the global tropopause.

The UGAMP trajectory model has been used to find the motion of emissions after their release into the stratosphere. It is a fully Lagrangian trajectory model, driven using ECMWF spectral wind records. Only stratospheric emissions were considered in lifetime and accumulation calculations, as only stratospheric emissions have the potential to be significant compared to the natural background humidity. Hence, it was necessary to determine when trajectories left the stratosphere. 50 day long trajectories were used to identify a suitable temporal removal criterion. Particles were labelled as having lost their stratospheric properties after spending 24 consecutive hours below the tropopause. This criterion takes account of the typical length of time that particles spend on temporary excursions into the troposphere, and also represents a good approximation to typical mixing timescales.

e-folding lifetimes were calculated for emissions released into the NAFC stratosphere in January and June 2004, and into the USA stratosphere in January 2004, using 50 day trajectories calculated from ERA-Operational data. For particles released at the median cruise pressure of 225 hPa ($\sim$11 km) e-folding lifetimes are 40 days (January NAFC), 25 days (June NAFC), and 29 days (January USA). These values demonstrate the sensitivity of stratospheric e-folding lifetimes to the distance of the release pressures above the tropopause.

e-folding lifetimes were also calculated using ERA-Interim for emissions released into the NAFC stratosphere in January 2004. Particles released at 225 hPa ($\sim$ 11 km) had an e-folding lifetime of 54 days. For releases at all pressure, lifetimes calculated from ERA-Interim trajectories were longer than those calculated using ERA-40 and ERA-Operational data. This is consistent with the excessive vertical dispersion found by Liu (2009) in trajectories calculated using ERA-40 and the UGAMP trajectory model.

Analysis of the trajectories calculated from all ERA data sets has demonstrated that the assumption of e-folding lifetimes is not appropriate for emissions released far above the tropopause. Exponential decay in stratospheric particle numbers can be seen in releases just above the tropopause, but the decay becomes steadily more linear with increased initial altitude above the tropopause.

Chapter 7

High resolution accumulation calculations

The accumulated mass of water vapour emissions in the stratosphere can be calculated directly from trajectories, without the need for assumptions about lifetimes, if each trajectory is assigned a mass of water vapour from an emissions inventory. If trajectories are released on a regular grid, covering at least one hemisphere, then any mass can be passed on to a new trajectory whenever the trajectory to which it is assigned ends (see Figure 7.1). This method means that mass is never lost when a trajectory ends. Hence, if sets of trajectories are calculated for long enough periods of time the total mass of water vapour in the stratosphere will reach equilibrium, from which a perturbation to the natural background can be calculated.

The memory requirement of a trajectory job is determined by the length of the trajectories, the number of trajectories and the number of attributes that are assigned to them. Ideally, trajectories should be released on the same grid as the AEDT inventory, so that the benefits of using a high resolution inventory are not lost. It was assumed, and later tested, that there would be little transport of emissions across the equator, so trajectories were only calculated for the Northern Hemisphere to allow the release of trajectories on a higher resolution grid. For trajectories to be calculated from the AEDT release grid

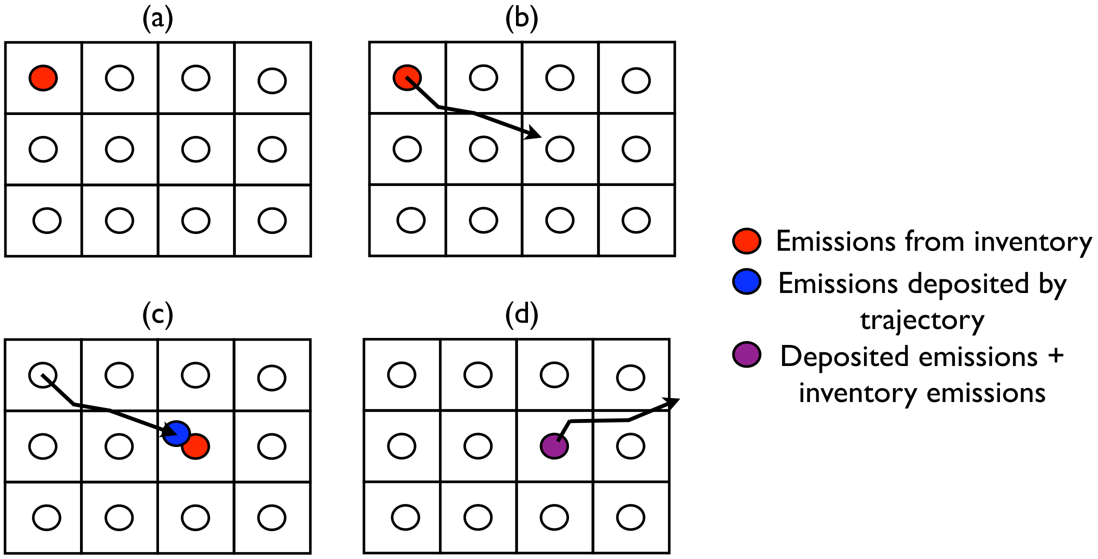


Figure 7.1: *Operation of the accumulation code.*
(a): Emissions from an emissions inventory are input into the model
(b): Emissions move along trajectories calculated from ERA winds
(c): When a trajectory ends, the emissions assigned to it are deposited into the local grid cell
(d): A new trajectory is then initialised in that grid cell, and assigned the deposited mass in addition to mass from the inventory. This mass is then advected along the new trajectory
Emissions are input, trajectories are initialised, and accumulation is calculated every 6 hours

for the whole Northern Hemisphere, trajectories could only have a maximum length of 2 days before the memory requirement of the calculations became too large.

2-day long trajectories were released on a 1°×1°×152 m×6 hr grid, covering the whole Northern Hemisphere. Trajectories were released on pressure levels corresponding to the AEDT pressure-height levels, which are based on the assumption of the ICAO standard atmosphere. Releases were made approximately every 4 hPa (152 m) between 466 hPa and 151 hPa. 151 hPa corresponds to a pressure-altitude of 13.562 km, just above the AEDT cap at 13.5 km, above which emissions are assumed to be erroneous (Chapter 2). 466 hPa corresponds to the altitude of the climatological minimum tropopause in December to March (6.1 km), when most of the trajectory analysis will be performed and the minimum climatological tropopause occurs. As the climatological minimum tropopause occurs in March, this lower bound is also applicable to other months.

Once calculated, the 2-day trajectories were passed into an accumulation routine, which assigned a mass of water vapour from the AEDT emissions inventory to each trajectory, passed mass from finishing trajectories to new trajectories, removed any tropospheric mass using the blended tropopause and the 24 hour removal criterion, and found the accumulated mass at each point (Figure 7.1). The operation, development and testing of the accumulation code is described briefly in the following section.

7.1 The accumulation code

Groups of trajectories were released every 6 hours on a 1°×1° grid on pressure levels corresponding to each of the AEDT pressure-height levels between 466 hPa and 151 hPa (a spacing of $\sim$4 hPa). The choice of the lowest altitude (highest pressure) to be used for trajectory releases was vital to the accurate calculation of the accumulated emissions: too high in altitude and not all the trajectories will reach the tropopause, too low in altitude and computationally expensive calculations will be performed unnecessarily.

Trajectories are released between 466 hPa and 151 hPa at all latitudes, regardless of whether their release occurs in the stratosphere or the troposphere. The blended tropopause removal criterion is also applied to all trajectories. This means that although mass can be input directly into the troposphere, it will be removed within 24 hours. However, this can result in a small amount of tropospheric accumulation along the main flight corridors, where up to 24 hours of input emissions may be present. It is necessary to allow this to occur, as these trajectories may also be receiving emissions that have recently been transported from the stratosphere.

The majority of water vapour emissions occur at pressure-altitudes above 9 km (pressures less than 308 hPa). However, this is far above the climatological minimum altitude of the tropopause, particularly in the Arctic region. If 9 km was chosen to be the lower bound for trajectory calculations there would be a significant mass of water vapour lost from the calculations without crossing the tropopause as trajectories would end in the stratosphere, but above 9 km. The sensitivity of the accumulated mass to the lower bound altitude for trajectory releases, and the mass loss out of the top of the grid and across the equator, was tested by assigning a unit mass to each trajectory.

The accumulation routine was run for January 2004 trajectories with a lower bound altitude of 6.1 km (466 hPa), which is the minimum altitude of the monthly mean tropopause in DJFM, to ensure that it was mass conserving. It was found that 5% of input mass was lost without crossing the tropopause. The tropopause at individual grid points can be several kilometres below the minimum height of the monthly-mean tropopause. However, the required computer time to release trajectories at enough pressure levels to capture all very low tropopause occurrences would be prohibitive. In all of the cases examined, mass lost in this way was below the level of the tropopause at nearby grid points. Additionally, the trajectory to which it was assigned was travelling downwards, suggesting that it would have quickly been removed across the tropopause had it not been lost out of the bottom of the accumulation grid.

In unit mass tests of the accumulation routine it was found that 9.2% of input mass was lost out of the top of the model. This is reduced to 3.8% when actual emissions are used, as only a relatively small mass of emissions enter the atmosphere at the top end of the

AEDT altitude range. It was found that this value was not notably reduced by including an additional 10 release levels above the AEDT cap.

An additional 2.5% of input mass in unit mass test was lost when it was advected across the equator. In some cases using actual emissions, no emissions cross the equator: there are relatively few emissions released in the tropics, and when emissions are released here they are deposited in the troposphere.

When trajectories remained within the boundaries of the release grid, mass was conserved. The small losses across the boundaries of the accumulation grid combine to give up to a 10% loss of input mass at cruise altitudes. However, as the mass lost downwards was generally close to being removed across the tropopause it is unlikely to cause too much of an underestimate in the calculated perturbation. Thus, the underestimate in accumulation due to emissions being lost across the boundaries of the accumulation grid is likely to be closer to 5%. A further 5% underestimate in global accumulation arises as the result of neglecting SH emissions.

7.2 Perturbation concentrations

Two-day long trajectories were released from a $1^\circ \times 1^\circ \times 152$ m×6 hr grid for the whole NH between pressures of 466 hPa (6.1 km) and 151 hPa (13.6 km). These trajectories were then combined with 6 hourly emissions from AEDT 2006 to calculate the accumulation of emissions in the stratosphere. Several case studies are considered, and compared to the perturbations found in DJFM 2003/2004.

The effect of flight routing on emissions accumulation is investigated through the comparison of perturbations calculated using the AEDT, QUANTIFY and AERO2k inventories. The sensitivity of the perturbation to the reanalysis data used in the trajectory calculations is investigated through the comparison of perturbations calculated from trajectories calculated using ERA-Operational data, and ERA-Interim data at full resolution and ERA-40 resolution.

The impact of different meteorological conditions on the size and structure of the perturbation is investigated through the comparison of perturbations for DJFM 2003/2004, DJFM 1994/1995, DJFM 1995/1996, and DJFM 1997/1998, representing different phases of the NAO and ENSO. A comparison between summer and winter perturbations is made using calculations for June to September (JJAS) 2006. The annual cycle in accumulated emissions is also discussed for the year 2006.

Typically, it takes 2 to 3 weeks for the accumulated mass in the stratosphere to reach equilibrium. Hence, true monthly-mean perturbations are not presented for the first month of each period considered. However, this month has an influence on the perturbation in the following month.

7.2.1 DJFM 2003/2004 perturbations

Zonal-mean monthly-mean perturbations for January to March 2004 are shown in Figure 7.2. The maximum zonal-mean monthly-mean perturbations are 60, 64, and 70 ppbv respectively, occurring at 46, 51, and 53°N. In all months, the maximum perturbation is found at an altitude of just less than 11 km, which corresponds to the altitude of maximum emissions. It can be seen in Figure 7.2 that although only stratospheric emissions are considered, the zonal-mean perturbation extends below the zonal-mean tropopause. This can be attributed to the use of the 24 hour removal criteria, which means that emissions can be found below the tropopause for up to 24 hours after they cross it, and the 6 hourly addition of emissions, which means that up to 24 hours of emissions that are input into the troposphere directly will remain there. There is also an apparent extension of the perturbation below the tropopause in the zonal-mean as a result of averaging.

There is some variation in the structure of the perturbation across the three months. This is particularly apparent in February, when the perturbation follows a dip in the zonal-mean tropopause near 75°N (Figure 7.2). The magnitude of the perturbation increases from January to March, although the perturbation is smaller near to the pole in February. These changes are due primarily to changes in the meteorology. There is little differ-

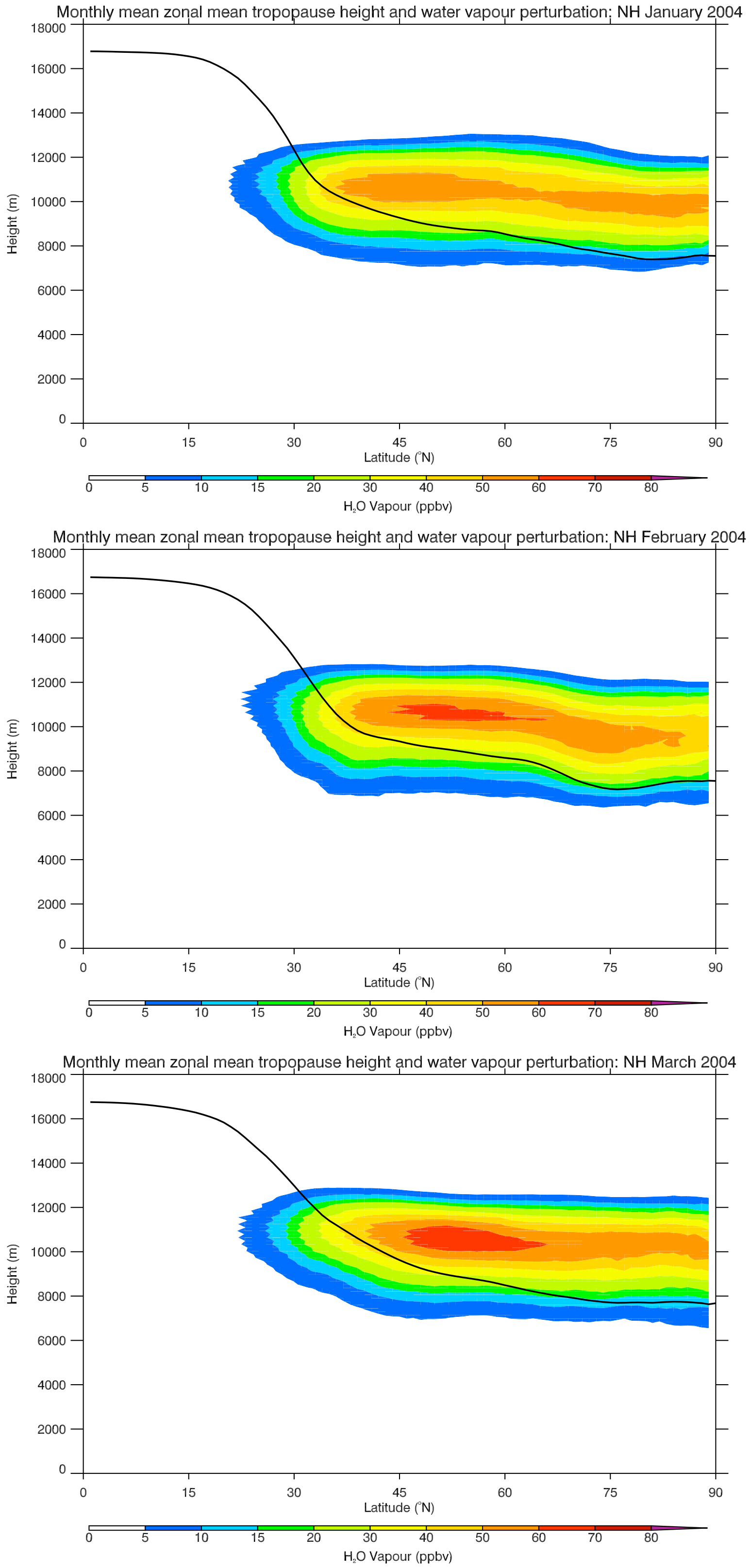


Figure 7.2: *Zonal-mean monthly-mean perturbations (ppbv) for January (top), February (middle), and March (bottom) 2004. The position of the zonal-mean monthly-mean blended tropopause is indicated by the position of the black line. Input emissions are from AEDT 2006.*

ence in the structure of the emissions in these months, and the minimum in global-total emissions that occurs in February is not reflected in the perturbation. The zonal-mean monthly-mean perturbation is actually smaller in January compared to other months.

In January there is descent[1] between 30 and 45°N, in the region of large input emissions (Figure 7.3). This will result in many of these emissions quickly being removed from the stratosphere. In contrast to this, there is only very weak descent, or strong ascent, in this region in February and March, allowing more of these input emissions to be transported to higher altitudes and latitudes.

There is a region of ascent between 45 and 75°N in January, with downwards velocities over the Arctic (Figure 7.3). In February, this region of ascent extends from 35 to 75°N, and has a larger magnitude, resulting in the larger perturbation. However, the descent in the Arctic is comparable, and it appears that emissions are more effectively removed from the February Arctic stratosphere compared to January.

The largest perturbation occurs in March. This is the month with the strongest ascent in the extratropics, with upwards velocities extending from 35 to 70°N in the zonal-mean (Figure 7.3). There is also ascent in the Arctic in this month, resulting in the large perturbations here.

The zonal-mean and column total monthly-mean accumulated water vapour in January 2004 is shown in Figure 7.4. The monthly-mean accumulated mass for each month is given in Table 7.1. It can be seen in Figure 7.4 that the distribution of accumulated mass is very similar to the distribution of the perturbation for January 2004, with a maximum zonal-mean of 19×10^6 kg near 40°N. Column-total accumulated mass has a maximum in the North Pacific jet entrance region over Japan. This is also the case for February and March, although the maxima are slightly larger in those months compared to January.

---

[1] In this Chapter, 'descent' refers to positive zonal-mean monthly-mean vertical velocities [Pa s^{-1}]. Similarly, 'ascent' refers to a negative zonal-mean monthly-mean vertical velocities.

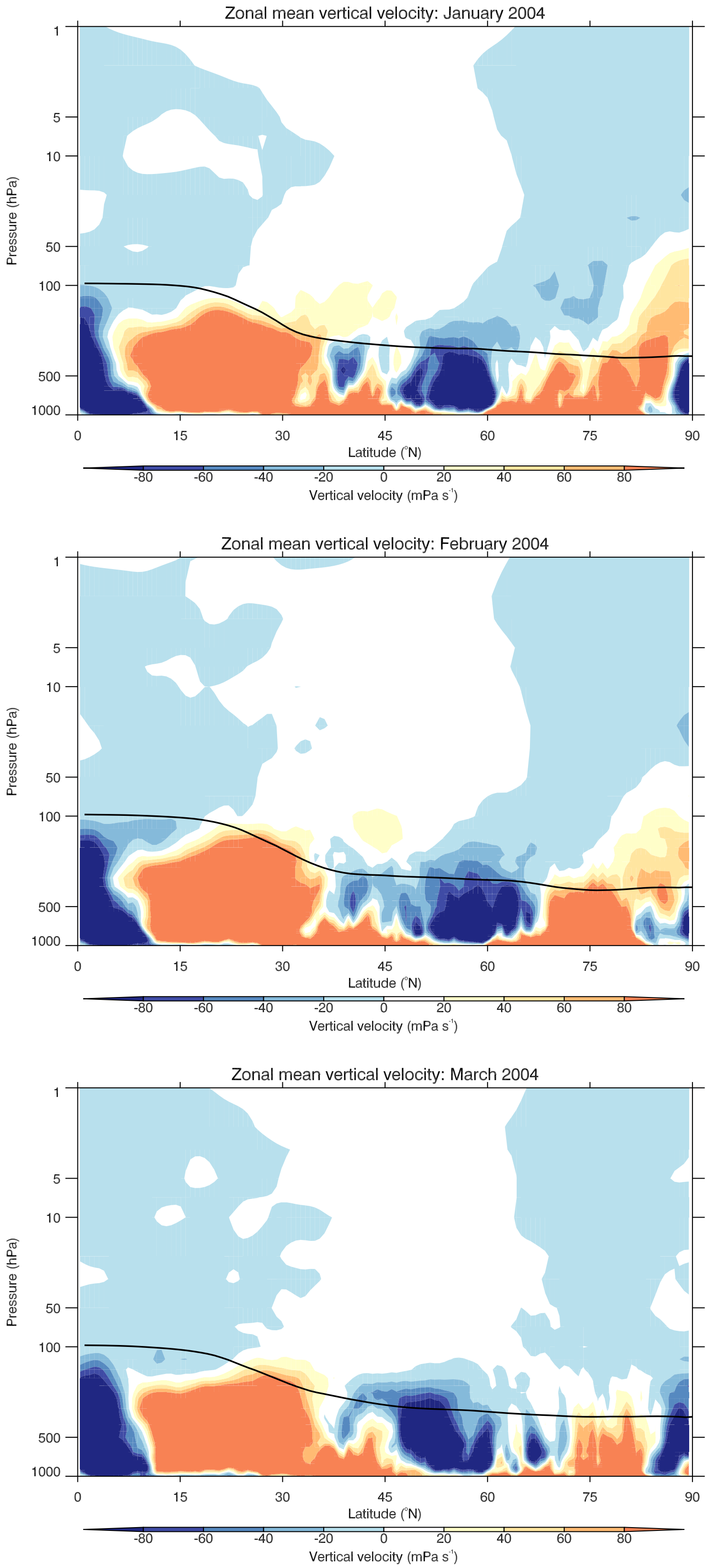


Figure 7.3: *Zonal-mean monthly-mean vertical velocities (Pa* s^{-1} *for January (top), February (middle), and March (bottom) 2004. The position of the zonal-mean monthly-mean blended tropopause is indicated by the position of the black line.*

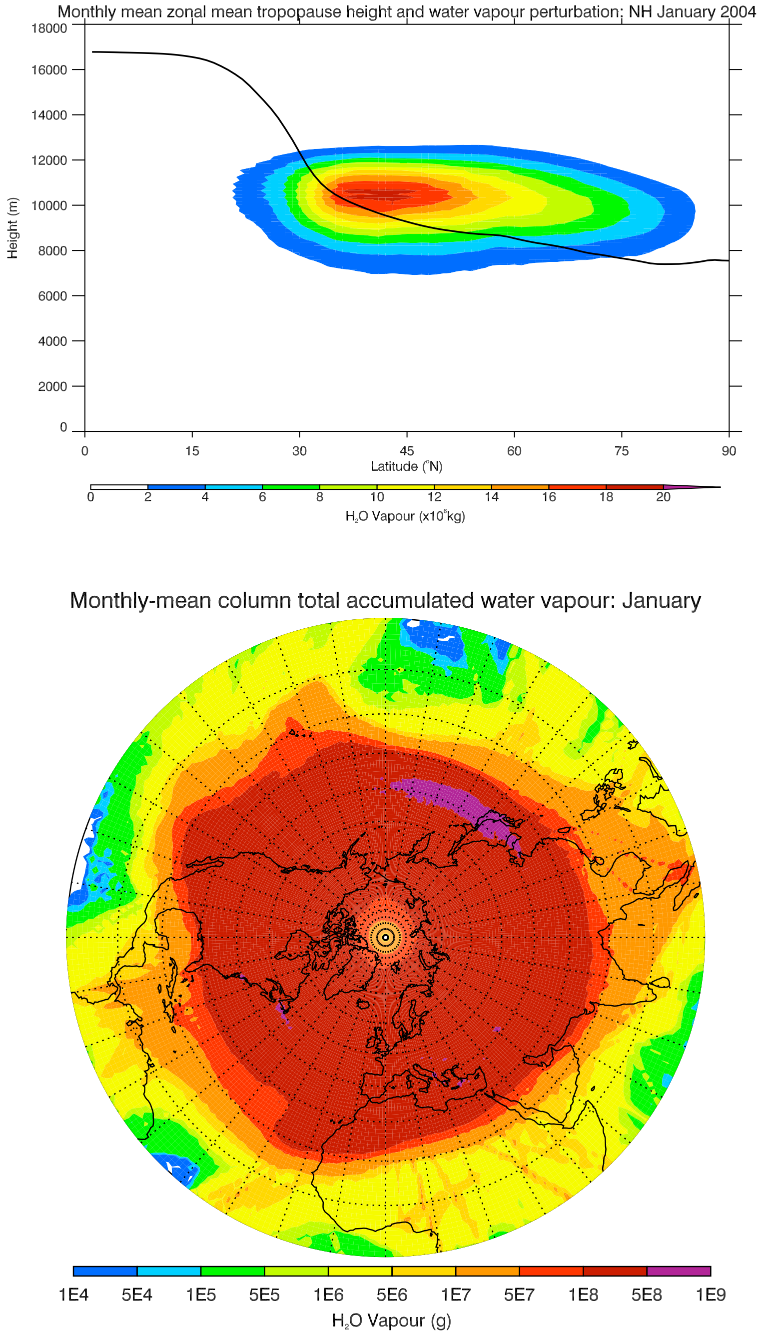


Figure 7.4: *Top: Zonal-mean monthly-mean accumulated mass of water vapour for January 2004. The position of the zonal-mean monthly-mean blended tropopause is indicated by the position of the black line.*

Bottom: Column-total monthly-mean accumulated mass of water vapour for January 2004.

Input emissions are from AEDT 2006 in both cases.

	Jan	Feb	Mar
DJFM 2003/2004	60, *46*, **5.49**	64, *51*, **5.72**	70, *53*, **5.37**
DJFM 1994/1995	60, *47*, **5.67**	63, *60*, **5.98**	70, *51*, **6.05**
DJFM 1995/1996	66, *51*, **5.53**	65, *88*, **6.04**	75, *48*, **5.96**
DJFM 1997/1998	66, *47*, **5.56**	66, *89*, **5.67**	71, *89*, **5.74**

Table 7.1: Maximum zonal-mean monthly-mean perturbation (ppbv), *the latitude at which it occurs (°N)*, and **the monthly-mean mass accumulated in the stratosphere (Tg)**, *for the four winters considered.*

7.2.2 Sensitivity to emissions inventory

In addition to the perturbations calculated using the AEDT inventory, calculations were performed for DJFM 2003/2004 using the QUANTIFY and AERO2k inventories. As discussed in Chapter 2, the three inventories have very different emission distributions, particularly in the vertical. In Chapter 6 it was shown that these differences have an influence on the mass of the emissions that enter the stratosphere directly. These differences have also been found to have a large impact on the mass of accumulated emissions in the stratosphere.

The difference between the perturbations calculated using the AEDT inventory (Figure 7.2) and the QUANTIFY inventory is shown in Figure 7.5. It can be seen in Figure 7.5 that the perturbation from QUANTIFY emissions is smaller than that from AEDT emissions. The pattern of the difference follows the pattern of the perturbation. There is a maximum difference of up to one third of the maximum zonal-mean monthly-mean AEDT perturbation in this region.

Emissions are more evenly distributed with height in the AERO2k inventory, compared to in AEDT and QUANTIFY, and a much smaller mass of emissions is emitted at cruise altitude. This can clearly been seen in the difference between the perturbations calculated from AEDT and AERO2k emissions (Figure 7.6). The AERO2k perturbation is uniformly smaller than the AEDT perturbation, with a maximum difference of up to half of the maximum zonal-mean monthly-mean AEDT perturbation.

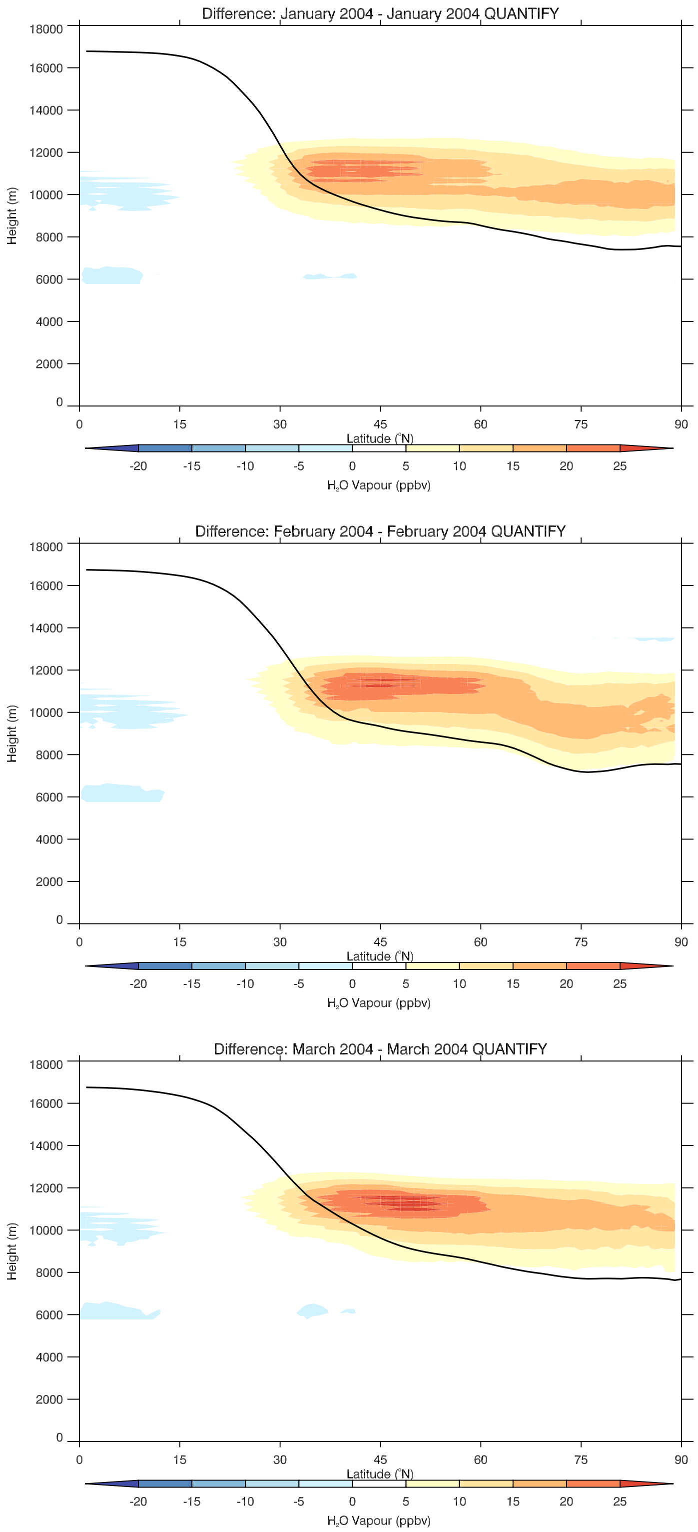


Figure 7.5: *Difference between zonal-mean monthly-mean perturbations (ppbv) calculated using AEDT and QUANTIFY emissions for January (top), February (middle), and March (bottom) 2004 trajectories. The position of the zonal-mean monthly-mean tropopause is indicated by the position of the black line.*

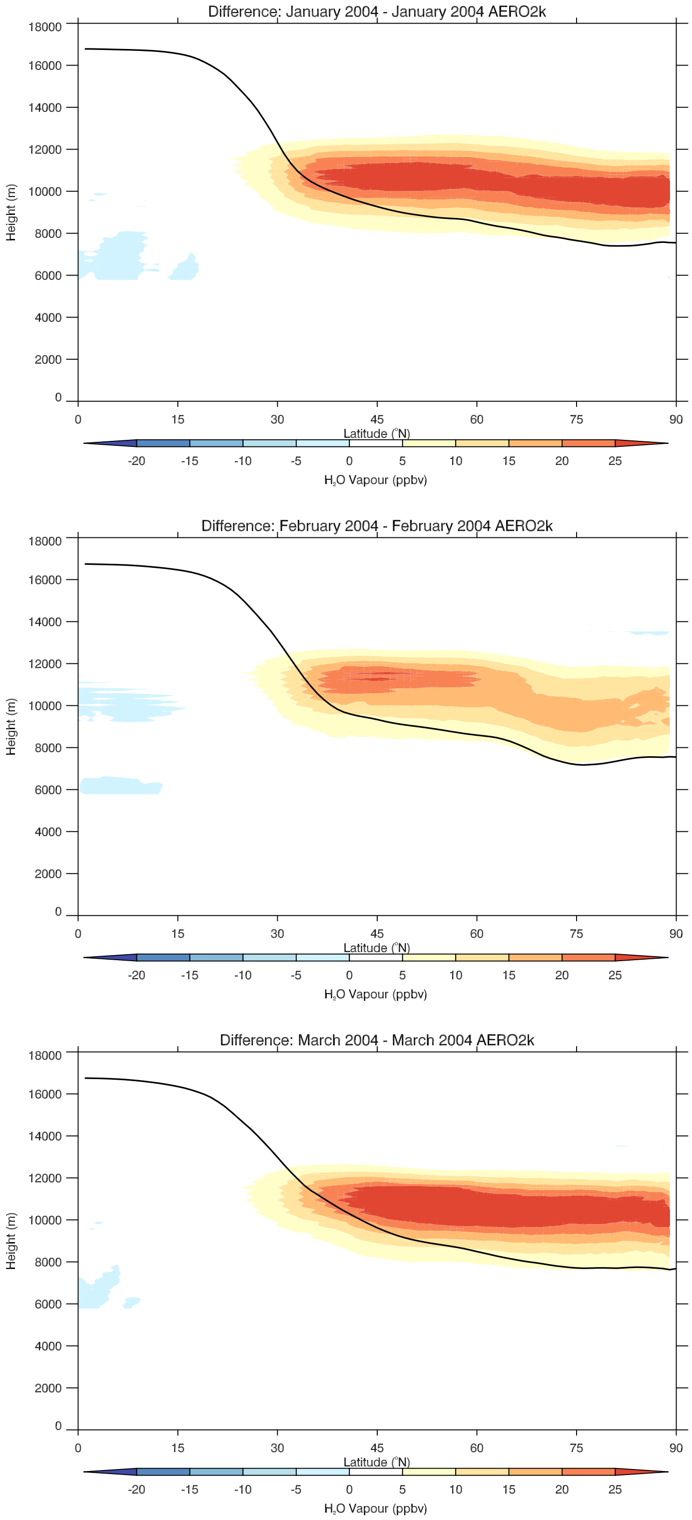


Figure 7.6: *Difference between zonal-mean monthly-mean perturbations (ppbv) calculated using AEDT and AERO2k emissions for January (top), February (middle), and March (bottom) 2004 trajectories. The position of the zonal-mean monthly-mean tropopause is indicated by the position of the black line.*

7.2.3 Sensitivity to reanalysis data

Two-day trajectories were calculated using ERA-Interim data, ERA-Operational data, and ERA-Interim data at ERA-40 resolution, for DJFM 2003/2004. As only the NH is being considered, it is expected that there will be reasonable agreement between the three results due to the high density of observations used in the reanalyses. However, Liu (2009) identified that excessive vertical dispersion occurs when the UGAMP trajectory model is used with ERA-40 data. It was shown in Chapter 6 that this results in shorter e-folding lifetimes compared to ERA-Interim calculations. Therefore, it is expected that the perturbations calculated using ERA-Operational trajectories will be smaller than those calculated with ERA-Interim trajectories.

Comparison of the zonal-mean monthly-mean perturbations from the two sets of ERA-Interim calculations reveals no significant differences. Differences in the zonal-mean perturbations never exceed 5 ppbv, which is an order of magnitude smaller than the maximum zonal-mean perturbation. Additionally, there is no coherent pattern in the difference field.

There are small differences between the perturbations calculated using ERA-Interim and ERA-Operational data. The use of ERA-Operational data results in a slightly larger perturbation at high altitudes in the high latitudes, and a slightly smaller perturbation at lower altitudes, compared to perturbations calculated using ERA-Interim trajectories. However, these differences are an order of magnitude smaller than the perturbations themselves in February 2004. In the other months they are smaller still.

7.2.4 Sensitivity to inter-annual variability in meteorology

Perturbations have been calculated for DJFM 1994/1995, DJFM 1995/1996, and DJFM 1997/1998, and compared to those calculated for DJFM 2003/2004. DJFM 2003/2004 is a winter with a very slightly negative NAO index (-0.07), while DJFM 1994/1995 has a strongly positive NAO index (3.96), and DJFM 1995/1996 has a strongly negative

NAO index (-3.78). A large El Niño occurred in DJFM 1997/1998. The NAO index had a value of 0.72 in this year.

Although the stratospheric circulation in the NH is influenced by several factors, whose interaction may not always be simple, it is expected that the influence of the NAO on the perturbations might be identifiable. 20% of monthly total global AEDT emissions occur at altitudes above 6.1 km over the North Atlantic and Europe, and it was shown in Chapter 6 that global stratospheric deposition is strongly correlated with NAFC deposition. Although there was no correlation between NAFC deposition and the NAO index, there was a strong correlation between NAFC deposition and average NAFC tropopause height. It is also possible that the changes in wind field in the North Atlantic region, and over the pole, with the NAO might have an influence on the global perturbation through their influence on the trajectories immediately following deposition.

7.2.4.1 DJFM 1994/1995: Strongly positive NAO index

The differences between the perturbations calculated using DJFM 1994/1995 meteorology and AEDT 2006 emissions, and those calculated using DJFM 2003/2004 meteorology, are shown in Figure 7.7. The differences between the perturbations resulting from DJFM 2003/2004 and DJFM 1994/1995 meteorology are largest at high latitudes, particularly in January and February. In all months, the difference at high latitudes has a dipole structure. Perturbations from 2004 meteorology are smaller at high altitudes and larger close to the tropopause compared to 1995 meteorology. There is also a smaller difference near 35°N, again in each month shown, where the 1995 perturbation is larger than that from 2004 meteorology.

The zonal-mean monthly-mean tropopause in January, February, and March 1995 is higher at high latitudes than the 2004 tropopause (dashed line compared to solid line in Figure 7.7). This means that emissions start to be removed into the troposphere at higher altitudes compared to 2004, resulting in a smaller perturbation near the tropopause. Similarly, tropopause position is likely to be the cause of the negative difference near 35°N.

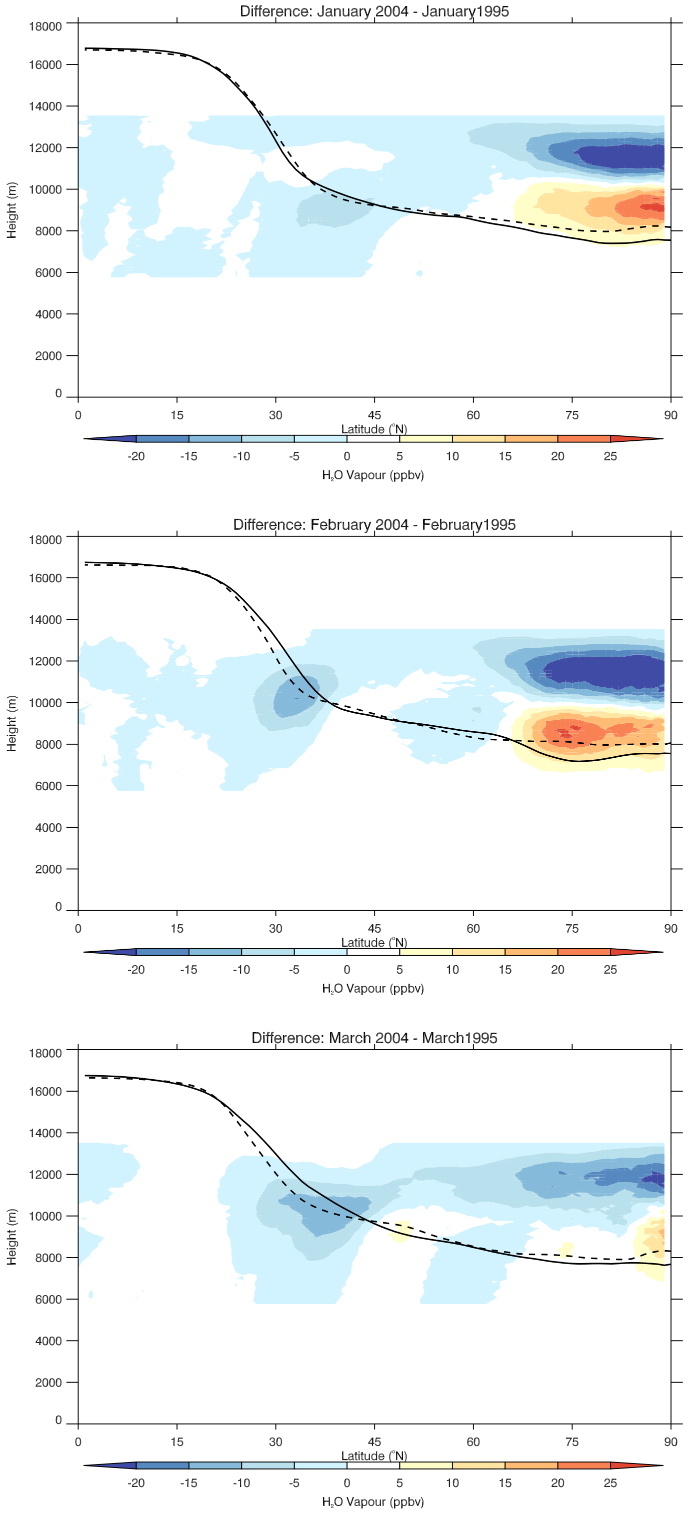


Figure 7.7: *Difference in January to March zonal-mean monthly-mean perturbations calculated DJFM 2003/2004 and DJFM 1994/1995 meteorology. The solid line indicates the position of the 2004 monthly-mean tropopause, the dashed line indicates the position of the 1995 tropopause. Input emissions are from AEDT 2006.*

The zonal-mean tropopause here is lower in 1995 compared to 2004, so the perturbation extends to lower altitudes.

The large negative difference at high latitudes in Figure 7.7 is primarily due to differences in the atmospheric circulation, rather than differences in the position of the tropopause, between the two years. In January and February 1995 there are stronger zonal-mean zonal winds at high latitudes, compared to 2004, and descent near the pole. These differences probably caused greater transport of emissions to high latitudes, and, for both transported emissions and locally input emissions, longer residence times and accumulation in the Arctic stratosphere, in January and February 1995 compared to 2004.

In March, the differences between the perturbations at high latitudes from 1995 and 2004 meteorology are smaller compared to those in January and February, although they have a similar dipole structure at high latitudes. The positive difference is smaller in March as the 1995 and 2004 zonal mean tropopauses are closer than in the other months. There is strong ascent between 60°N and 80°N in 1995, compared to weak descent in 2004, which will act to loft the emissions higher in 1995, giving them longer stratospheric lifetimes. However, there is also strong descent over the pole in 1995, compared to weak ascent in 2004. This means that more emissions are likely to be transported to the polar stratosphere, but that they will descend faster there than in 2004, resulting in a smaller difference than in January and February. The effect of the strong ascent in the extratropics in 1995 is reflected in the larger negative difference in the extratropics in March, compared to January and February.

Peak zonal-mean perturbations of 60, 63, and 70 ppbv occur in January to March respectively. The accumulated mass is also largest in March, with monthly-mean accumulations of 5.67, 5.98, and 6.05 Tg occurring in January to March (Table 7.1).

7.2.4.2 DJFM 1995/1996: Strongly negative NAO index

The differences between the perturbations calculated using DJFM 1995/1996 meteorology and AEDT 2006 emissions, and those calculated using DJFM 2003/2004 meteorology, are shown in Figure 7.8. As in the DJFM 1994/1995 example, the largest differences between the perturbations from 1996 and 2004 meteorology are at high latitudes. Again, the differences have a dipole structure here, with the 1996 perturbation being larger than the 2004 perturbation at high altitudes, and smaller near the tropopause.

The differences in the January and February perturbations resulting from the use of 1995 and 1996 meteorology compared to those from the DJFM 2003/2004 base case have a similar structure. The two years also have a somewhat similar zonal-mean background circulation. The causes of the structure of the difference between the perturbations from the two years relative to the DJFM 2003/2004 case are similar, but there are variations in the 1995 and 1996 meteorology that result in differences in the magnitudes. Therefore, the January and February 1996 perturbations are discussed relative to those from January and February 1995 meteorology. The larger magnitudes of the negative polar anomaly seen in Figure 7.8 compared to Figure 7.7 appear to be due to larger stratospheric deposition in January, and larger regions of ascent in February. In January 1996 there is a lower tropopause over most of the North Atlantic, with comparable tropopause heights over North America, which will result in more direct deposition into the stratosphere. There is also ascent near the pole in 1996, compared to descent in 1995, leading to longer removal times and a larger perturbation.

In February the region of ascent has greater latitudinal extent in 1996 compared to 1995, and comparable magnitudes. In February 1995, there is ascent from 50°N to 70°N. However, in February 1996, the region of ascent extends from near 50°N to the pole, with the exception of a region of weak descent just above the tropopause from 75–85°N.

In March there are negative differences between the perturbations from 2004 and 1996 meteorology just above the tropopause in the subtropics and at the pole, and a positive difference in the extratropics (Figure 7.8). This reflects the structure of the zonal-mean

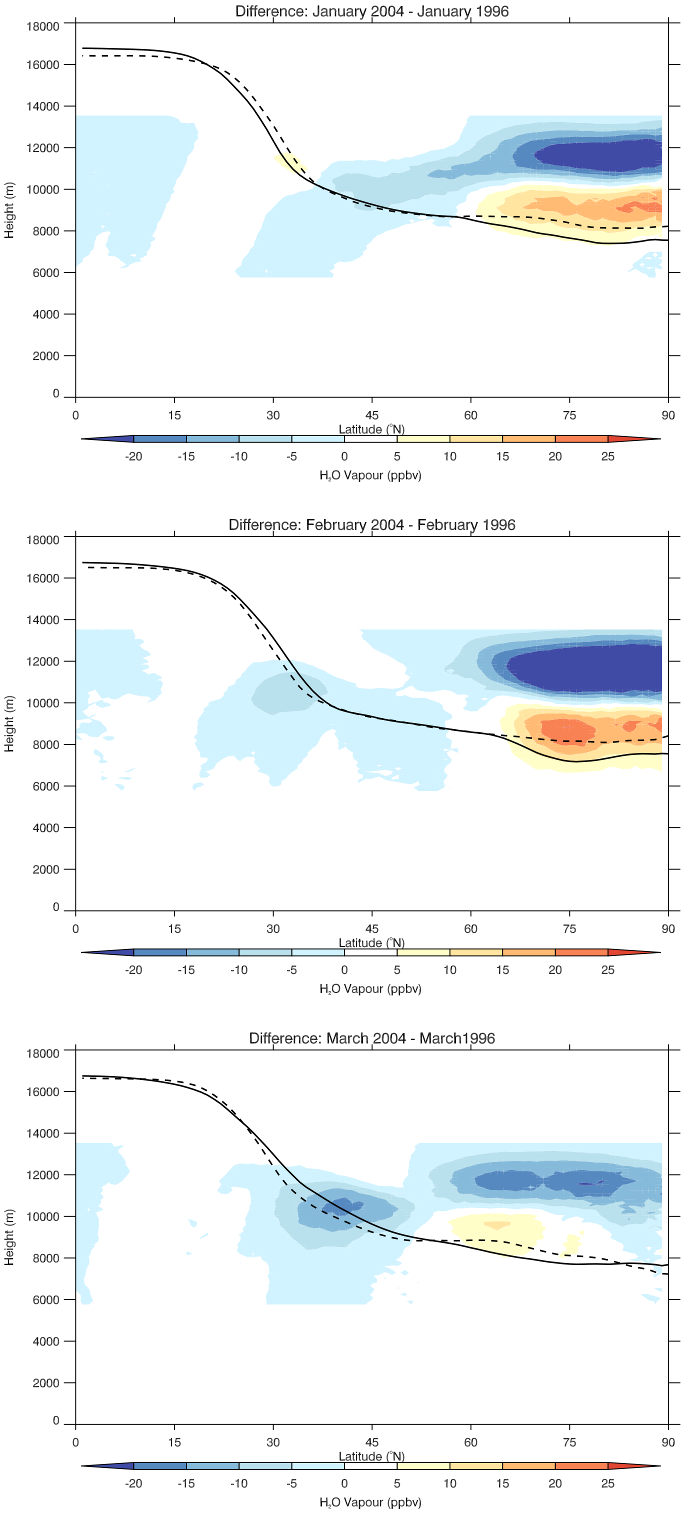


Figure 7.8: *Difference in January to March zonal-mean monthly-mean perturbations calculated DJFM 2003/2004 and DJFM 1995/1996 meteorology. The solid line indicates the position of the 2004 monthly mean tropopause, the dashed line indicates the position of the 1996 tropopause. Input emissions are from AEDT 2006.*

tropopauses, with negative differences where the 1996 zonal-mean tropopause is lower than its 2004 equivalent, resulting in more accumulation at lower altitudes, and vice versa. The negative difference between the two perturbations at altitude is the result of greater stratospheric deposition, and stronger ascent between 60°N and 80°N, in 1996. In March 1996 there is a lower tropopause over the North Atlantic, Europe, and North America compared to 2004, resulting in more emissions entering the stratosphere directly. Greater ascent lofts these emissions to higher altitudes compared to 2004. There is still a negative difference between the 2004 and 1996 perturbation near the pole, despite there being descent here in 1996, compared to ascent in 2004. This is probably the result of larger concentrations of emissions being transported to high latitudes and subsequently being removed when 1996 meteorology is used.

Peak zonal-mean perturbations of 66, 65, and 75 ppbv occur in January to March respectively. The accumulated mass is largest in February, with monthly-mean accumulations of 5.53, 6.04, and 5.96 Tg occurring in January to March (Table 7.1).

7.2.4.3 DJFM 1997/1998: El Niño

In addition to the strong El Niño, the NAO index was slightly positive in DJFM 1997/1998. The differences between the perturbations calculated using DJFM 1997/1998 meteorology and AEDT 2006 emissions, and those calculated using DJFM 2003/2004 meteorology, are shown in Figure 7.9. The perturbations calculated using DJFM 1997/1998 meteorology are more similar to the DJFM 2003/2004 case than those calculated using DJFM 1994/1995 or DJFM 1995/1996 meteorology. However, there are still some differences of note.

In January, the difference between the 2004 and 1998 perturbations has a similar structure to the differences seen in Figures 7.7 and 7.8, although the magnitude of the differences are smaller. The perturbation from 1998 meteorology is larger than the 2004 base case, with the structure of the difference following the structure of the perturbation in the extratropics (Figure 7.9). This larger perturbation is likely to be the result of stronger

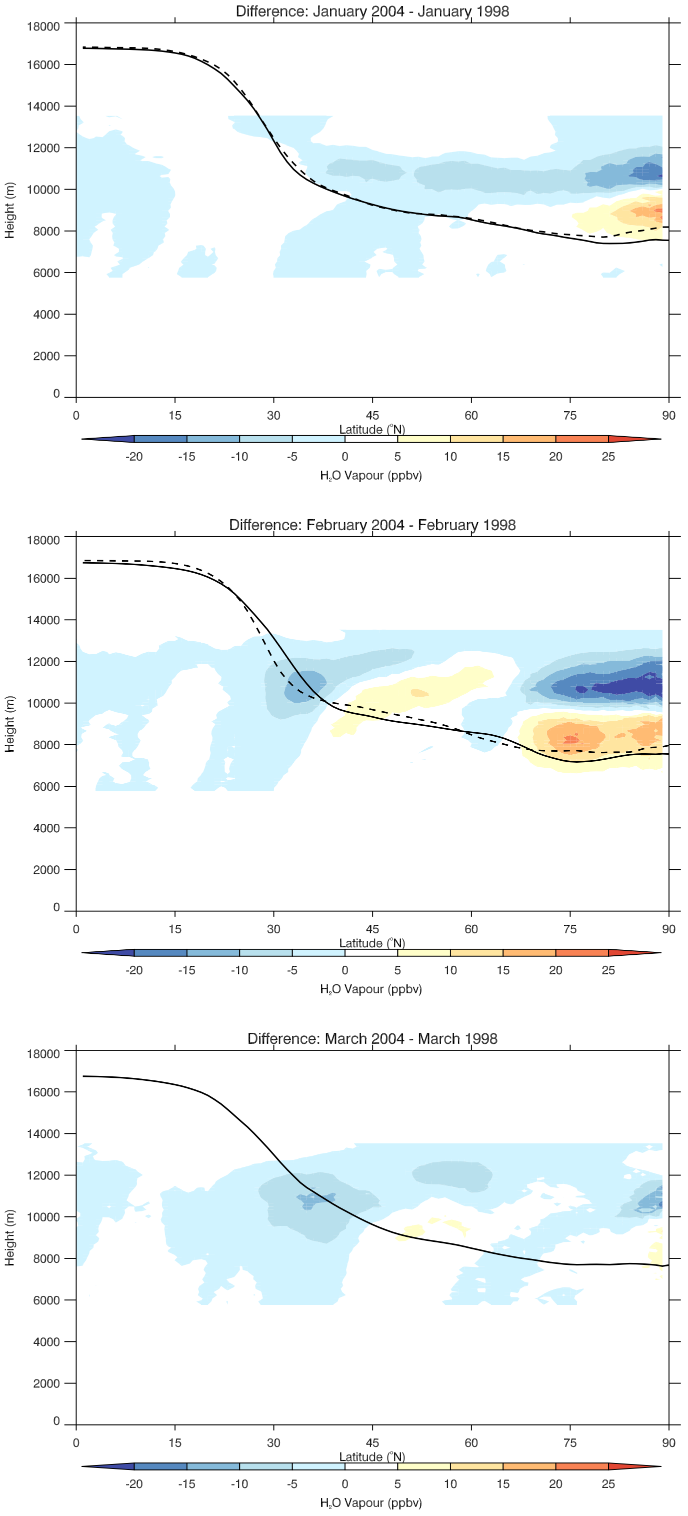


Figure 7.9: *Difference in January to March zonal-mean monthly-mean perturbations calculated DJFM 2003/2004 and DJFM 1997/1998 meteorology. The solid line indicates the position of the 2004 monthly mean tropopause, the dashed line indicates the position of the 1998 tropopause. Input emissions are from AEDT 2006.*

ascent from 35°N to 45°N, where a large amount of deposition occurs, and weaker descent at high latitudes, relative to 2004. The zonal-mean tropopause is similar in the two cases. However, the 1998 tropopause is slightly higher in the high latitudes, causing the positive anomaly here.

The dipole structure in the high latitudes is still present in February (Figure 7.9). It is the result of ascent and a high tropopause near the pole, compared to descent and a lower tropopause in 2004. There is a negative difference at the tropopause near 30°N, resulting from the lower tropopause in 1998, and a positive difference near 45°N, where the 1998 zonal-mean tropopause is higher than its 2004 equivalent. However, this perturbation is not centred on the tropopause; it extends upwards and polewards. This is likely to be the result of the descent, and therefore more rapid removal, centred near 45°N in 1998, compared to ascent in 2004.

In March, the differences between the 2004 and 1998 perturbations are generally small. In most places they are an order of magnitude less than the original perturbation (Figure 7.9). The structure seen in February can still be seen in March, but the magnitude is reduced.

There is a degree of inter-annual variability in the magnitude and structure of the perturbations. The primary drivers of this variability in the zonal-mean perturbation appear to be the position of the zonal-mean tropopause and the relative signs of the zonal-mean vertical velocities. However, there is no obvious link between the phase of the NAO and the structure of the zonal-mean perturbation. The strongly negative and positive NAO case studies resulted in perturbations that were more similar to each other than they were to the base case.

The maximum zonal-mean monthly-mean perturbations for each of the months considered, and the latitudes at which they occur, are summarised in Table 7.1. The perturbations have a spread of 10% of the maximum perturbation in January, and 5% in February and March. The latitudes of these maxima are extratropical in January. However, in February and March, there is a suggestion of a bi-modal structure to the latitudes, with maxima either in the extratropics or near the pole. The monthly-mean accumulated mass in the stratosphere is also shown in Table 7.1. The accumulated mass has a spread of 3% of the maximum accumulation in January, 6% in February, and 13% in March.

Peak zonal-mean perturbations of 66, 66, and 71 ppbv occur in January to March respectively. The accumulated mass is also largest in March, with monthly-mean accumulations of 5.56, 5.67, and 5.74 Tg occurring in January to March (Table 7.1).

7.2.4.4 JJAS 2006

Perturbations have been calculated for JJAS 2006, using high resolution ERA-Interim trajectories and AEDT 2006 meteorology. The perturbations for July to September are shown in Figure 7.10. It can be seen from comparison of Figures 7.10 and 7.2 that the perturbations have a similar structure to the JFM cases, although they have a smaller magnitude in the summer months. The shape of the perturbation is still very much influenced by the position of the zonal-mean tropopause. This is particularly apparent in the August perturbation (Figure 7.10). As in the winter examples, the zonal-mean perturbation extends below the zonal-mean tropopause.

The maximum zonal-mean perturbations are smaller in JAS compared to JFM, and located at higher latitudes. The maximum zonal-mean perturbations are 47, 60, and 56 ppbv respectively, at 57, 59, and 66°N.

7.2.5 The annual cycle in perturbations

Perturbations have also been calculated using low-resolution ERA-Interim trajectory calculations for 2006, and 6 hourly AEDT emissions. The zonal-mean annual-mean perturbation is shown in Figure 7.11. An equivalent plot from Fichter (2009), calculated using the TRADEOFF 2000 inventory and the E39CA GCM that includes a fully Lagrangian advection scheme, ATTILA (Atmospheric Tracer Transport In a LAgrangian model), is also shown. It should be noted that while the perturbation presented in this thesis is a volume mixing ratio (in ppbv), the result from Fichter (2009) is a mass mixing

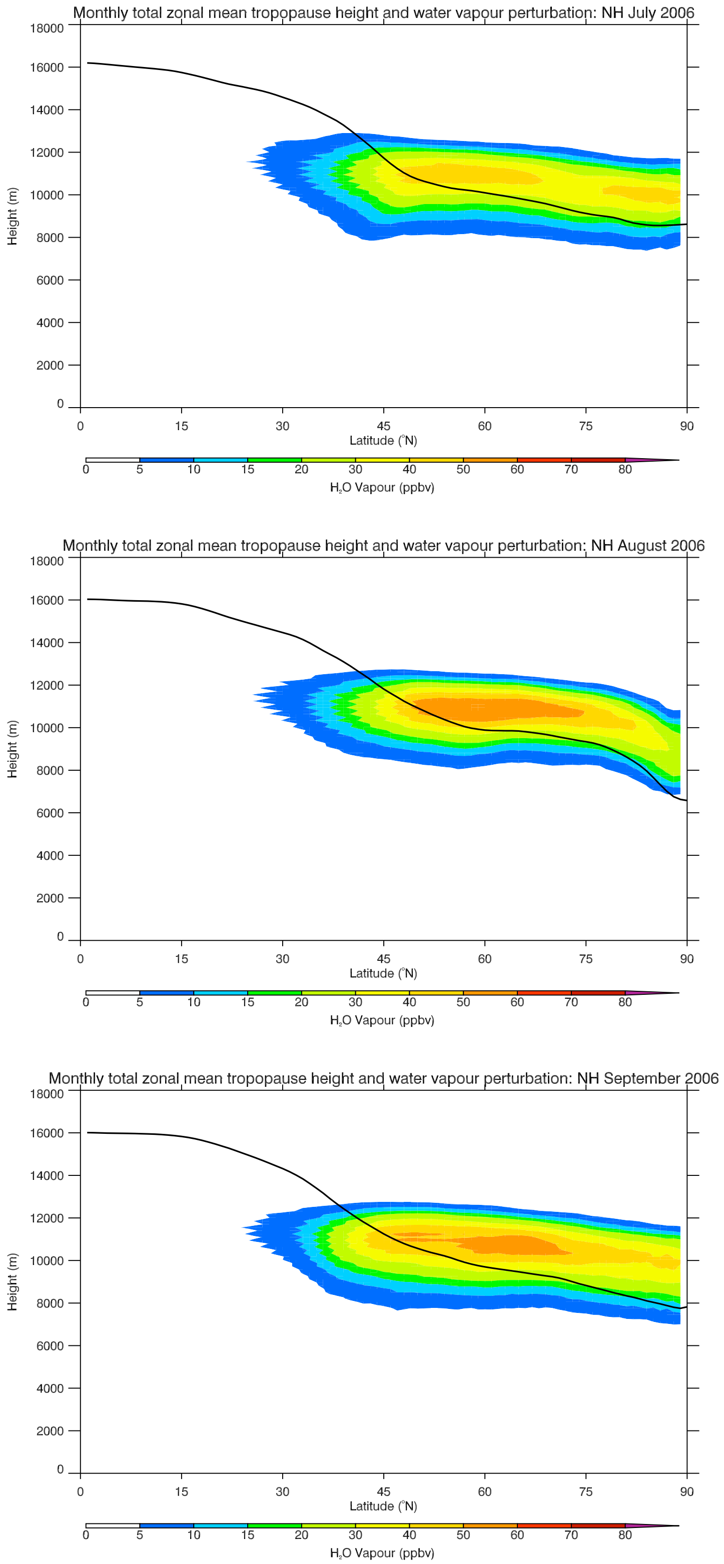


Figure 7.10: *Zonal-mean monthly-mean perturbations (ppbv) for July (top), August (middle), and September (bottom) 2006. The position of the zonal-mean monthly-mean tropopause is indicated by the position of the black line. Input emissions are from AEDT 2006.*

ratio (in ppbm). There is a factor of 1.6 between these two units, making the maximum contour at 35 ppbm equivalent to 56 ppbv in the lower panel of Figure 7.11.

It can be seen from comparison of the two panels in Figure 7.11 that the perturbations have broadly the same structure. However, there are some notable differences. The zonal-mean tropopause in this work always slopes downwards from the Equator to the pole, and the perturbation follows this. The E39CA tropopause is more horizontal in the midlatitudes, and this is again reflected in the shape of the perturbation. The perturbations both extend from near 25°N to the pole, and have some extension below the tropopause at all latitudes. However, the perturbation from Fichter (2009) appears to extend further below the tropopause than that calculated here. This is not surprising, as there is no explicit removal of tropospheric water vapour emissions in E39CA. Emissions have the same loss processes as natural water vapour: cloud physics, rainout, and chemical conversion.

The perturbation calculated in this thesis is slightly smaller than that calculated by Fichter (2009) (a maximum zonal-mean annual-mean perturbation of 64 ppbv compared to 69 ppbv). This is a somewhat surprising result as the TRADEOFF inventory has a smaller total mass of water vapour emissions (190 Tg(H_2O) yr^{-1} compared to 232 Tg(H_2O) yr^{-1} in AEDT 2006). It also has a similar emissions distribution to the QUANTIFY inventory, which was found to result in a smaller perturbation compared to AEDT in Section 7.2.2. The perturbation calculated here using the TRADEOFF inventory and the UGAMP trajectories is shown in Figure 7.11. This perturbation is, as expected, 30% smaller than the perturbation calculated using AEDT emissions.

The 30% smaller perturbation found here using TRADEOFF emissions, and the UGAMP trajectory model, compared to the results of Fichter (2009) implies that emissions are removed from the stratosphere faster in this work. A small difference in the rate of removal may arise from the difference in meteorology, but the difference is likely primarily to be due to differences in transport within the two models, although both of the models use Lagrangian transport of emissions. Additionally, the E39CA tropopause is at a slightly higher altitude compared to the blended tropopause in the high latitudes,

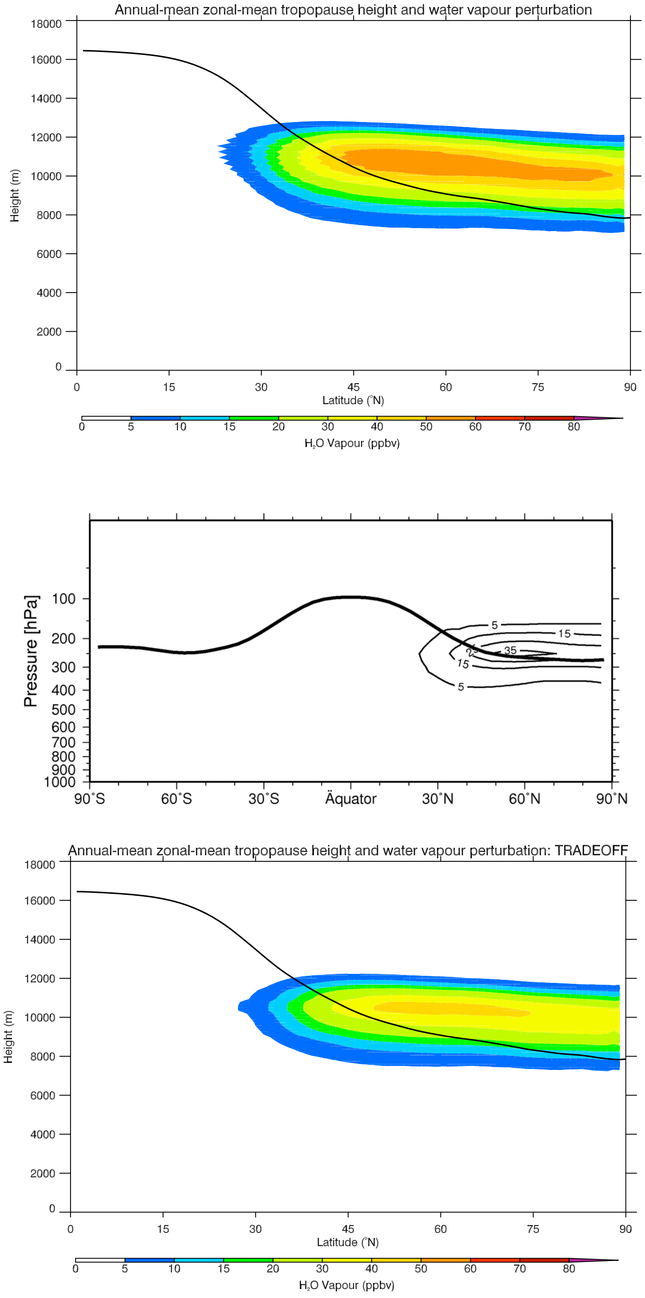


Figure 7.11: *Top: Zonal-mean annual-mean perturbation (ppbv) for AEDT 2006. The position of the zonal-mean monthly-mean tropopause is indicated by the position of the black line.*

Middle: Zonal-mean annual-mean perturbation (ppbm) for TRADEOFF 2000 emissions, as calculated using E39CA (from Fichter, 2009).

Bottom: Zonal-mean annual-mean perturbation (ppbv) for TRADEOFF 2000, using 2006 trajectories.

but at comparable altitude in the subtropics, which can also be expected to result in less accumulation of emissions in the stratosphere. This is not the case.

One likely explanation for this apparent discrepancy between the perturbation calculated in this thesis and that calculated by Fichter (2009) is the difference in the spatial resolution of the two models. In this thesis, transport is modelled on a $1^\circ \times 1^\circ \times 152$m grid (50 levels between 466 and 151 hPa). Fichter uses T30 horizontal resolution ($3.75^\circ \times 3.75^\circ$), with 39 vertical levels up to 10 hPa. Kentarchos et al. (2000) performed a comparison of stratosphere to troposphere transport in the semi-Lagrangian ECHAM4 model at T30 and T63 ($1.875^\circ \times 1.875^\circ$) horizontal resolution, with 19 levels in the vertical. They identified finer scale structures in the output from the calculations with higher horizontal resolution, and found them to be in better agreement with analysis data. They also found stronger maximum jet speeds, and, as a result, stronger vertical mixing. This resulted in stratospheric air penetrating further into the troposphere. Kentarchos et al. (2000) found that the change from T30 to T63 resolution resulted in a $\sim$12% increase in the amount of stratospheric ozone transported from the stratosphere into the troposphere in their case study. Similarly, a 35% increase in downwards transport was found by Van Velthoven and Kelder (1996) when they increased the resolution of their offline transport model from $8^\circ \times 10^\circ$ to $4^\circ \times 5^\circ$.

Rogers et al. (2002) tested the sensitivity of emissions transport to both horizontal and vertical resolution. They also found an increase in tracer transport when higher horizontal resolutions were used. However, when they compared results from ECHAM4 with 19 and 39 vertical levels, they found that increasing the vertical resolution reduced the exchange of tracers from the stratosphere to the troposphere. The increased vertical resolution improved the ability of the semi-Lagrangian model to maintain the vertical gradients of the tracer. As excessive numerical diffusion does not affect Lagrangian models, such an effect will not necessarily play a role in the comparison between the results of two fully Lagrangian models.

There is a strong seasonal cycle in both the maximum zonal-mean monthly-mean perturbation, and the total mass of emissions in the stratosphere (see Figure 7.12). The largest maximum zonal-mean perturbation of 85 ppbv occurs in April, and the smallest occurs in July (50 ppbv). The annual mean is 64 ppbv. This cycle mirrors the cycle in the total mass of emissions in the stratosphere, except that the maximum mass in the stratosphere (6.08×10^{12}g) occurs in March.

There is no clear cycle in the latitude at which the maximum zonal-mean perturbation is located (Figure 7.12). However, there is a suggestion of two regimes: one with a maximum perturbation in the extratropics, and one with a maximum over the pole. In addition to the polar maxima seen in the annual cycle, polar maxima also occurred in February 1996 and 1998, and March 1998. Forster et al. (2003) also found that stratospheric emissions had a maximum concentration over the pole 10-20 days after their release.

The resolution of the meteorological input data appears to have a small influence on the latitude of the maximum zonal-mean monthly-mean perturbation. When low resolution ERA-Interim data was used to calculate trajectories for the whole of 2006, maximum zonal-mean monthly-mean perturbations of 50, 60, and 56 ppbv were found at 89, 64, and 66°N in July, August, and September. When full resolution ERA-Interim data was used the maximum zonal-mean monthly-mean perturbations were 47, 60, and 56 ppbv, at 57, 59, and 66°N. The resolution of the ERA-Interim data used in the calculation of the trajectories makes no difference to the September results, and only a small difference to the August results. However, in July the maximum zonal-mean monthly-mean perturbation is smaller, and located at extratropical latitudes instead of the pole, when the high resolution data is used. It can be seen in Figure 7.10 that there is a bi-modal structure to the perturbation in July when high resolution data is used to calculate the trajectories. This means that a small change in the concentration of water vapour emissions at each location has the potential dramatically to change the latitude at which the maximum zonal-mean concentration occurs. When the difference in the latitude at which the zonal-mean monthly-mean maximum perturbation occurs is viewed in the context of Figure 7.10, then it is easy to view it as a small difference.

A similar analysis of the annual cycle in stratospheric perturbations was performed by

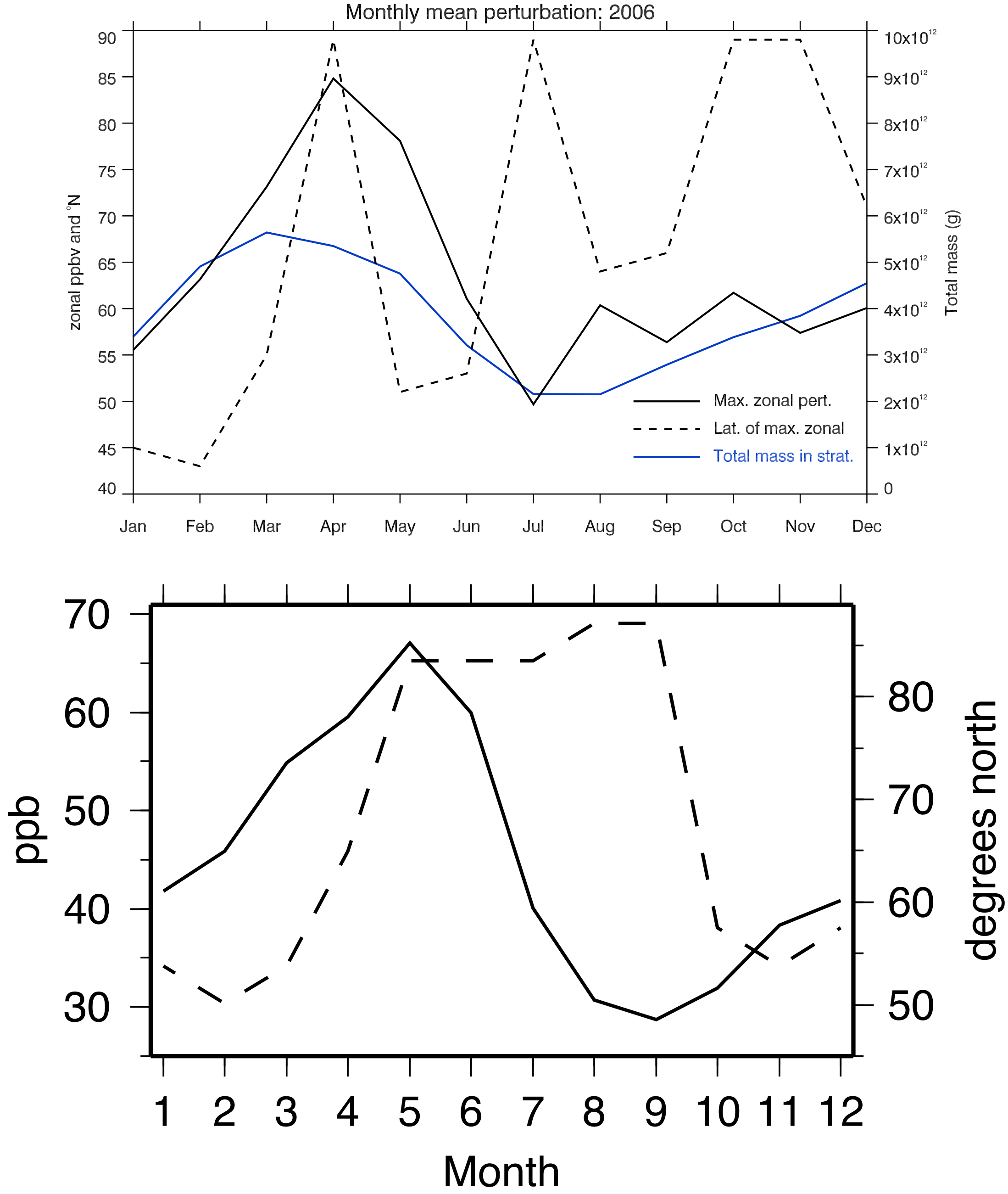


Figure 7.12: *Top: Maximum zonal-mean monthly-mean perturbation, and the latitude at which it occurs, for perturbations calculated using AEDT 2006 emissions and low resolution ERA-Interim trajectories. The monthly total mass in the stratosphere is also shown.*

Bottom: Maximum zonal-mean monthly-mean perturbation, and the latitude at which this occurs, from Fichter (2009).

Fichter (2009), who found a maximum zonal-mean perturbation equivalent to 107 ppbv in May, and a minimum equivalent to 45 ppbv in September (Figure 7.12). This cycle has a larger amplitude, and is shifted by two months, relative to the cycle found in this thesis. Fichter also identified a bi-modal distribution in the latitude of the maximum zonal-mean perturbation. However, her results are more indicative of an annual cycle. When TRADEOFF emissions were used with 2006 trajectories, the latitude of the maximum zonal-mean perturbation had a similar bi-modal structure to when AEDT emissions were used. Hence, the latitude at which the maximum perturbation is located is determined by meteorology and not the location of input emissions.

7.3 Comparison with earlier estimates

Estimates of the perturbation to background humidity due to aviation water vapour emissions have been given in a number of previous studies, in addition to Fichter (2009). A wide range of estimates, calculated using a variety of techniques, have been presented.

Hoinka et al. (1993) used Civil Aviation Authority (CAA) flight tracks across the North Atlantic, and the assumption of a 0.5 year stratospheric lifetime, to find a 'daily surplus' of 380 ppbv water vapour in the NAFC stratosphere. They defined their NAFC as the region bounded by 45°N, 65°N, 10°W, and 50°W. An even larger estimate of the stratospheric water vapour perturbation was derived by Fortuin et al. (1995) who calculated mass loading from the Organisation for Economic Co-operation and Development (OECD) records of fuel burn for 1971 to 1991. In this calculation they included a range of assumed emissions lifetimes, ranging from 0.5 to 2.5 years in the stratosphere, which gave a range of stratospheric perturbations from 76 to 380 ppbv. The values presented in both of these studies are very large compared to the perturbations found in this thesis. However, this is not unexpected given the very long emissions lifetimes used, which are at least 6 times larger than the typical values calculated in this work.

Ponater et al. (1996) and Rind et al. (1996) performed similar studies, where they multiplied an emissions scenario by increasing factors in an attempt to detect an impact on climate. In both cases they give the perturbation as a number of orders of magnitude less than the background concentration, which makes a simple comparison difficult.

A model intercomparison was performed by Danilin et al. (1998), using four 2-D models, and seven 3-D models, with the NASA 1992 inventory. In all the models used, the only sink was globally uniform removal of the emissions below 400 hPa. Danilin et al. (1998) identified the maximum tracer mixing ratio between 50 and 65°N, near to 10 to 12 km altitude, in all but one of their models. Such a distribution is consistent with the emissions distribution in the NASA inventory, and agrees well with the location of the maximum zonal-mean perturbations calculated in this work. However, their Figure 1 reveals some large differences in the structure of the perturbation between models. In most cases, the maximum perturbation occurs at cruise altitude, and extends through most of the extratropical and high latitudes, but there are marked differences in the vertical extent of the perturbations between some of the models.

Danilin et al. (1998) were investigating whether aviation water vapour emissions could be responsible for observed trends in stratospheric water vapour at 40°N. The largest perturbation at this latitude across all the models was 55 ppbv at 10 km, which decreased to 12 ppbv at 24 km. This perturbation is very slightly larger, and at lower altitude, compared to the perturbations calculated in this thesis, possibly due to the use of 400 hPa as a removal criterion, which is lower than the annual-mean blended tropopause.

Gauss et al. (2003) evaluated the impact of water vapour emissions from subsonic aircraft, and from a hypothetical fleet of liquid hydrogen fuelled ‘cryoplanes’. The NASA 2015 inventory was used to provide the input distribution of the subsonic emissions, and a chemical transport model was used to calculate the perturbation due to several future scenarios. For the cryoplane base case (NASA 2015 emissions multiplied by 2.55), the water vapour perturbation had a zonal-mean maximum of 479 ppbv and a zonal-mean annual-mean maximum of 279 ppbv. For the subsonic base case (H_2O-K1), Gauss et al. (2003) found a maximum zonal-mean perturbation of 188 ppbv, and a zonal-mean annual-mean maximum of 109 ppbv. Both of these values are larger than the perturbations found in this thesis, as is to be excepted given the future scenarios. The NASA

2015 inventory includes an emission rate of 400 Tg H_2O yr^{-1}, compared to 233 Tg H_2O yr^{-1} in AEDT 2006. Applying a factor of 0.58 to scale the NASA 2015 values to AEDT 2006 values, gives a maximum zonal-mean annual-mean perturbation of 63 ppbv. This is almost identical to the annual-mean value of 64 ppbv found in this thesis.

The structure of the perturbation in Figure 5 of Gauss et al. (2003) is similar in the extra-tropics to that of the perturbations presented in this thesis. However, in the subtropics, the perturbation extends upwards along the tropopause, with a very small fraction of the emissions being lofted high into the stratosphere in the tropics.

Morris et al. (2003) used a trajectory model to evaluate the impact of aviation water vapour on the lower stratosphere. They used the Goddard trajectory model to advect water vapour as a passive tracer, as we have used the UGAMP model. They removed the trajectories when they were 0.5 km or more below the tropopause. Morris et al. (2003) define their tropopause as the lesser of the height at which the minimum temperature occurs in a profile and the height of the 3 PVU surface. Using the NASA 2015 inventory, they found a perturbation to the natural background just above the tropopause, and mostly polewards of 30°N, as was found in this thesis. Like the result from Gauss et al. (2003), the perturbation presented by Morris et al. (2003) has a small degree of extension along the tropopause in the subtropics. However, the magnitude of the zonal-mean perturbation in this region is at least an order of magnitude smaller than that just above the tropopause in the extratropics.

Morris et al.'s perturbation had maximum values of up to 100 ppbv. Applying the 0.58 scaling, this value has an AEDT 2006 equivalent of 58 ppbv, which is slightly smaller than the perturbation found in this thesis. This is likely to be due, at least in part, to their use of a 3 PVU tropopause, which will be higher than the blended tropopause (2 PVU in the extratropics). 58 ppbv is within the range of the annual cycle of maximum zonal-mean perturbations found in this thesis, which has a minimum of 50 ppbv.

The perturbations found here are consistent with those presented by Gauss et al., (2003), Morris et al. (2003), and the maximum perturbation from the models compared by Danilin et al. (1998). The perturbations presented in this thesis have a similar structure to those from Fichter (2009). However, when the same emissions inventory is used, the concentrations are about 30% smaller. Some of this difference may arise from interannual variability in meteorology. However, it is likely that the main cause of the difference is due to differences in the transport models themselves, possibly the large differences in vertical and horizontal resolution.

7.4 Lifetime estimates

Estimates of the stratospheric lifetime of stratospheric emissions, based on the assumption of exponential decay, are presented in some previous studies. E-folding lifetimes calculated from trajectories were presented in Chapter 6 for comparative purposes. However, analysis of the trajectories revealed that the assumption of an exponential decay in emissions was not necessarily appropriate, especially for high altitude emissions where decay was found to be linear.

A more realistic estimate of the stratospheric lifetime of emissions can be calculated by dividing the accumulated emissions by the input emissions. As this estimate is found from the perturbation directly, there is no need for assumptions about the nature of the decay in the input emissions.

It can be seen in Figure 7.13 that the distribution of accumulated water vapour (in grams) is smoother compared to the distribution of input emissions (grams per month), and has a peak at a slightly lower altitude, which is consistent with the zonal-mean perturbations shown in earlier Figures, and the general descent of emissions after their release. The accumulated emissions are about half as large as the input emissions, suggesting an average stratospheric lifetime of 15 days. This is much less than the e-folding lifetimes calculated for NAFC trajectories, reflecting the fact that a significant proportion of emissions enter the stratosphere very close to the tropopause.

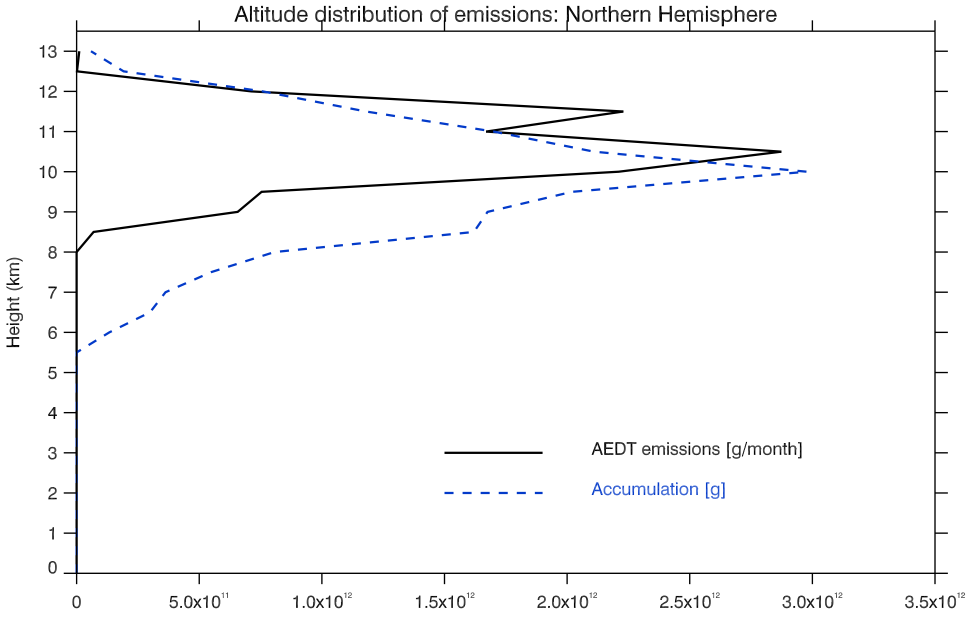


Figure 7.13: *Total AEDT emissions for January 2006 (solid black line) and calculated stratospheric accumulation of water vapour, from January 2004 trajectories and AEDT January 2006 emissions, multiplied by 3 for comparison (dashed blue line).*

7.4.1 Summary

Limits on the memory requirements of trajectory calculations mean that in order to have a release grid with the same resolution as the AEDT inventory, trajectories can have only a maximum length of 2 days. Because a regular release grid is used, such short trajectories can be combined to advect emissions until they are removed from the stratosphere. The accumulation of emissions in the stratosphere can be calculated directly from such trajectories, removing the need to make assumptions about emission lifetime.

Two day long trajectories have been calculated using ERA-Interim for DJFM 2003/2004, DJFM 1994/1995, DJFM 1995/1996, DJFM 1997/1998, and JJAS 2006, with a release grid of $1^\circ \times 1^\circ \times 152$ m$\times$6 hr for the whole Northern Hemisphere between 466 hPa and 151 hPa. Trajectories were also calculated for DJFM 2003/2004 using ERA-Operational data, and ERA-Interim data at ERA-40 resolution. Trajectories were calculated for the whole of 2006 using ERA-Interim data at ERA-40 resolution.

The perturbation due to aviation water vapour emissions was found to be sensitive to the emissions inventory used in the calculation. The use of a different inventory resulted in differences of up to 50% of the AEDT perturbation. This sensitivity is a reflection of the very different vertical profiles of emissions found in the inventories considered.

Some differences were apparent between perturbations calculated using ERA-Interim and ERA-Operational trajectories. ERA-Interim perturbations were slightly larger than those from ERA-Operational trajectories near to the tropopause in the high latitudes. However, these differences were much smaller than the differences between the perturbations calculated using different emissions inventories.

There was found to be a degree of inter-annual variability in both the magnitude and structure of the calculated perturbation across the four winters considered. In February and March, maximum zonal-mean monthly-mean perturbations were up to 5% smaller than the maximum for that month in the years considered. This increased to 10% for January. Again, these differences are smaller than the differences arising from the choice of emissions inventory. The differences between the perturbations were consistent with the differences between the large-scale wind fields, but there was no obvious correspondence between the differences between the perturbations and the NAO index. In fact, the two extreme opposite NAO cases considered resulted in perturbations with quite similar structures.

A strong annual cycle was identified in both the total mass of emissions accumulated in the stratosphere, and the magnitude of the maximum zonal-mean perturbation. For 2006, the annual cycle in the mass of accumulated emissions has a maximum in March and a minimum in July. The cycle in the maximum zonal-mean perturbation is slightly skewed with respect to the total accumulation cycle, with a maximum of 85 ppbv in April and a minimum of 50 ppbv in July. The cycle in the maximum zonal-mean perturbation is comparable to the cycle identified by Fichter (2009) using the TRADEOFF inventory and EC39A, although it has a slightly smaller amplitude.

The structure and magnitude of the zonal-mean emissions are comparable with the recent work of Fichter (2009), Gauss et al. (2003), and Morris et al. (2003), and the magnitude of the maximum perturbation found by Danilin (1998). The magnitude of the perturbations were much smaller than those presented by Hoinka et al. (1993), and Fortuin et al. (1995), who assumed that water vapour emissions had extremely long lifetimes in the stratosphere.

The stratospheric lifetime of emissions has been estimated by comparing the accumulated mass to the input mass for January. This avoids the need to make any assumptions about the nature of the time-evolution of emissions after their release. The lifetime was found to be 15 days. This is shorter than most e-folding lifetime estimates of the stratospheric lifetime of emissions, and is in best agreement with the e-folding lifetimes calculated for low altitude emissions in the NAFC and over the USA.

Chapter 8

Radiative forcing due to aviation water vapour emissions

Radiative forcing is established as a useful predictor of climate change. It is commonly used in intercomparisons of the climate impact of aviation (e.g. Lee et al., 2009, Sausen et al., 2005) and in the IPCC assessments. It is accepted to give a reasonable first-order approximation of climate change, without the need for computationally expensive climate model calculations.

The definitions of radiative forcing used in this thesis, and the application of radiative forcing, are introduced in Section 8.1. The mechanisms and characteristics of radiative forcing due to stratospheric water vapour are discussed in Section 8.2. The model and method used to calculate radiative forcing in this work is given in Section 8.3, before the results of the calculations and a comparison with previous estimates is given in Section 8.4.

8.1 Definitions of radiative forcing

The concept of radiative forcing was developed for considering simple climate change mechanisms in one-dimensional models. It has since been extended to cover different spatial and temporal scales, and more complex changes (Forster et al., 2007). This has led to some debate as to the relevance of radiative forcing. However, it still appears to be a useful first-order method of comparing climate change mechanisms (Forster et al., 1997, Hansen et al., 2005).

The simplest definition of radiative forcing is the 'instantaneous radiative forcing': the change in net irradiance following a perturbation, with no changes to any other atmospheric variables. It is typically evaluated at the tropopause. Hansen et al. (2005) note that the instantaneous radiative forcing is, in a sense, a true forcing, because the climate is held fixed when it is calculated. However, it can be unrepresentative of the climate response, and it is less relevant for considering the response to forcings that act for long time periods.

The 'standard definition' of radiative forcing, used by the IPCC, and many other studies, is the 'adjusted radiative forcing'. The adjusted radiative forcing is the change in net irradiance following a perturbation, after the stratospheric temperature has returned to equilibrium. The adjustment of the stratosphere means that this definition gives a better indication of the expected climate response (Hansen et al., 2005). The adjustment of the stratosphere is a fast process, with the stratosphere typically returning to equilibrium in months following the application of the perturbation. Only the stratosphere is adjusted because the troposphere is influenced by the large thermal inertia of the oceans, and can take decades to reach equilibrium.

The adjusted radiative forcing $\Delta\mathrm{F}_a$ is related to the change in surface temperature $\Delta\mathrm{T}_s$ by the climate sensitivity, λ:

$$\Delta T_s \simeq \lambda \Delta F_a \tag{8.1.1}$$

It is more useful to compare changes in the adjusted forcing than in surface temperatures because the climate sensitivity parameter is poorly known, and varies between models

(Solomon et al., 2007). The climate sensitivity is often roughly constant for a given model, regardless of the nature of the forcing. However, it can vary significantly for some forcing mechanisms within models (Joshi et al., 2003). The adjusted radiative forcing has the additional advantage that it is more readily computed than the change in surface temperature, which involves multidecadal GCM runs, especially for small forcings.

Both the instantaneous and adjusted radiative forcing are sensitive to the position of the tropopause (Forster et al., 1997, Solomon et al., 2010). Alternative definitions of radiative forcing have been suggested that avoid such a sensitivity. Shine et al. (2003) and Hansen et al. (2005), for example, discuss alternative definitions where the top-of-atmosphere perturbation is calculated in a GCM, both the troposphere and stratosphere are adjusted, but sea surface, or surface, temperatures are kept fixed. This avoids introducing a tropopause dependence, and the fixed sea surface temperatures mean that very long GCM integrations are not required.

8.2 Radiative properties of water vapour

The water molecule absorbs and emits energy by changing its internal energy. This can occur via several processes, including changing the rotational kinetic energy of the molecule and changing its vibrational energy. The asymmetric shape of a water molecule, and its resulting permanent electric dipole moment, means that it has major rotational absorption bands. Rotational energy structures are closely spaced, so only low energy photons can cause a change in a molecules rotational kinetic energy (Petty, 2004). For water vapour the rotational energy bands are stronger than the vibrational bands. Hence, stratospheric water vapour absorbs radiation primarily at infrared (IR) wavelengths (see Figure 8.1) in the spectral region from 0 to 350 cm^{-1}.

The energy states of a molecule are quantised. In theory, this means that a photon can only be absorbed if it has the correct amount of energy (inversely proportional to its wavelength) to cause a transition between a pair of allowed states. However, in reality

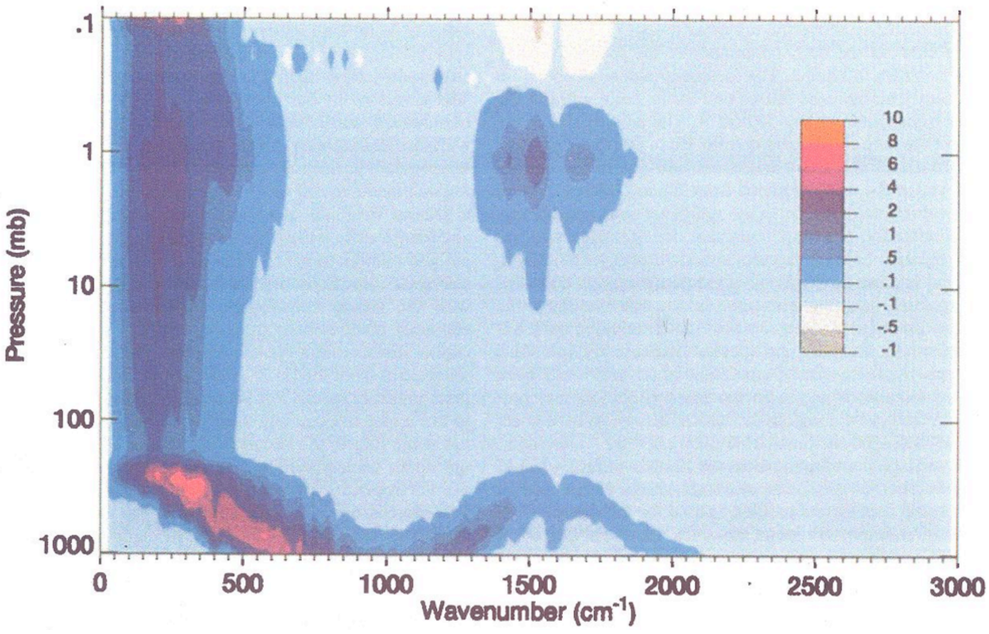


Figure 8.1: *Spectral cooling rate profile for water vapour in a midlatitude summer atmosphere with clear skies, from Clough et al. (1995). The colour scale is* $\times 10^{-3}$ *K* d^{-1} *(*cm^{-1}*)*$^{-1}$*. In the stratosphere, cooling mostly occurs at wavenumbers less than 350* cm^{-1}*, which are characteristic of transitions in the rotational energy of molecules.*

absorption does not just occur at a single energy (wavelength). The process of ‘broadening’ means that radiation at similar energies can also be absorbed by a molecule. In the troposphere and stratosphere, the primary broadening mechanism, ‘pressure broadening’ occurs when collisions between molecules disrupt the basic transitions between energy states, allowing photons with wavelengths either side of the theoretical value to be absorbed (Petty, 2004). Broadening can also occur via ‘Doppler broadening’ where the velocity of molecules introduces a Doppler shift into the frequency of photons that it can emit and absorb. Doppler broadening is the primary broadening mechanism in the mesosphere and above. The process of ‘natural broadening’ also occurs, resulting from the finite width imposed on the absorption line by the Heisenberg uncertainty principle. However, the broadening from this mechanism is negligible compared to that from pressure and Doppler broadening (Petty, 2004).

The emission of IR radiation by stratospheric water vapour acts to decrease stratospheric temperatures, while the increased emission of IR by the stratosphere acts to increase

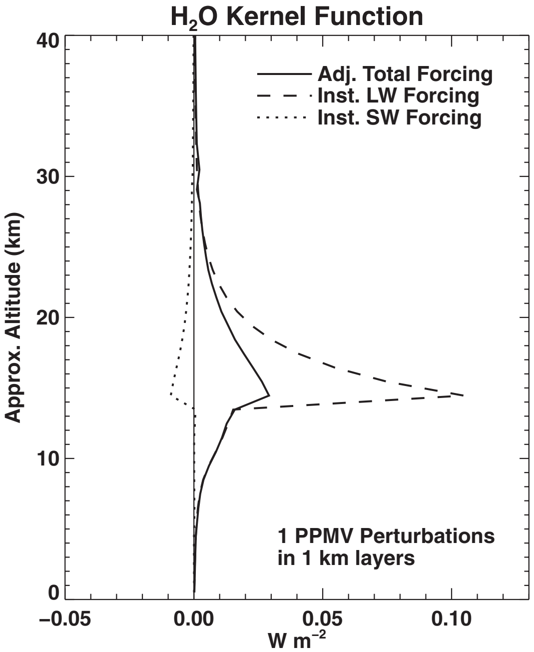


Figure 8.2: *Instantaneous longwave and shortwave radiative forcing, and adjusted net forcing, versus altitude for uniform 1 ppmv perturbations in 1 km layers at 35°N, from Solomon et al. (2010).*

tropospheric temperatures. The magnitude of the temperature change in the stratosphere is dependent on the emission by the stratosphere and absorption of radiation emitted by the troposphere. The radiative forcing due to water vapour perturbations depends on the change in emission by water vapour, and the induced temperature change.

Forster and Shine (1999, 2002) identified that climate is most sensitive to water vapour perturbations just above the tropopause. Solomon et al. (2010) investigated the effect on radiative forcing of stratospheric water vapour perturbations at different altitudes by calculating the radiative forcing per layer, the ‘kernel function’, for 1 ppmv perturbations in 1 km thick layers. It can be seen in Figure 8.2 that the largest sensitivity occurs close to the tropopause (typically 12 km at 35°N). It can also be seen that the stratospheric temperature adjustment acts to reduce the impact of local changes in the lower stratosphere, as the reduction in temperatures there will reduce the amount of re-radiated IR from the stratosphere to the troposphere.

8.3 Calculation of radiative forcing

In this thesis, the flux perturbations due to aviation water vapour emissions are calculated using the Edwards and Slingo radiative transfer model, ES (Edwards and Slingo, 1996). ES is a broadband code that calculates longwave and shortwave radiative fluxes and heating rates for an atmospheric column, given the properties of that column.

ES is based on the two-stream approximation to the radiative transfer equations. There are several variants of the two-stream equations, and many of these can be used in ES (Edwards and Slingo, 1996). In this thesis the recommended settings are used. For shortwave calculations this is the ‘Practical Improved Flux Method’ from Zdunkowski (1980), described by Edwards and Slingo (1996). For longwave calculations the 1985 version of the Practical Improved Flux Method is used, with Elsasser’s value of 1.66 for the diffusivity factor (Edwards and Slingo, 1996).

ES is used with 9 bands in the longwave region, and 6 in the shortwave. The spectral characteristics follow those used in the 60 level HadGEM2 model (Collins et al., 2008). Scattering is considered, and clouds are represented using the maximum random overlap approximation. The Gaussian integration method, with six intervals, is used to integrate the shortwave irradiance over daylight hours for the 15^{th} day of each month, which is used as a monthly-mean.

Zonal-mean monthly-mean climatological fields of pressure, temperature, specific humidity, ozone, cloud cover, and cloud liquid and ice water contents have been produced from ERA-Interim, and used as input to ES. Well mixed concentrations of methane (704 ppbv), carbon dioxide (320 ppmv), nitrous oxide (260 ppbv), and oxygen (21.3%) are also specified. ES is used with a horizontal resolution of 1°, with 37 pressure levels up to 1 hPa. The pressure levels were chosen to be the same as the levels in the ERA-Interim gridded data (see Section 2.5).

The tropopause pressure used in ES is based on the ERA-Interim blended tropopause pressure (Chapter 4), but must correspond to a fixed pressure level for forcing calculations. For each radiative forcing calculation, ES was run twice, with the tropopause specified as the pressure level below and above the blended tropopause. The radiative forcing is then calculated from the change in flux found by ES, and interpolated to the position of the actual tropopause.

The fixed dynamical heating (FDH) approximation is adopted when calculating the adjusted forcings (Fels et al., 1980; Gray et al., 2009; Maycock et al., 2010). It is assumed that diabatic (Q) and dynamical (D) heating are in balance for each atmospheric column so that:

$$Q(T, H_2O, \phi, z) + D(\phi, z) = 0 \tag{8.3.1}$$

where T is temperature, H_2O is the background water vapour, ϕ is latitude, and z is altitude. Q is found using ES, which also gives D when using the climatological fields. When the water vapour is perturbed (H_2O') it causes a change in diabatic heating so that Equation 8.3.1 is no longer true:

$$Q(T, H_2O', \phi, z) + D(\phi, z) \neq 0 \tag{8.3.2}$$

Equilibrium is restored by adjusting stratospheric temperatures. A new temperature is found by timestepping the radiation code using the unperturbed value of D, so

$$Q(T', H_2O', \phi, z) + D(\phi, z) = 0 \tag{8.3.3}$$

The temperature was assumed to have converged when the temperature change with each timestep was less than 5 μK. All tropospheric variables are held at climatology while FDH is applied in the stratosphere.

There has been some debate about the usefulness of broadband radiation models for quantifying the radiative effects of stratospheric water vapour perturbations (e.g. Forster et al., 2001; Oinas et al., 2001). Myhre et al. (2009) compared a series of broadband models to more accurate line-by-line and narrowband models for a uniform perturbation of 3 to 3.7 ppmv in the stratosphere, and for perturbations due to subsonic and supersonic aircraft emissions. They found that the difference in net instantaneous radiative forcing between models was in excess of a factor of two, and that the results from ES in particular deviated from the results from other models (Figure 8.3). However, the adjusted radiative forcing from ES compared well with the results from a narrowband model (see Figure 8.4). Although more sophisticated codes are available, the large number of calculations that were required for this work meant that a fast code was needed. As adjusted radiative forcing will be considered, the ES code was used.

The impact of a uniform perturbation of 3 to 3.7 ppmv in the stratosphere has also been evaluated in this thesis using the configuration of ES used in the calculation of the radiative forcing due to aviation water vapour emissions. For the 3 to 3.7 ppmv calculation, the most notable difference is the use of the blended tropopause in the calculations. This surface tends to be lower in altitude than other tropopause definitions, so the change in water vapour will also extend to lower altitudes in this work. For January to March 2004 meteorology, the instantaneous radiative forcing due to the uniform perturbation was found to have an average value of 0.59 W m^{-2}. The adjusted value is 0.35 W m^{-2}. This is larger than the 0.49 Wm^{-2} (instantaneous) and 0.26 W m^{-2} (adjusted) found by Myhre et al. (2009). Some difference between the two sets of results is not unexpected. Myhre et al. (2009) use only 8 longwave bands, compared to the 9 used in this thesis, and different input climatologies have been used. Additionally, both the instantaneous and adjusted values presented in this thesis are global values. In Myhre et al. (2009) the adjusted radiative forcing is a global number, while the instantaneous value is for a single profile.

The temperature change at 60°N resulting from the 3.0 to 3.7 ppmv perturbation is shown in Figure 8.5. The maximum temperature change of -1.47 K occurs just above the tropopause at 150 hPa.

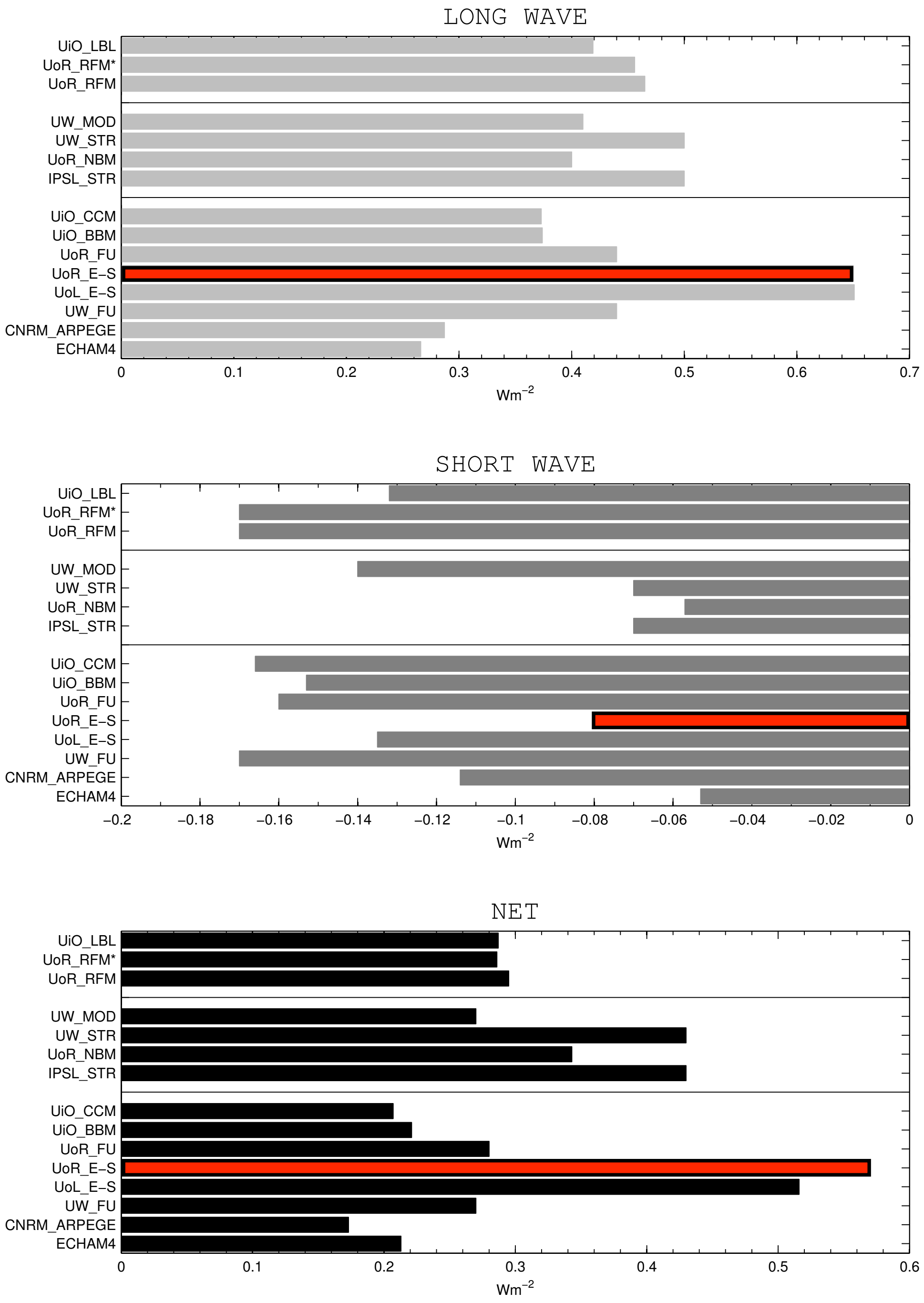
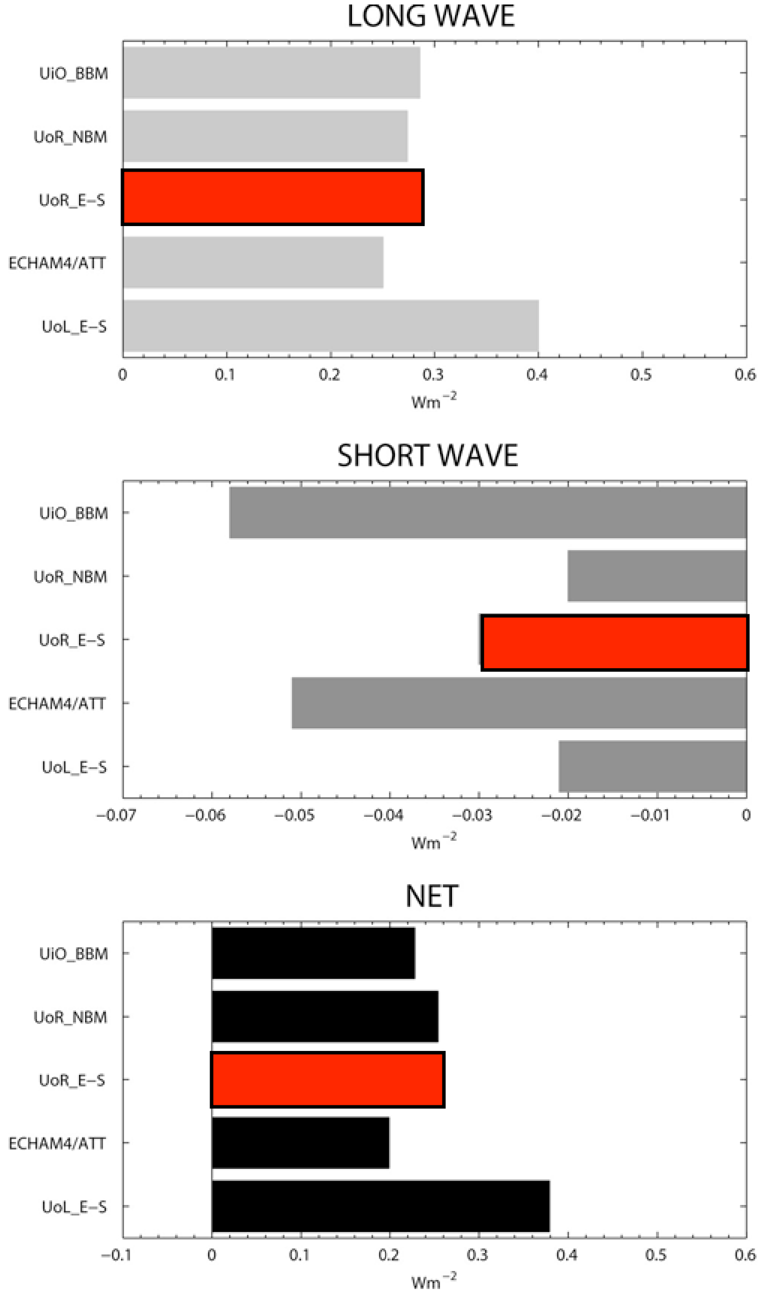


Figure 8.3: *Longwave (top), shortwave (middle) and net (bottom) instantaneous radiative forcing at the tropopause (179 hPa) for a 3 to 3.7 ppmv stratospheric water vapour perturbation from different radiative transfer codes. The calculations were performed for a single profile in mid-latitude summer. The top group are line-by -line models, the middle group are intermediate complexity models, and the bottom group are broadband models. Results from ES, as used here, are highlighted in red. All plots are from Myhre et al. (2009).*


LONG WAVE

SHORT WAVE

NET

Figure 8.4: *Longwave (top), shortwave (middle) and net (bottom) adjusted radiative forcing at the tropopause for a 3 to 3.7 ppmv stratospheric water vapour perturbation from different radiative transfer codes. Results from ES, as used here, are highlighted in red. In contrast to the instantaneous results shown in Figure 8.3, the adjusted values shown in this Figure are global values. Spatial variability in the input climatologies, tropopause, and the temperature adjustments, has also been introduced. All plots are from Myhre et al. (2009).*

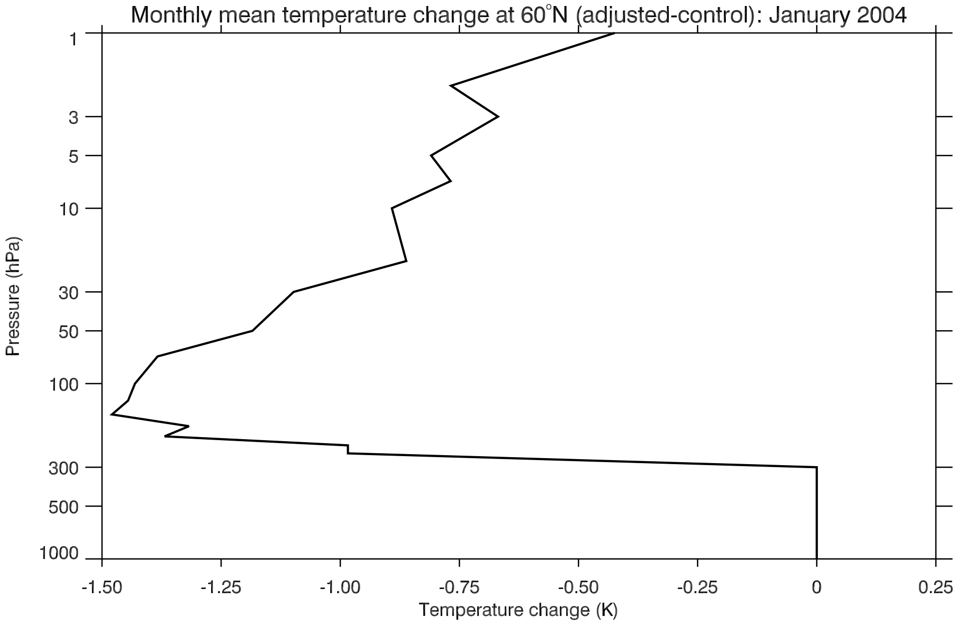


Figure 8.5: *The zonal-mean monthly-mean temperature change at 60°N resulting from a uniform stratospheric water vapour perturbation of 3.0+0.7 ppmv.*

8.4 Radiative forcing due to aviation water vapour emissions

Radiative forcing calculations have been performed for the perturbations introduced in Chapter 7. The sensitivity of the radiative forcing due to aviation water vapour emissions to the background meteorology is investigated using January, February, and March perturbations for four winters: 2004, 1995, 1996, and 1998. In each case trajectories for the winter being investigated are used with emissions from the AEDT 2006 inventory. Radiative forcing calculations are performed using the relevant background climatology and tropopause for that season. As for the perturbation calculations presented in Chapter 7, results for December are not considered as the perturbation typically does not reach equilibrium until the second or third week of the month. The effect of flight routing on radiative forcing is investigated through the comparison of perturbations calculated using the AEDT, QUANTIFY and AERO2k inventories. Summer values of radiative forcing are calculated using the JJAS 2006 perturbations produced using full resolution

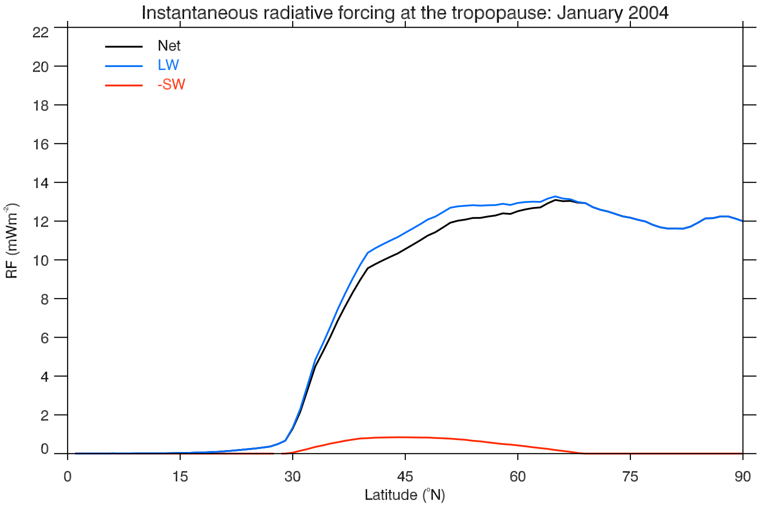


Figure 8.6: *Zonal-mean monthly-mean instantaneous radiative forcing for January 2004 meteorology and AEDT 2006 water vapour emissions. Note that the shortwave component has been multiplied by -1.*

ERA-Interim data, and the seasonal cycle in radiative forcing, and an annual average, is presented from perturbations calculated using low-resolution ERA-Interim trajectories and the AEDT 2006 inventory. In all cases, radiative forcing calculations are performed using zonal-mean monthly-mean climatological fields from ERA-Interim, and the zonal-mean monthly-mean ERA-Interim blended tropopause.

The zonal-mean net instantaneous radiative forcing with latitude for January 2004 meteorology (using January 2006 emissions from AEDT) is shown in Figure 8.6, along with the longwave and shortwave contributions. Shortwave radiative forcing is only ∼5% of the net radiative forcing. Solomon et al. (2010) and Maycock et al. (2010), amongst others, have also demonstrated that the shortwave component of the net radiative forcing due to stratospheric water vapour is small compared to the longwave component (see e.g. Figure 8.2).

The temperature change due to the January 2004 perturbation is shown in Figure 8.7. As in the 3.0 to 3.7 ppmv case, the maximum temperature change is found just above the tropopause. However, in this case the peak in the profile is much more pronounced due to the localised nature of the forcing. The maximum temperature change is also 30 times smaller (-0.048 K).

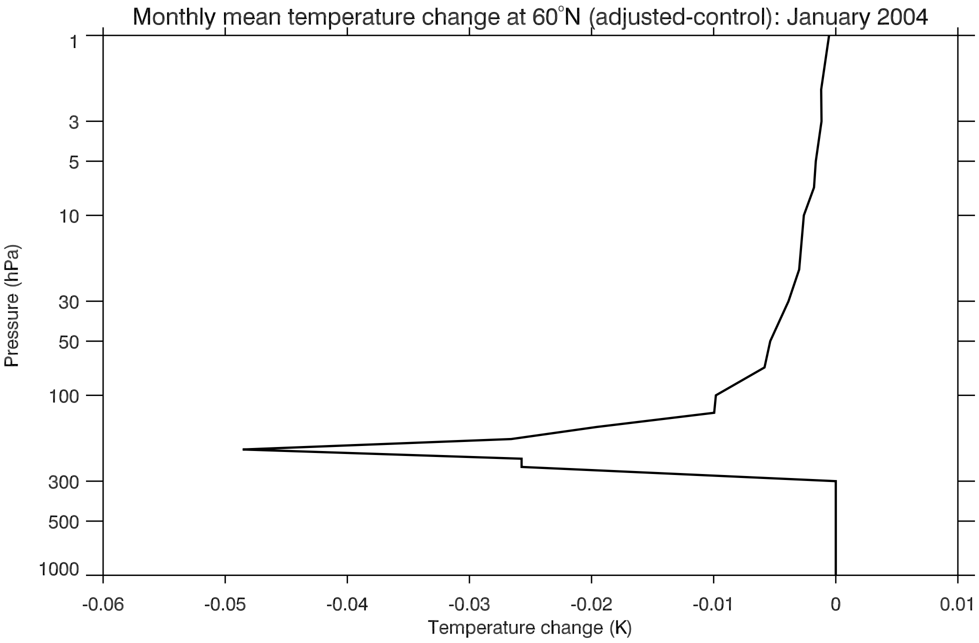


Figure 8.7: *The zonal-mean monthly-mean temperature change at 60°N resulting from the January 2004 aviation water vapour perturbation.*

The interpolated instantaneous and adjusted radiative forcing due to aviation water vapour emissions for January, February, and March 2004, 1995, 1996, and 1998 are shown in Figure 8.8. The forcings are comparable equatorward of ∼45°N, but they diverge poleward of this latitude. In 1995 and 1996 the distribution of forcings is very similar, although the magnitude is larger in January and February of 1996. The largest spread in forcings at any given latitude, of up to a third of the maximum forcing, is found in February. The amplitude of this spread is reduced, along with the magnitude of the forcing itself, in the adjusted case, although the proportions remain the same.

Both the instantaneous and adjusted radiative forcings tend to zero near 20°N in all years, as emissions equatorward of this point are in the troposphere, and there is no significant transport from the high latitudes in the stratosphere. Instantaneous forcings have maxima of up to 18 mW m^{-2} at a given latitude, while adjusted forcings have maxima of up to 10 mW m^{-2}. For January 2004, the global-mean monthly-mean instantaneous radiative forcing due to aviation water vapour emissions is 2.59 mW m^{-2}. The adjusted value is 1.38 mW m^{-2}. The forcing is smaller in February (1.35 mW m^{-2} adjusted), and smaller again in March (1.17 mW m^{-2} adjusted).The global-mean monthly-mean

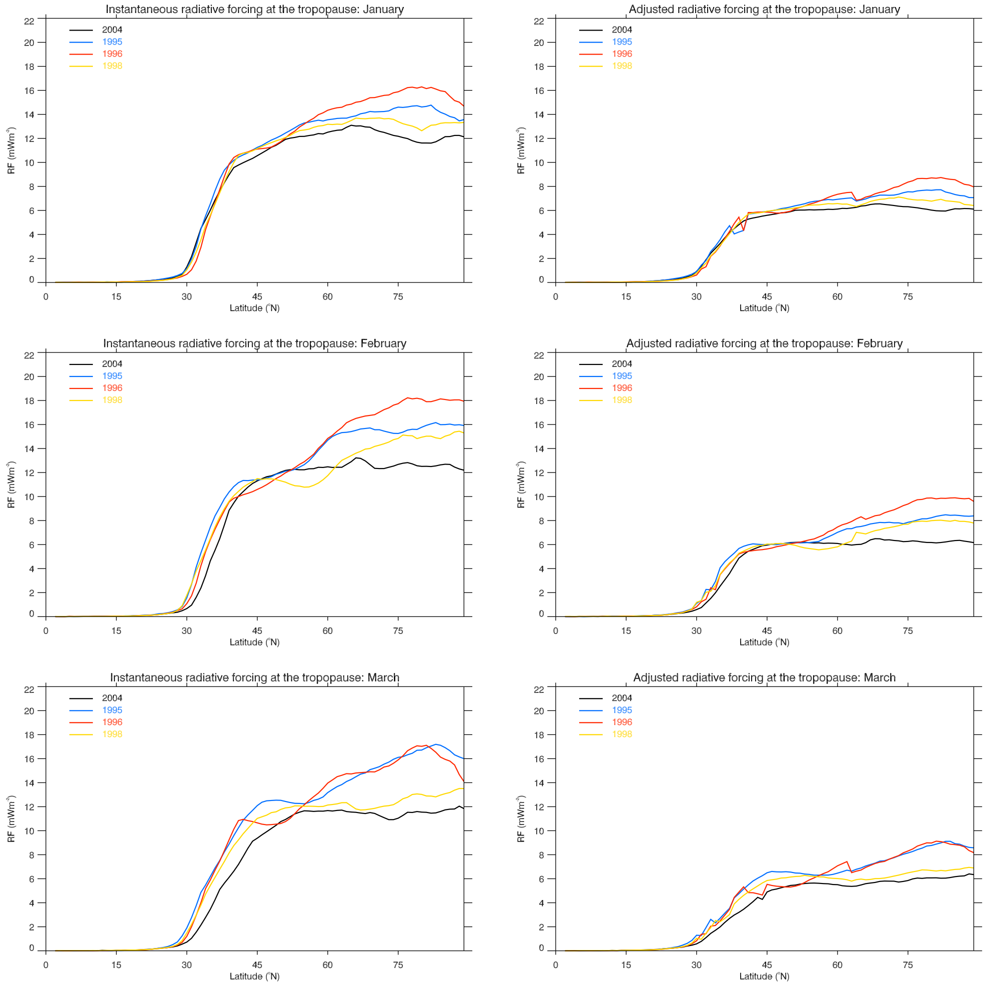


Figure 8.8: *Instantaneous (left) and adjusted (right) radiative forcing for January (top), February (middle) and March (bottom) 2004,1995, 1996, and 1998.*

	Instantaneous (mW m^{-2})			Adjusted (mW m^{-2})		
	Jan	Feb	Mar	Jan	Feb	Mar
JFM 03/04	2.59	2.55	2.22	1.38	1.35	1.17
JFM 94/95	2.82	2.97	2.88	1.53	1.57	1.54
JFM 95/96	2.82	2.94	2.77	1.53	1.59	1.48
JFM 97/98	2.68	2.70	2.53	1.43	1.46	1.36
AEDT	2.59	2.55	2.22	1.38	1.35	1.17
QUANTIFY	1.60	1.59	1.34	0.86	0.87	0.72
AERO2k	1.25	1.23	1.06	0.68	0.67	0.58
AEDT	2.59	2.55	2.22	1.38	1.35	1.17
QUANTIFY (norm)	1.95	1.89	1.67	1.05	1.03	0.90
AERO2k (norm)	1.48	1.49	1.58	0.81	0.81	0.86
ERA-Interim	2.59	2.55	2.22	1.38	1.35	1.17
ERA-Interim (low res)	2.60	2.56	2.24	1.38	1.36	1.17
ERA-Operational	2.30	2.30	2.08	1.20	1.22	1.09
	Jul	Aug	Sep	Jul	Aug	Sep
JJAS 2006	0.56	0.70	0.87	0.30	0.35	0.48

Table 8.1: *Monthly-mean global-mean instantaneous and adjusted radiative forcings (mW m^{-2}) at the blended tropopause due to aviation water vapour emissions. All calculations use the AEDT 2006 inventory, except for the AERO2k and QUANTIFY calculations.*

instantaneous and adjusted radiative forcings at the blended tropopause for JFM from all four winters are given in Table 8.1.

Comparison of Table 8.1 and Table 7.1 shows that there is no simple link between the total mass of emissions in the stratosphere or the maximum zonal-mean perturbation: neither of these values has a maximum in January 2004, relative to February and March 2004.

It can be seen in Table 8.1 that the largest range of radiative forcing values across the four winters considered occurs in March, with an absolute spread of 0.66 mW m^{-2} between global mean instantaneous radiative forcing values for the four years considered. This is 23% of the maximum radiative forcing. A similar percentage spread is found in the adjusted radiative forcing. The smallest range of values is found in January, with an range of 0.23 mW m^{-2} between global-mean monthly-mean instantaneous radiative forcings (8% of the maximum January-mean forcing). In January, February, and March, the maximum global-mean monthly-mean forcing occurs in 1995, but with only smaller differences between the 1995 and 1996 values. This is to be expected from Figure 8.8.

The adjusted radiative forcing at a given latitude, and in the global-mean, is close to half of the magnitude of the corresponding instantaneous forcing because of the stratospheric cooling due to the change in water vapour. The factor of 2 between the instantaneous and adjusted radiative forcings is consistent with the 3.0 to 3.7 ppmv results from ES presented by Myhre et al. (2009), and the 3.0 to 3.7 ppmv results in this thesis.

Instantaneous and adjusted radiative forcing calculated using perturbations derived from the AEDT, QUANTIFY and AERO2k inventories for January to March 2004 are shown in Figure 8.9. Global-mean values are given in Table 8.1. In all cases the radiative forcing due to AEDT emissions is just over a factor of two larger than that due to AERO2k emissions, and a factor of one and a half larger than that due to QUANTIFY emissions. The inventories are for different years, and have different monthly total emissions, which makes a simple comparison difficult. To allow a comparison to be made the forcings have been normalised using the ratio of monthly total emissions from a given inventory to those from AEDT 2006. This removes any differences resulting from different total emissions, leaving only those due to different emissions distributions. However, even when the forcings are normalised with respect to AEDT emissions in this way, there is still a large spread in the radiative forcing (see Figure 8.10 for a comparison of the instantaneous radiative forcing). The largest reduction in spread following normalisation occurs in March. In the instantaneous case maximum spreads of 7, 6, and 6 mW m^{-2} are found in January, February, and March respectively, compared to 6, 5, and 4 mW m^{-2} after normalisation. This indicates that variation in estimates of radiative forcing using different emissions inventories is due primarily to the difference in the distribution of the emissions within the inventories.

Global-mean monthly-mean values of instantaneous and adjusted radiative forcing due to emissions from different inventories are shown in Table 8.1 for January to March. Radiative forcing calculated from AERO2k emissions is a factor of 2 smaller than that calculated from AEDT emissions. After normalisation, the difference is a factor of $\sim$1.35.

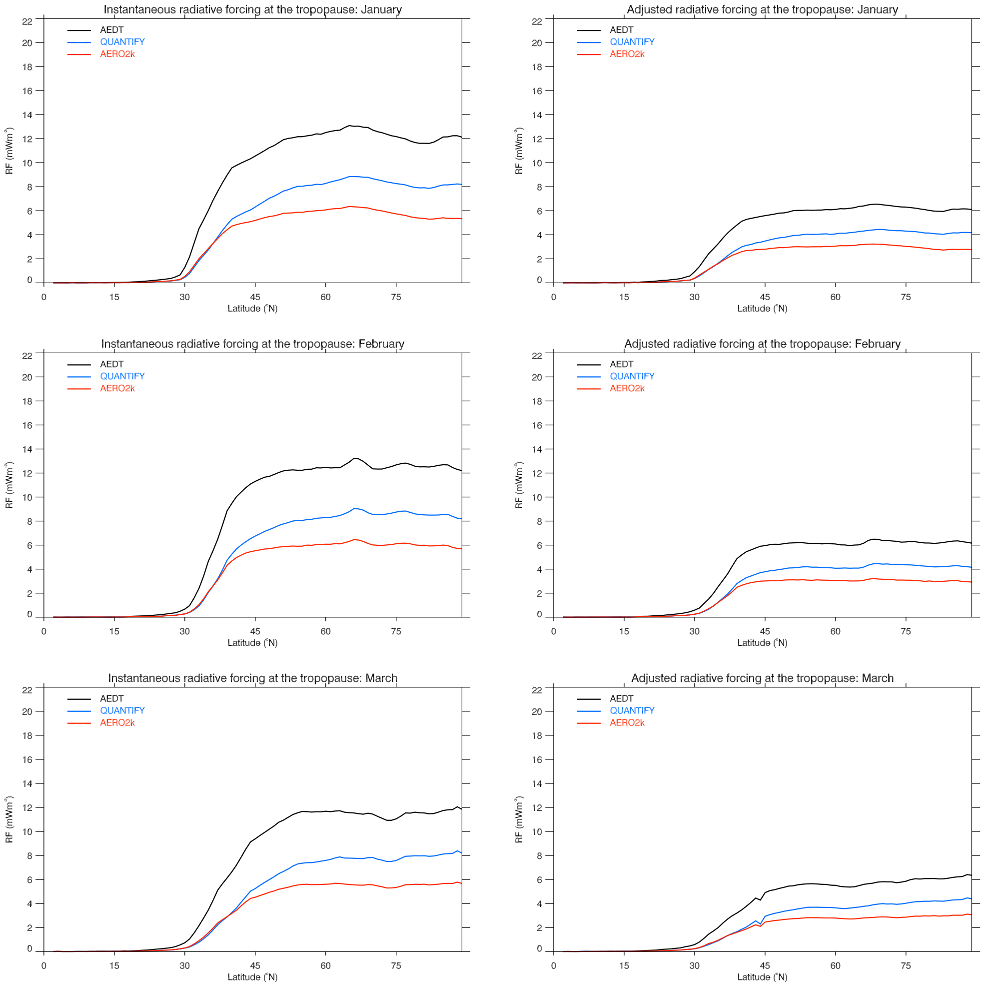


Figure 8.9: *Instantaneous (left) and adjusted (right) radiative forcing for January (top), February (middle) and March (bottom) 2004 meteorology for perturbations calculated using the AEDT (2006), QUANTIFY (2000), and AERO2k (2002) inventories.*

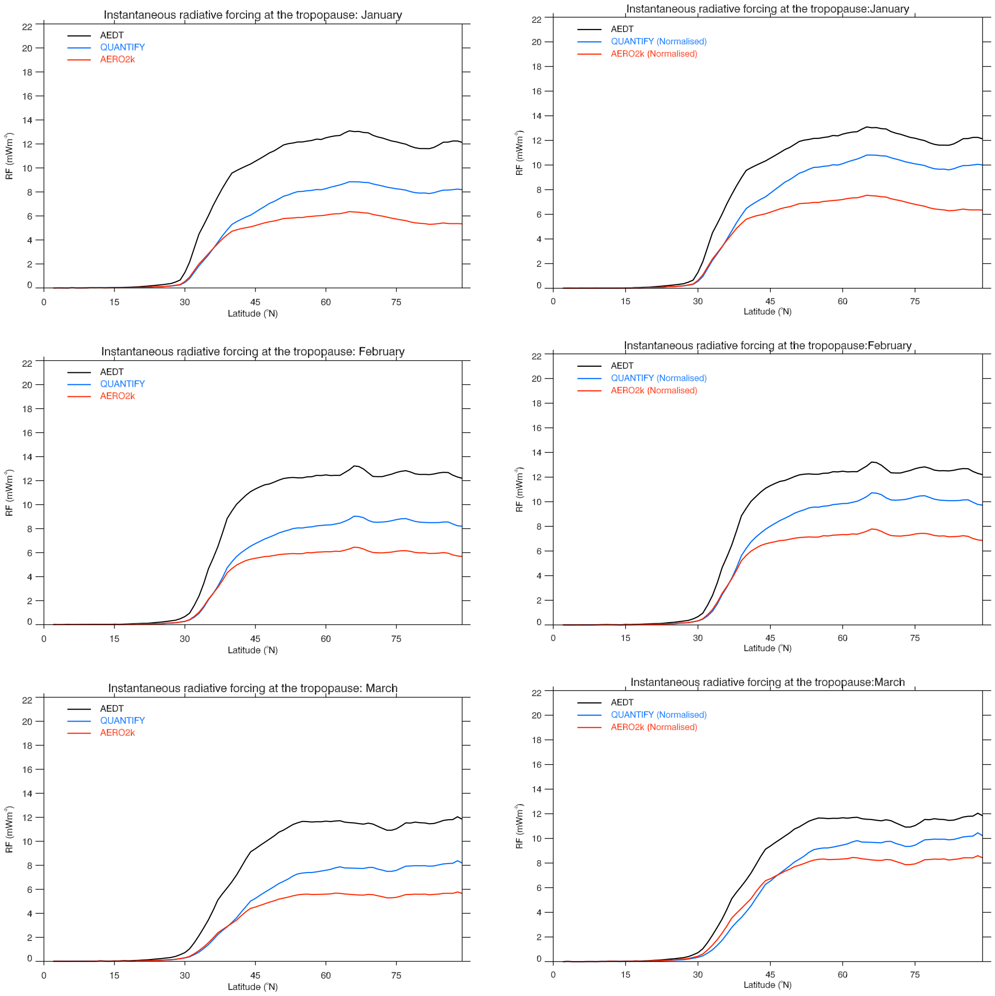


Figure 8.10: *Instantaneous radiative forcing for January (top), February (middle) and March (bottom) 2004 for perturbations calculated using the AEDT, QUANTIFY, and AERO2k inventories. Panels in the left column are as in Figure 8.9. In the panels on the right, the radiative forcing has been normalised with respect to the AEDT inventory.*

For QUANTIFY the radiative forcing is a factor of $\sim$1.60 smaller than that due to AEDT, and a factor of $\sim$1.28 smaller after normalisation. This is to be expected as QUANTIFY emissions have a closer vertical distribution to AEDT compared to AERO2k, and a similar percentage of column total emissions that enter the stratosphere directly (see Chapter 6).

The radiative forcing due to perturbations calculated using full- and low-resolution ERA-Interim data (T255 and T159), and ERA-Operational data, has also been calculated (Figure 8.11, Table 8.1). In all cases an ERA-Interim climatology and blended tropopause is used. The choice of resolution of ERA-Interim data had little influence on the calculated water vapour perturbation (Chapter 7). It follows that it should also have only a small influence on the radiative forcing. This can be seen in Figure 8.11. The use of ERA-Operational data with the UGAMP trajectory model results in over-dispersion in the vertical, leading to shorter lifetimes and smaller perturbations compared to ERA-Interim calculations (Chapters 6 and 7). This is also reflected in the radiative forcing calculations (Figure 8.11, Table 8.1). However, the difference is very small in March (a factor of 1.03 in the global-mean). The sensitivity of radiative forcing to the meteorological data used to calculate the water vapour perturbation is less than that due to interannual variability in the meteorology, and much less than that due to the choice of emissions inventory.

The radiative forcing due to water vapour emissions is smaller in the summer months compared to the winter months, as little as 20% of the winter values (see Figure 8.12, Table 8.1). This is primarily due to the smaller perturbations resulting from the higher tropopause along the busiest flight routes. There is also a small contribution from the increase in shortwave forcing at high latitudes (not shown): the forcing from the shortwave alone is 30% of the net forcing, compared to just 5% in the winter months.

The monthly-mean, global-mean adjusted radiative forcing at the blended tropopause is shown in Figure 8.13. The annual cycle is the same as the annual cycle in total mass of water vapour emissions in the stratosphere (Figure 7.12), with a maximum of 1.40 mW m^{-2} in March, and a minimum of 0.36 mW m^{-2} in August. The annual-mean

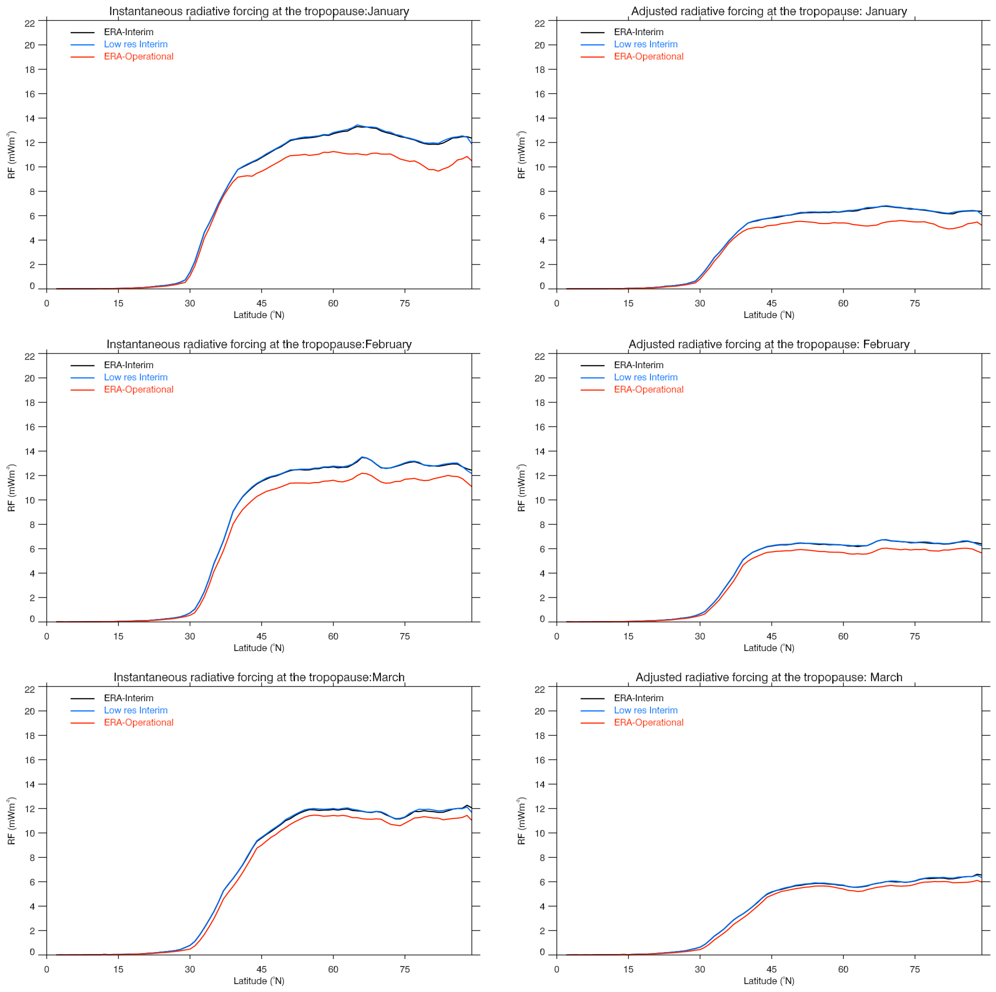


Figure 8.11: *Instantaneous (left) and adjusted (right) radiative forcing for January (top), February (middle) and March (bottom) 2004 for perturbations calculated using full resolution ERA-Interim, low resolution ERA-Interim, and ERA-Operational trajectories. Note that the results from the two sets of ERA-Interim trajectories are almost identical.*

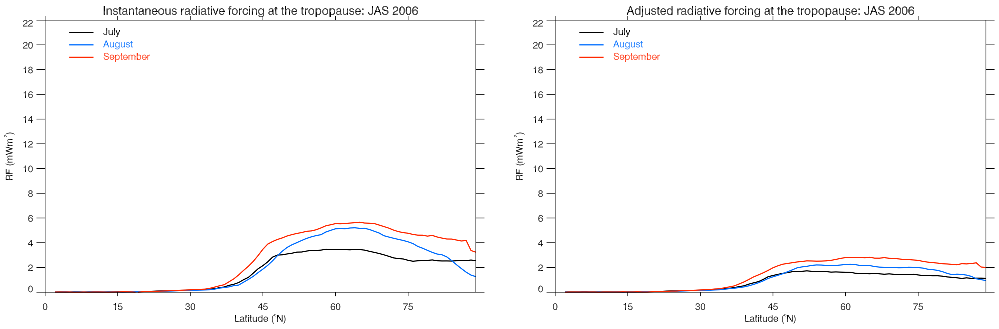


Figure 8.12: *Instantaneous (left) and adjusted (right) radiative forcing for July, August and September 2006.*

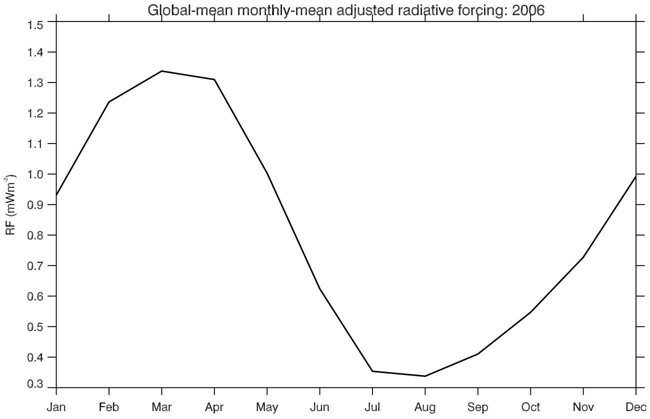


Figure 8.13: *Global-mean monthly-mean adjusted radiative forcing for 2006.*

global-mean adjusted forcing at the blended tropopause is 0.86 mWm^{-2} for 2006.

If the uncertainties associated with interannual variability, the choice of inventory and meteorological data sets, the choice of tropopause definition, and the uncertainties associated with the radiative calculation are assumed to be independent, then, for the cases considered here there is a maximum uncertainty of 56% associated with the annual mean adjusted radiative forcing. The uncertainty from the choice of inventory is ±30% based on the difference between radiative forcing from AEDT and AERO2k emissions. Interannual variability in the meteorology introduces an uncertainty of ±25%, based on the range of radiative forcing estimates for different winters. This represents an upper limit as the uncertainty from interannual variability is based only on winter calculations here, and interannual variability is expected to be smaller in the summer months. Following Myhre et al. (2009), a further ±25% comes from the radiative calculations. An uncertainty of ±10% arises from the choice of meteorological data set, based on the comparison results from ERA-Operational and ERA-Interim trajectories and tropopause height. The choice of tropopause definition itself resulted in an uncertainty of ±30% in stratospheric deposition in the NAFC. It is expected that it will result in a smaller uncertainty in global-mean radiative forcing. However, this has not been tested, so it has been assumed that the choice of tropopause definition also results in a ±30% uncertainty in the radiative forcing, to give the largest uncertainty range. There is also up to a 5% underestimate in global-mean radiative forcing that results from neglecting Southern

Hemisphere emissions. The square root of the sum of the squares of these uncertainties gives an overall uncertainty of ±56%. Hence, the annual-mean global-mean adjusted radiative forcing for 2006 is 0.86±0.48 mW m^{-2}.

8.5 Comparison with previous estimates

The annual-mean global-mean adjusted radiative forcing of 0.86±0.48 mW m^{-2} is small compared to some previous estimates of the radiative forcing due to aviation water vapour emissions (e.g. IPCC (1999), Sausen et al. (2005), and Lee et al. (2009)). However, it is comparable to the estimates of Gauss et al. (2003), Ponater et al. (2006), and Fichter (2009). It is almost double the estimate presented by Marquart et al. (2001). Previous estimates range from 0.8 mW m^{-2} for future fleets (Marquart et al., 2001) to 131 mW m^{-2} (Fortuin et al., 1995) for a present day fleet, although most are in the region of 1 to 2 mW m^{-2} for present-day emissions.

Morris et al. (2003) also presented estimates of the radiative forcing due to subsonic emissions at 54 and 82°N. In winter they found radiative forcings of 2 mW m^{-2} and 4 mW m^{-2} respectively. However, in summer they find negative radiative forcings of -1 mW m^{-2} and -6 mW m^{-2}. It seems very unlikely that such values are correct.

The largest estimate of the radiative forcing due to water vapour emissions was presented by Fortuin et al. (1995). They found an instantaneous radiative forcing of up to 241 mW m^{-2}, and an adjusted radiative forcing of up to 131 mW m^{-2} in the summer. However, this estimate is unlikely to be realistic. Fortuin et al. (1995) also found a very large perturbation of up to 380 ppbv as a result of their assumption that water vapour emissions had a stratospheric lifetime between 0.5 and 2.5 years. This lifetime is at least 6 times larger than the lifetime found in this thesis, and is also much larger than the e-folding lifetimes used in other studies.

Many previous estimates of the climate impact of water vapour emissions are made for future scenarios, including forecast inventories and hypothetical fleets of aircraft. Gauss et al. (2003) present one such study, based on the NASA 2015 inventory, finding an instantaneous radiative forcing of 2.6 mW m^{-2} for their H_2O-K1 kerosene base case. As radiative forcing scales with emissions, this value can be converted to an AEDT 2006 equivalent value by multiplying by a factor of 0.58 (the ratio of AEDT 2006 emissions to NASA 2015 emissions). The annual-mean global-mean instantaneous radiative forcing for 2006 was found in this work to be 1.61 mW m^{-2}. The equivalent radiative forcing from scaling the result from Gauss et al. (2003) is 1.51 mW m^{-2}, which is similar to the value found here.

Marquart et al. (2001) also estimate the radiative forcing due to water vapour emissions in the NASA 2015 inventory. They found an adjusted radiative forcing of 0.8 mW m^{-2}. Scaling this value by a factor of 0.58 gives an AEDT 2006 equivalent adjusted radiative forcing of 0.46 mW m^{-2}, which is almost just over half the value found in this thesis. It was suggested by Ponater et al. (2006) that their value is small because their parameterised water vapour sink was too strong in the upper troposphere/lower stratosphere, as a result of an overestimated background water vapour concentration in their version of ECHAM4.

Using the same inventory, Ponater et al. (2006) found a radiative forcing of 2.8 mW m^{-2} for a fleet of cryoplanes. These aircraft were assumed to follow the same routes as existing aircraft, but emit 2.55 times more water vapour. Again assuming that scalings in the emissions result in a similar scaling of the radiative forcing, this corresponds to a radiative forcing of 0.64 mW m^{-2}, which is smaller than the adjusted radiative forcing of 0.86 mW m^{-2} found in this thesis.

The most recent estimate of the radiative forcing due to aviation water vapour emissions was made by Fichter (2009). Fichter (2009) found an annual-mean global-mean adjusted radiative forcing of 1.0 mW m^{-2} from TRADEOFF 2000 water vapour emissions. She used the ECHAM/ATTILA model, which includes a Lagrangian transport model, to evaluate the radiative forcing due to water vapour emissions in the TRADEOFF inventory. She scales her radiative forcing by 1.5 following the findings of Forster et al. (2001) to account for uncertainties in the radiative transfer calculations, to give a final adjusted radiative forcing of 1.5 mW m^{-2}.

A calculation of the annual-mean global-mean radiative forcing due to water vapour emissions from the TRADEOFF inventory has also been performed here, using 2006 meteorology. It was shown in Section 7.2.5 that the idealised emissions distribution in the QUANTIFY inventory results in a perturbation that is 30% smaller than that found to result from AEDT emissions. Hence, a smaller radiative forcing compared to the AEDT results was found. The global-mean adjusted radiative forcing is 0.53 mW m^{-2}, which is roughly half of the unscaled result presented by Fichter (2009). These large differences in perturbation and radiative forcing suggest that there are some significant differences in the transport and radiative calculations between the two methodologies (see Chapter 7).

In addition to focused evaluations of the impact of water vapour emissions, estimates of the radiative forcing due to water vapour emissions are also presented in the assessments of IPCC (1999), Sausen et al. (2005) and Lee et al. (2009). The forcings are 1.5 mW m^{-2} (1992), 2.0 mW m^{-2} (2000) and 2.8 mW m^{-2} (2005) respectively. After scaling Sausen et al. (2005) found that the 2000 equivalent of the IPCC (1999) forcing was 2.0 mW m^{-2}, and Lee et al. (2009) found a 2000 value of 2.5 mW m^{-2}.

Sausen et al. (2005) present mean values for the year 2000, found from the results of the TRADEOFF project. No emissions inventory was available for 2000 at the time. They used a 1991/92 movements base year, corrected by ICAO statistics, to give 2000 emissions. This gave 207.87 Tg(H_2O) yr^{-1}, compared to 232.68 Tg(H_2O) yr^{-1} in AEDT 2006.

Sausen et al. (2005) scaled their CO_2, O_3 and CH_4 results by 1.15, following IPCC (1999), to account for the underestimate inherent in using an inventory. In IPCC (1999), this scaling is also applied to the water vapour radiative forcing. However, Sausen et al. (2005) do not discuss water vapour, so it is not entirely clear whether this scaling factor was applied. With or without the scaling factor, the 1.5 mW m^{-2} for 2000 emissions is larger than the 0.86 mW m^{-2} for 2006 emissions found in this thesis.

Lee et al. (2009) use scaling arguments to calculate the radiative forcing of the non-CO_2 emissions from aviation. They scale Sausen et al.'s (2005) year 2000 result by the ratio of total IEA fuel to the Sausen et al. fuel, as, even after Sausen et al.'s factor of 1.15 scaling, their fuel still represented an underestimate compared to the fuel sales data of IEA. They then scaled this value, again using IEA fuel sales, to give an estimate of radiative forcing for the year 2005. Lee et al.'s (2009) best estimate was 2.8 mW m^{-2}. This is much larger than the value found in this thesis, and also falls outside the upper end of the uncertainty range.

Lee et al. (2009) also assigned numerical uncertainties to their best estimates using probability distribution functions (PDFs). For water vapour emissions, a lognormal distribution of values was assigned, and a Monte Carlo routine was used to calculate a discrete PDF. The uncertainties presented by Lee et al. (2009) only relate to uncertainties in the model evaluation, and do not include uncertainties in the underlying fuel usage. Their 90% likelihood range had a lower limit of 0.39 mW m^{-2}, and an upper limit of 20.3 mW m^{-2}. The upper end of this range is over an order of magnitude larger than the estimate from this thesis. Although Lee et al. (2009) acknowledge that their uncertainty ranges could be overly large as they assumed that the uncertainties for each radiative forcing component were entirely independent, the upper end of this uncertainty range still seems implausible.

8.6 Summary

The Edwards and Slingo (1996) radiative transfer model has been used to calculate the radiative forcing due to aviation water vapour emissions at the ERA-Interim blended tropopause using zonal-mean monthly-mean stratospheric water vapour perturbations and a background climatology from ERA-Interim.

Monthly-mean global-mean radiative forcings were calculated for 2006 using 6 hourly AEDT emissions, and 6 hourly ERA-Interim, at low spatial resolution. This reveals an annual cycle in radiative forcing, with a maximum of 1.40 mW m^{-2} in March, and a minimum of 0.36 mW m^{-2} in August. The annual-mean global-mean adjusted radiative forcing at the blended tropopause was found to be 0.86 ± 0.48 mW m^{-2}.

The sensitivity of the calculated radiative forcing to interannual variability in the background meteorology, emissions inventory, and meteorological data set and resolution, are consistent with the sensitivities seen in the perturbation concentrations. There is a spread of up to a quarter of the maximum value in the forcings for the four winters considered.

The choice of emissions inventory was found to influence the global-mean radiative forcing by as much as a factor of two. Even when the forcings are normalised with respect to the AEDT result to remove any differences due to different input mass, the difference is still a factor of 1.35. This is due only to the difference in the distribution of the water vapour emissions within the different inventories.

Radiative forcings calculated using perturbations found using ERA-Interim trajectories are larger than those found using ERA-40 trajectories. This is consistent with the excessive vertical dispersion known to occur when ERA-40 data is used with the kinematic version of the UGAMP trajectory model (Chapter 6). However, the difference is small: global-mean adjusted radiative forcings are a factor of 1.03-1.09 larger when ERA-Interim perturbations are used compared to ERA-40 perturbations. The difference in the radiative forcing due to perturbations from ERA-Interim at different resolutions is negligible.

It can be seen from comparison of the sensitivity experiments that it is the choice of emissions inventory that has the largest impact on the final radiative forcing estimate. Both interannual variability and the choice of emissions inventory have more of an influence on the estimate of radiative forcing due to water vapour emissions than the meteorological data set used for the trajectory calculations.

The best estimate of the adjusted radiative forcing due to aviation water vapour emissions presented in this thesis is small compared to those presented in the comprehensive assessments of IPCC (1999), Sausen et al. (2005), and Lee et al. (2009, 2010). However, none of these studies give any detail of their methods, and Lee et al. (2009) relies entirely on scaling arguments to obtain their best estimate. The best estimate presented here is closer to those given by Fichter (2009), Ponater et al. (2006), and particularly Gauss et al. (2003), than those presented in these comprehensive assessments.

The most recent estimate of the radiative forcing due to aviation water vapour emissions, 2.8 mW m^{-2} was presented by Lee et al. (2009, 2010), along with a 90% likelihood range. This best estimate is large compared to the estimate presented in this thesis, even if the upper end of the uncertainty range is considered. The upper limit of Lee et al.'s (2009, 2010) likelihood range was 20.3 mW m^{-2}. This is more than 20 times larger than the best estimate presented in this thesis. As it is difficult to scale the best estimate found here to Lee et al.'s best estimate, this work suggests that it is very unlikely that water vapour emissions from aircraft could cause a radiative forcing as large as 20 mW m^{-2}.

Chapter 9

Conclusions

The aim of this thesis is to quantify the impact of present-day aviation water vapour emissions on climate, and its dependence on the variability of the atmospheric circulation, and on emissions inventory. Emissions inventories have been used to determine the mass and location of emissions when they enter the atmosphere. As only stratospheric emissions are likely to produce a significant perturbation to the background atmospheric humidity, a new blended tropopause definition has been developed, which allows water vapour emissions and their subsequent location to be labelled as either tropospheric or stratospheric. The accumulation of water vapour emissions in the stratosphere was calculated using the UGAMP Lagrangian trajectory model, with the assumption that emissions lost their stratospheric properties after excursions into the troposphere lasting 24 consecutive hours or more. This stratospheric water vapour perturbation, along with the blended tropopause, could then be input to the Edwards and Slingo radiation code to calculate the radiative forcing due to aviation water vapour emissions. The dependence of the radiative forcing on the variability of the atmospheric circulation was tested using meteorology from different winters, and fixed input emissions.

The large differences identified between emissions inventories provided a means to test the sensitivity of the climate impact to flight path. The AEDT, QUANTIFY and AERO2k inventories were studied in detail. The NASA and TRADEOFF inventories were also used to allow direct comparisons with previous studies. There were large differences between the distribution of emissions within the different inventories, particularly in the vertical. These differences have been shown to be crucial to the calculated climate impact of water vapour emissions.

The opportunity to produce a tropopause climatology based on the new tropopause definition was taken from the need to produce a field of tropopause variables four times a day for use in accumulation calculations. Key features were identified in this climatology, and comparisons were made between climatologies derived from ERA-Interim and ERA-40. Linear trends were calculated for the global-mean and zonal-mean tropopause, and at each grid point to reveal the geographical structure of the trends. One of the most striking features of this structure was the large height trends in the sub-tropics, indicative of an expansion of the region with a high tropopause that is characteristic of the tropics. Two new metrics were developed to quantify this expansion.

The main findings of this thesis are summarised in this Chapter, before suggestions are made for further research.

9.1 Emissions inventories

- Significant differences were identified in the construction of the inventories, with QUANTIFY being the most idealised, and AEDT using the most input data. These differences resulted in differences in the total emissions and their distributions, particularly in the vertical.
- Globally, AEDT emissions have a maximum at higher altitudes compared to QUANTIFY and AERO2k. Emissions are distributed more evenly with height in AERO2k than in other inventories. It is thought that such a distribution is not representative of the vertical profile of actual emissions.
- Differences between the emissions distribution within the inventories are smaller over the USA and North Atlantic, where more input data is used, and there is some

overlap in the source of that data.

- It has been shown here that the direct deposition of emissions into the stratosphere, their accumulation there, and their radiative impact are sensitive to the emissions inventory used.

9.2 Tropopause definitions and climatologies

- There is no universally accepted definition of the tropopause. Two common tropopause definitions were considered in this thesis: the WMO thermal tropopause, and a dynamic tropopause at the position where |PV|=2 PVU.
- There are difficulties in applying both the thermal and dynamic tropopause definitions globally.
- In this thesis a flow-dependent blending was developed and used. In this blended definition the smooth transition between the dynamic and thermal tropopause occurs in the region bounded by the 350 K and 370 K isentropic surfaces.
- Climatologies of the tropopause have been produced, using the thermal, dynamic and blended definitions, and ERA-Interim and ERA-40. The height of the blended, thermal, and dynamic tropopauses, and the pressure, potential temperature, specific humidity, and zonal and meridional winds at the blended tropopause, have been archived at $1^\circ \times 1^\circ \times 6$ hour resolution for the whole of ERA-40, and for ERA-Interim up to and including 2007.[1]

9.3 Tropopause trends

- Trends in tropopause variables were evaluated for the global and zonal-mean tropopause, and also for each grid point. Trends were assumed to be linear, and the least squares method was used for fitting.

---

[1] No analysis of specific humidity was performed here.

- Trends in the global-mean annual-mean blended tropopause height were +47 m decade^{-1} (-0.97 hPa decade^{-1}) for ERA-Interim (1989-2007) and +39 m decade^{-1} (-0.8 hPa decade^{-1}) for ERA-40 (1979-2001).

- There is a large difference between the thermal and blended tropopause trends in ERA-40, with the height of the global-mean annual-mean thermal tropopause increasing at a rate of +104 m decade^{-1} (-2.61 hPa decade^{-1}). However, in the later, improved ERA-Interim such a discrepancy between thermal and blended tropopause trends is not apparent.

- Trends were identified in the Antarctic that were consistent with patterns in ozone loss and recovery.

- A broadening of the tropics was identified from trends in tropopause height, and tropopause-level zonal wind. This broadening was quantified using three metrics. The first, based on a critical tropopause altitude, was adapted from Seidel and Randel (2007). The second metric, which uses the ratio of subtropical height trends to subtropical height gradients, and the third metric, based on easterly winds and their tendencies, are both unique to this thesis.

- For 1989-2007, the rate of expansion was found to lie in the range 0.9° decade^{-1} to 2.2° decade^{-1}. When the critical height and wind tendency metrics were used, $\sim$70% of the total expansion was into the SH. However, when height tendencies were used it was found that the amount of expansion into the SH was dependent on season, with 72% into the SH in DJF and 55% in JJA.

9.4 Stratospheric emissions

- The calculated amount of direct deposition of emissions into the stratosphere has been shown to be sensitive to tropopause definition, emissions inventory, and inventory resolution:

  - For February 2004 meteorology, 61% of monthly mean AEDT and QUANTIFY emissions, and 52% of AERO2k emissions, enter the NAFC stratosphere directly.
  - 61% of AEDT emissions enter the NAFC stratosphere directly in February 2004 when the blended tropopause is used. Only 44% of emissions enter the stratosphere directly when the thermal tropopause is used.
  - Positioning emissions at the top or bottom of a 1 km deep grid box results in a difference in direct deposition into the stratosphere of 20%. This is reduced to 2% when 152 m deep grid boxes are used.

- Calculated direct deposition of NAFC emissions into the stratosphere has been found to be strongly correlated with average tropopause height, but not with the NAO index.

- Seasonal cycles in direct deposition have been identified. In 2006 a maximum of 64% of AEDT emissions in the NAFC entered the stratosphere directly in March, decreasing to 38% in June. Globally, a maximum of 31%, and a minimum of 16%, of emissions entered the stratosphere directly in March and August respectively.

- e-folding lifetimes, calculated from 50-day long trajectories, increase rapidly with increasing initial height above the tropopause.

- The e-folding lifetime of NAFC emissions released at the median cruise level of 225 hPa ($\sim$11 km) is 40 days in January 2004, and 25 days in June 2004, when ERA-Operational data is used. For emissions released over the USA, the e-folding lifetime is 29 days.

- Longer e-folding lifetimes were found when ERA-Interim data was used to calculate the trajectories: 54 days for NAFC releases at 225 hPa ($\sim$11 km). This difference is due to the over dispersion in vertical transport that occurs in the UGAMP model when it is driven using ERA-Operational winds.

- It has been shown that the assumption of an e-folding lifetime is not necessarily appropriate for emissions released far above the tropopause. Sets of particles released near to the tropopause exhibited an exponential decay in particle numbers, but as the initial distance above the tropopause was increased the decay became steadily more linear.

- Only the accumulation of emissions in the stratosphere was considered here, as significant accumulation relative to the background humidity is not likely to occur in the troposphere. Emissions were determined to have lost their stratospheric properties after spending 24 consecutive hours in the troposphere. Although this choice was somewhat subjective, it is believed that the choice of a 24 hour criterion gives a good balance: permitting short-term excursions into the troposphere, whilst not giving too much potential for spurious mixing.

9.5 Accumulated emissions

- A routine has been developed to calculate the accumulation of aviation water vapour in the stratosphere from 2-day long trajectories. Mass was assigned to the trajectories from an emissions inventory. Trajectory mass was set to zero if the trajectory stayed below the tropopause for 24 consecutive hours.
- For JFM 2004, the zonal-mean monthly-mean water vapour perturbations had maxima of 60, 64, and 70 ppbv in JFM respectively. In each case the maximum was located at 11 km, near the altitude of maximum emissions. The shape of the perturbation followed the position of the tropopause. The mean stratospheric lifetime of emissions in this case study was found to be 10 days.
- A degree of internanual variability was apparent in the four winters considered (JFM 1995,1996,1998). The largest range in maximum zonal-mean perturbations was found in January, when values ranged from 60 to 66 ppbv.
- There was no clear relationship between the NAO index and the structure and magnitude of the perturbation.

- Apart from the sensitivity to the tropopause definition discussed in the previous Section, the largest sensitivity was to the emissions inventories. The choice of emissions inventory influenced the magnitude of the maximum zonal-mean perturbation by up to a factor of two. After normalising the accumulated mass to the AEDT inventory, the difference was a factor of 1.4.

- The calculated perturbations were found to be sensitive to whether ERA-40 or ERA-Interim trajectories were used, but not to the resolution of the ERA-Interim data. The column total and zonal-mean perturbations were larger when trajectories were calculated from ERA-Interim data, which is consistent with the longer e-folding lifetimes found from ERA-Interim trajectories compared to ERA-Operational.

- In 2006 there was a strong seasonal cycle in the total mass of emissions deposited in the stratosphere, and the maximum zonal-mean perturbation. The peak zonal-mean perturbation has a maximum of 85 ppbv in April, and a minimum of 50 ppbv in July. The latitude at which the maximum zonal-mean perturbation occurs has a bimodal distribution, occurring either in the midlatitudes, or near the pole.

9.6 Radiative forcing due to emissions

- The annual-mean radiative forcing due to AEDT 2006 water vapour emissions was found to be 0.86±0.48 mW m^{-2}. If the uncertainties associated with interannual variability, the choice of inventory and meteorological data sets, the choice of tropopause definition, and the uncertainties associated with the radiative calculation are assumed to be independent, then, for the cases considered in this thesis there is a maximum uncertainty of 56% associated with the annual mean adjusted radiative forcing.

- The global-mean annual-mean instantaneous radiative forcing was found to be 1.61 mW m^{-2}.

- The global-mean annual mean adjusted radiative forcing found here is less than half of the recent estimate of Lee et al. (2009, 2010). Additionally, their best estimates of the radiative forcing due to linear contrails, carbon dioxide, and aviation induced cirrus, are 14, 33, and 38 times larger than the estimate of forcing due to aviation water vapour emissions found here. Their upper limit of the forcing due to aviation induced cirrus is over 100 times larger. (Figure 9.1)

- Interannual variability in the global-mean monthly-mean radiative forcing due to AEDT emissions was investigated using radiative forcings for JFM 1995, 1996, 1998, and 2004. Variability had a maximum in March, out of the three months considered, with forcings from 1.11 to 1.45 mW m^{-2}. In January and February radiative forcing ranged from 1.32 to 1.44 mW m^{-2} and 1.29 to 1.51 mW m^{-2} respectively.
- The radiative forcing was found to be sensitive to the emissions inventory used to calculate the perturbation. The radiative forcing due to AEDT emissions are a factor of 2 larger than those from AERO2k. After normalisation this factor reduced to just over a third.

9.7 Discussion

The radiative forcing found here is small compared to the best estimates given in the assessments of Lee et al. (2009, 2010) and Sausen et al. (2005). However, it is comparable to the estimates of Gauss et al. (2003), Ponater et al. (2006) and Fichter (2009), all of whom produced more detailed estimates than Sausen et al. (2005) and Lee et al. (2009, 2010).

The calculations of Gauss et al. (2003) yield a global-mean annual-mean instantaneous radiative forcing of 1.51 mW m^{-2} (after scaling to AEDT 2006). This is in good agreement with the 1.61 mW m^{-2} found in this thesis. Ponater et al. (2006) found an adjusted radiative forcing which, when scaled to AEDT 2006, is 0.64 mW m^{-2}. This is smaller than the 0.86 mW m^{-2} found here, but within the range of uncertainty.

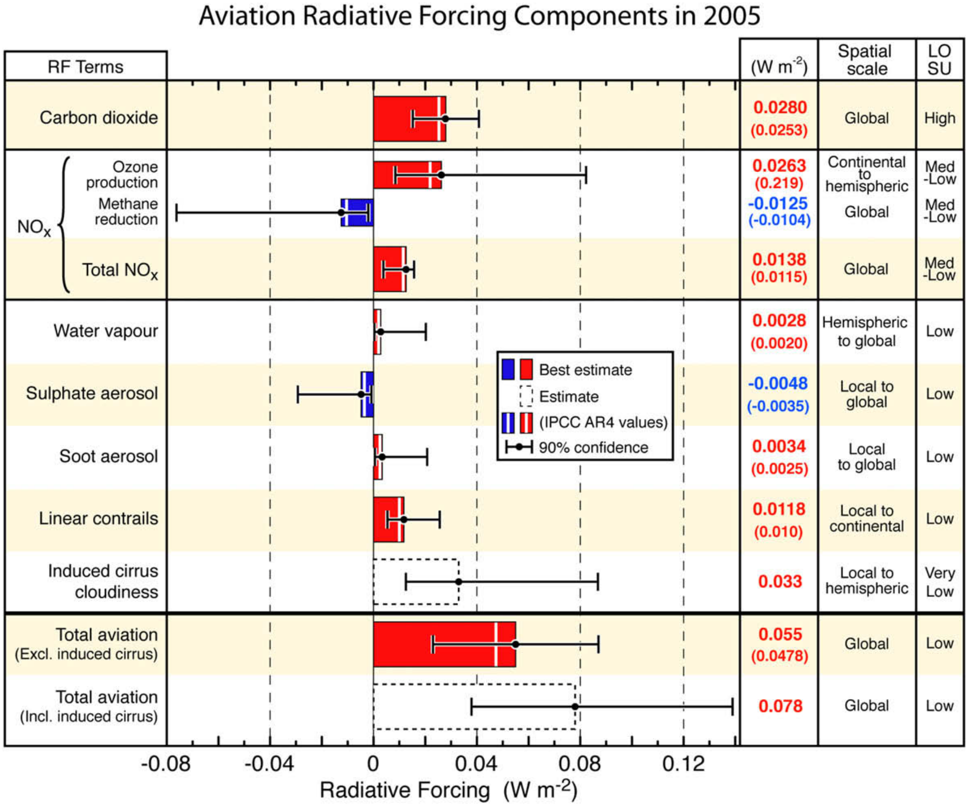


Figure 9.1: *Radiative forcing components from global aviation, from Lee et al. (2009). The coloured bars indicate the best estimate, and the white lines show the IPCC AR4 estimate. Numerical values of these estimates are given on the right, IPCC AR4 estimates are in brackets. The bars indicate the 90% confidence interval. 'LOSU' is the level of scientific understanding.*

The recent estimate of Fichter (2009) is produced using a Lagrangian transport model, but at a lower resolution than the calculations presented in this thesis. She found a radiative forcing of 1.0 mW m^{-2} (1.19 mW m^{-2} if normalised to AEDT 2006), which she scaled to 1.5 mW m^{-2} to account for radiative forcing uncertainties. This is larger than the radiative forcing found in this thesis, but is still within the uncertainty range.

The adjusted radiative forcing presented in this thesis is three times smaller than the best estimate of 2.8 mW m^{-2} presented by Lee et al. (2009). Additionally, the Lee et al. best estimate is two to four times larger than the recent estimates of Fichter (2009), Ponater et al. (2006), and Gauss et al. (2003). As Lee et al. find their estimate from scaling arguments alone, this suggests that their best estimate is too large.

Additionally, Lee et al. (2009) present a 90% likelihood range with their best estimate (Figure 9.1). This has a maximum value of 20 mW m^{-2}, which is far larger than the estimate presented in this thesis. The only previous study where such a large value of radiative forcing was found to result from aviation water vapour emissions is Fortuin et al. (1995), who assumed an unrealistically large 0.5 year lifetime. As it is difficult to reach even the best estimate of Lee et al. (2009) by scaling the result from this thesis, an upper bound to the uncertainty range that is at least an order of magnitude larger than all other recent estimates does not seem plausible.

9.8 Possible topics of future research

The largest source of uncertainty in this work has been the emissions inventories used. The differences between the distribution and mass of emissions within the different inventories is large. Such differences are compounded by the fact that the inventories are for different years. Furthermore, the reliance of inventories on the use of scheduled flight data and idealised routing, in some cases more than others, the assumption of a standard atmosphere with no winds, and the use of load factor assumptions, means that the total fuel burn in the inventories is an underestimate. The underestimate can be accounted for by scaling inventory emissions to IEA fuel sales. However, this makes the implicit assumption that the distribution of emissions within the inventory is realistic. This is not necessarily the case, as there is currently no inventory available that does not rely on cruise altitude assumptions for large parts of the world. It has been shown in this thesis that the climate impact of water vapour emissions is strongly dependent on their vertical distribution. Ideally, future work should be performed using actual emissions data, or at least actual movement data, to avoid such uncertainties.

The interannual variability of stratospheric perturbations due to aviation water vapour emissions was investigated in this thesis. However, the source of this variability could not be attributed to a single cause. It is possible that the release of emissions from specified geographical regions may provide more information about the impact of interannual variability on the perturbation due to aviation water vapour emissions. It would certainly be simpler to determine whether meteorology local to the input region, or zonal-mean meteorology had a greater influence on the accumulation of emissions.

In this thesis, e-folding lifetimes were calculated for different release altitudes. The assumption of an exponential decay in particle numbers was shown to be inappropriate for emissions released far above the tropopause. The stratospheric lifetime of emissions can be calculated directly from the mass of accumulated emissions, without the need for such assumptions. Such a calculation was performed in this work, but only for the stratosphere as a whole, giving an average emission lifetime. If the accumulation code was run for a series of emission releases at specified altitudes, then a vertical profile of emissions lifetimes could be constructed. This would allow a comparison with the e-folding lifetime profile, and reveal how much of an error is associated with the assumption of an e-folding lifetime.

The broadband Edwards and Slingo radiative transfer model was used in this work. There are uncertainties associated with the use of any broadband radiative transfer model, as many approximations need to be made. If time allows in any future studies, it would be instructive to perform at least a portion of the radiative forcing calculations using a line-by-line code.

Of all the aviation induced radiative forcings, that due to water vapour is small and relatively well constrained, especially compared to aviation induced cirrus (see Figure 9.1). Therefore, it is more important that further work on quantifying the climate impact of aviation focuses on other components of the net aviation radiative forcing, and contrail cirrus in particular, instead of water vapour emissions. However, it will still be important to consider water vapour emissions in future scenarios like changes in flight routing and fuel modifications, when water vapour emissions can be the dominant forcing (IPCC, 1999).

It would be possible to reduce the climate impact of aviation water vapour emissions by adopting lower cruise altitudes, as this would reduce the direct deposition of emissions into the stratosphere. Fichter (2009) showed that increasing or decreasing cruise altitudes by just 2000 ft caused 40% increases or decreases in the radiative forcing due to water vapour emissions. However, when considering all of the components of the climate impact of aviation together, the solution is not so simple, as a reduction in cruise altitude will cause increases in the climate impact of other emissions. For example, emissions of carbon dioxide would increase as more fuel would need to be burnt to maintain cruising speeds at their present values, and it is likely that more contrails would be formed as more aircraft would be flying just below the tropopause. As the estimates of the individual aspects of the climate impact of aviation are constrained, the focus of future work can be shifted to identifying operational procedures that minimise the overall climate impact of aviation.

In this thesis, tropopause trends were only evaluated for the ERA-40 and ERA-Interim data sets. It would be useful to compare these results to trends in the blended tropopause using other reanalyses. This would reveal whether the comparatively small blended tropopause trends were a feature unique to the tropopause definition itself, or whether this result was in some way dependent on the reanalysis data. It would also be interesting to examine the spatial structure of tropopause trends in the SH using a different reanalysis (e.g. NCEP), as reanalyses are still quite dependent on their underlying model in this region, which may result in different trends. It would also be interesting to see the results of using the broadening metrics with different reanalysis, or even observational, data.

The mechanisms behind the expansion of the tropics are poorly understood, and the observed rate of expansion exceeds model projections, sometimes significantly. At present, there is no consensus as to the primary cause(s) of tropical broadening, so further study is required to understand the mechanisms.

Bibliography

Ambaum, M. H. P., and Hoskins, B. J., 2002: The NAO Troposphere-Stratosphere Connection, *J. Clim.*, **15**(14), 1969–1978.

Andersson, E., and Thépaut, J. -N., 2008: ECMWF's 4D-Var data assimilation system - the genesis and ten years in operations, *ECMWF Newsletter*, **115**, 8-12.

Andrews, D. G., 2000: *An introduction to atmospheric physics.* Cambridge University Press, 229pp.

Barrie, L. A., Whelpdale, D. M., and Munn, R. E., 1976: Effects of Anthropogenic Emissions on Climate: A review of Selected Topics, *Ambio*, **5**(5/6), 209–212.

Beyersdorf, A., and Anderson, B, 2010: An Overview of the NASA Alternative Aviation Fuel Experiment (AAFEX), *Proceedings of the 2nd International Conference on Transport, Atmosphere and Climate (TAC-2)*, Sausen, R., van Velthoven, P. F. J., Brüning, C., and Blum, A., (Eds.).

Birner, T., 2010: Recent widening of the tropical belt from global tropopause statistics: Sensitivities, *J. Geophys. Res.*, submitted.

Brasseur, G. P., Cox, R. A., Hauglustaine, D., Isaksen, I., Lelieveld, J., Lister, D. H., Sausen, R., Schumann, U., Wahner, A., and Wiesen, P., 1998: European scientific assessment of the atmospheric effects of aircraft emissions, *Atmos. Environ.*, **32**(13), 2329–2418.

Bromwich, D. H., and Fogt, R. L., 2004: Strong Trends in the Skill of the ERA-40 and NCEP–NCAR Reanalyses in the High and Midlatitudes of the Southern Hemisphere, 1958–2001, *J. Climate*, **17**, 4603–4619. doi: 10.1175/3241.1

Charney, J. G., 1947: The dynamics of long waves in a baroclinic westerly current. *J.*

Atmos. Sci., **4**, 136–162.

Chen, J., Carlson, B. E., and Del Genio, A. D., 2002: Evidence for Strengthening of the Tropical General Circulation in the 1990s, *Science*, **295**(5556), 838–841. doi: 10.1126/science.1065835

Clough, S. A., and Iacono, M. J., 1995: Line-by-line calculation of atmospheric fluxes and cooling rates 2. Application to carbon dioxide, ozone, methane, nitrous oxide and the halocarbons, *J. Geophys. Res.*, **100**(D8), 16,519–16,535.

Collins, W. J., Belloiun, N., Doutriaux-Boucher, M, Gedney, N., Hinton, T., Jones, C. D., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior C., Totterdell, I., Woodward, S., Reichler, T., and Kim, J., 2008: Evaluation of the HadGEM2 model. *Hadley Centre Technical Note*, **74**.

Committee on Aviation Environmental Protection (CAEP), 2007: Environmental Goals Assessment, *International Civil Aviation Organization, Working Paper*, **CAEP/7-WP/18**, 36pp.

Danielsen, E. F., 1964: Project Springfield report. Washington D. C., Defense Atomic Support Agency Rept., DASA 1517, 97pp. (DDC).

Danielsen, E. F., 1968: Stratospheric-Tropospheric Exchange Based on Radioactivity, Ozone and Potential Vorticity, *J. Atmos. Sci.*, **25** (3), 502–518.

Danilin, M. Y., Fahey, D. W., Schumann, U., Prather, M. J., Penner, J. E., Ko, M. K. W., Weisenstein, D. K., Jackmann, C. H., Pitari, G., Köhler, I., Sausen, R., Weaver, C. J., Douglass, A. R., Conell, P. S., Kinnison, D. E., Dentener, F. J., Fleming, E. L., Berntsen, T. K., Isaksen, I. S. A., Haywood, J. M., and Kárcher, B., 1998: Aviation fuel tracer simulation: Model intercomparison and implications, *Geophys. Res. Lett.*, **25** (21), 3947–3950.

ECMWF, 2010: The evolution of the ECMWF analysis and forecasting system, `http://www.ecmwf.int/products/data/operational_system/evolution/`

(last accessed 8 November 2010).

Edwards, J. M., and Slingo, A., 1996: Studies with a flexible new radiation code. 1: Choosing a configuration for a large-scale model, *Q. J. R. Meteorol. Soc.*, **122**(531), 689–719.

Eyers, C.J., Norman, P., Middel, J., Plohr, M., Michot, S., Atkinson, K. & Christou, R.A., 2004: AERO2k global aviation emissions inventories for 2002 and 2025, QinetiQ Ltd, `http://www.cate.mmu.ac.uk/aero2k.asp` (last accessed 30 November 2007).

Eyring, V., Cionni, I, Bodeker, G. E., Charlton-Perez, A. J., Kinnison, D. E., Scinocca, J. F., Waugh, D. W., Akiyoshi, H., Bekki, S., Chipperfield, M. P., Dameris, M., Dhomse, S., Frith, S. M., Garny, H., Gettelman, A., Kubin, A., Langematz, U., Mancini, E., Marchand, M., Nakamura, T., Oman, L. D., Pawson, S., Pitari, G., Plummer, D. A., Rozanov, E., Shepherd, T. G., Shibata, K., Tian, W., Braesicke, P., Hardiman, S. C., Lamarque, J. F., Morgenstern, O., Pyle, J. A., Smale, D., and Yamashita, Y., 2010: Multi-model assessment of stratospheric ozone return dates and ozone recovery in CCMVal-2 models, *Atmos. Chem. Phys*, **10**, 9451-9472.

Fels, S., Mahlman, J. D., Schwarzkopf, M. D., and Sinclair, R. W., 1980: Stratospheric sensitivity to perturbations in ozone and carbon dioxide: radiative and dynamical response, *J. Atmos. Sci*, **37**, 2265–2297.

Fichter, C., 2009: Climate Impact of Air Traffic Emissions in Dependency of the Emission Location and Altitude, *Ph.D. Thesis*, Deutsches Zentrum für Luft- und Raumfahrt.

Forster, P. M. D., Freckleton, R. S., and Shine, K. P.: 1997: On aspects of the concept of radiative forcing, *Clim. Dyn.*, **13**, 547–560.

Forster, P. M. de F., and Shine, K. P., 1999: Stratospheric water vapour changes as a possible contributor to observed stratospheric cooling, *Geophys. Res. Lett.*,**26**(21), 3309–3312.

Forster, C., and Wirth, V., 2000: Radiative decay of idealized stratospheric filaments in the troposphere, *J. Geophys. Res.*, **105**(D8), 10,169–10,184.

Forster, P. M. D., Ponater, M., and Zhong, W. Y., 2001: Testing broadband radiation schemes for their ability to calculate the radiative forcing and temperature response to stratospheric water vapour and ozone changes, *Met. Zeitscrift*, **10**, 387–393.

Forster, P. M. de F., and Shine, K. P., 2002: Assessing the climate impact of trends in stratospheric water vapor, *Geophys. Res. Lett.*, **29**(6), 1086, 10.1029/2001GL013909.

Forster, C., Stohl, A., James, P., and Thouret, V., 2003: The residence times of aircraft emissions in the stratosphere using a mean emission inventory and emissions along actual flight tracks. *J. Geophys. Res.*, **108**, D12, 8524, 10.1029/2002JD002515.

Forster, P. M. de F., Shine, K. P., and Stuber, N., 2006: It is premature to include non-CO_2 effects of aviation in emission trading schemes, *Atmos. Environ.*, **40**(6), 1117–1121.

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

Fortuin, J. P. F., van Dorland, R., Wauben, W. M. F., and Kelder, H., 1995: Greenhouse effects of aircraft emissions as calculated by a radiative transfer model. *Ann. Geophysicae*, **13**, 413–418.

Fu, Q., Johanson, C. M., Wallace, J. M., and Reichler, T., 2006: Enhanced Mid-Latitude Tropospheric Warming in Satellite Measurements, *Science*, **312**(5777), p. 1179. doi: 10.1126/science.1125566

Fuglestvedt, J. S., Shine, K. P., Berntsen, T., Cook, J., Lee, D. S., Stenke, A., Skeie, R. B., Velders, G. J. M., and Waitz, I. A., Transport impacts on atmosphere and climate: Metrics, *Atmos. Environ.*, **44**, 4648–4677.

Gauss, M., Isaksen, I. S. A., Wong, S., and Wang, W. -C., 2003: The impact of H_2O emissions from kerosene aircraft and cryoplanes in the atmosphere, *J. Geophys. Res.*, **108**, doi:10.1029/2002JD002623.

Gettelman, A., 1998: The evolution of aircraft emissions in the stratosphere. *Geophys. Res. Lett.*, **25**, 2129–2132.

Gettelman, A., and Baughcum, S.L., 1999: Direct deposition of subsonic emissions into the stratosphere. *J. Geophys. Rev.*, **104**, 8317–8327.

Gettelman, A., Hegglin, M. I., Son, S.-W., Kim, J., Kujiwara, M., Birner, T., Kremser, S., Rex, M., Añel, J. A., Akiyoshi, H., Austin, J., Bekki, S., Braesike, P., Brühl, Butchart, N., Chipperfield, M., Dameris, M., Dhomse, S., Garny, H., Hardiman, S. C., Jöckel, P., Kinnison, D. E., Lamarque, J. F., Mancini, E., Marchand, M., Michou, M., Morgenstern, O., Pawson, S., Pitari, G., Plummer, D., Pyle, J. A., Rozanov, E., Scinocca, J., Shepherd, T. G., Shibata, K., Smale, D., Teyssédre, H., and Tian, W., 2010: Multimodel assessment of the upper troposphere and lower stratosphere: Tropics and global trends, *J. Geophys. Res.*, **115**, D00M08, doi:10.1029/2009JD013638.

Gold, E., 1909: The isothermal layer of the atmosphere and atmospheric radiation. *Proc. Roy. Soc. London Ser. A*, **82**, 43–70.

Gray, L. J., Rumbold, S. T., and Shine, K. P., 2009: Stratospheric Temperature and Radiative Forcing Response to 11-Year Solar Cycle Changes in Irradiance and Ozone, *J. Atmos. Sci.*, **66**, 2402–2417.

Grewe, V., and Stenke, A., 2008: AirClim: an efficient tool for climate evaluation of aircraft technology. *Atmos. Chem. Phys.*, **8**, 4621–4639.

Hagemann, S., Arpe, K, and Bengtsson, L., 2005: Validation of the hydrological cycle of ERA-40, *ERA-40 Project Report Series*, Number 24.

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., Russell, G., Aleinov, I., Bauer, M., Bauer, S., Bell, N., Cairns, B., Canuto, V., Chandler, M., Cheng, Y., Del Genio, A., Faluvegi, G., Fleming, E., Friend, A., Hall, T., Jackman, C., Kelley, M., Kiang, N., Koch, D., Lean, J., Lerner, J., Lo, K., Menon, S., Miller, R., Minnis, P., Novakov, T., Oinas, V., Perlwitz, Ja., Perlwitz, Ju, Rind, D., Romanou, A., Shindell, D., Stone, P., Sun, S., Tausnev, N., Thresher, D., Wielicki, B., Wong, T, Yao, M., and Zhang, S., 2005: Efficacy of climate forcings, *J. Geophys. Res.*, **110**, D18104

Haynes, P. H., Marks, C. J., McIntyre, M. E., Shepherd, T. G., and Shine, K. P., 1991: On the "Downward Control" of Extratopical Diabatic Circulations by Eddy-Induced Mean Zonal Forces, *J. Atmos. Sci.*, **48**(4), 651–678.

Held, I., 1982: On the height of the tropopause and the static stability of the troposphere. *J. Atmos. Sci.*, **39**, 412–417.

Highwood, E. J., and Hoskins, B. J., 1998: The Tropical Tropopause, *Q. J. R. Meteorol. Soc.*, **124**(549), 1579–1604.

Highwood, E. J., Hoskins, B. J., and Berrisford, P., 2000: Properties of the arctic tropopause, *Q. J. R. Meteorol. Soc.*, **126**(565), 1515–1532.

Hoerling, M. P., Schaack, T. K., and Lenzen, A. J., 1991: Global objective tropopause analysis. *Mon. Weather Rev.*, **119**, 1816–1831.

Hoinka, K. P., Reinhardt, M. E., and Metz, W., 1993: North Atlantic Air Traffic Within the Lower Stratosphere: Cruising Times and Corresponding Emissions. *J. Geophys. Res.*, **98(D12)**, 23,113–23,131.

Hoinka, K. P., 1998: Statistics of the Global Tropopause Pressure. *Mon. Wea. Rev.*, **126**, 3303–3325.

Hoinka, K. P., 1999: Temperature, Humidity, and Wind at the Global Tropopause. *Mon. Wea. Rev.*, **127**, 2248–2265.

Holton, J. R., Haynes, P. M., McIntyre, M. E., Douglass, A. R., Rood, R. B., & Pfister, L., 1995:Stratosphere-troposphere exchange. *Rev. Geophys.*, **33**, 403–439.

Hoskins, B. J., McIntyre, M. E., and Robertson, W. A., 1985: On the use and significance of isentropic potential vorticity maps. *Q. J. R. Meteorol. Soc.* , **111** (470), 877–946.

Hoskins, B. J., and Berrisford, P., 1988: A potential vorticity perspective on the storm of 1516 October 1987, *Weather*, **43**, 122–129.

Hoskins, B. J., 1991: Towards a PV-θ view of the general circulation, *Tellus A*, **43**(4), 27–35.

Hu, Y., and Fu, Q., 2007: Observed poleward expansion of the Hadley circulation since 1979, *Atmos. Chem. Phys.*, **7**(19), 5229–5236.

Hudson, R. D., Andrade, M. F., Follette, M. B., and Frolov, A. D., 2006: The total ozone field separated into meteorological regimes. Part II: Northern Hemisphere mid-latitude total ozone trends, *Atmos. Chem. Phys. Discuss.*, **6**(4), 6183–6209.

IPCC, 1999: *Aviation and the global atmosphere.* A special report of IPCC Working Groups I and III in collaboration with the scientific assessment panel to the Montreal Protocol on substances that deplete the ozone layer. Penner, J. E., Lister, D. H., Griggs, D. J., Dokken, D. J., and McFarland, M. (Eds). Intergovernmental Panel on Climate Change, Cambridge University Press, UK.

James, P., Stohl, A., Forster, C., Eckhardt, S., Seibert, P., and Frank, A., 2003: A 15-year climatology of stratosphere-troposphere exchange with a Lagrangian particle dispersion model: 2. Mean climate and seasonal variability. *J. Geophys. Res.*, **108(D12)**, 8522, doi:10.1029/2002JD002639.

Johanson, C. M., and Q. Fu (2009), Hadley Cell Widening: Model Simulations versus

Observations, *J. Clim.*, **22**, 2713–2725, doi:10.1175/2008JCLI2620.1.

Kentarchos, A. S., Roelofs, G. J., and Lelieveld, J., 2000: Simulation of Extratropical Synoptic-Scale Stratosphere-Troposphere Exchange Using a Coupled Chemistry GCM: Sensitivity to Horizontal Resolution, *J. Atmos. Sci*, **57**, 2824–2838.

Kim, B., Fleming, G., Balasubramanian, S., Malwitz, A., Lee, J., Ruggiero, J., Waitz, I., Klima, K., Stouffer, V., Long, D., Kostiuk, P., Locke, M., Holsclaw, C., Morales, A., McQueen, E., and Gillette, W., 2005: SAGE, System for assessing Aviation's Global Emissions, Version 1.5. Technical Manual. FAA-EE-2005-01.

Kim, B., Fleming, G., Balasubramanian, S., Malwitz, A., Lee, J., Waitz, I., Klima, K., Locke, M., Holsclaw, C., Morales, A., McQueen, E., and Gillette, W., 2005a: SAGE, System for assessing Aviation's Global Emissions, Version 1.5. Global aviation Emissions Inventories for 2000 through 2004. FAA-EE-2005-02.

Lamarque, J.-F., and Hess, P. G., 1994: Cross-tropopause mass exchange and potential vorticity budget in a simulated tropopause folding, *J. Atmos. Sci.*, **51**, 2246-2269.

Land, C., Feichter, J., and Sausen, R., 2002: Impact of vertical resolution on the transport of passive tracers in the ECHAM4 model, *Tellus*, **54B**, 344–360.

Lee, D. S., Owen, B., Graham, A., Fichter, C., Lim, L. L., Dimitriu, D., 2005: Allocation of international aviation emissions from scheduled air traffic - Present day and historical (report 2 of 3). CATE-2205-3(C)-2.

Lee, D. S., Fahey, D. W., Forster, P. M., Newtone, P. J., Wit, R. C. N., Lim, L. L., Owen, B., and Sausen, R., 2009: Aviation and global climate change in the 21st century, *Atmos. Environ.*, **43**(22-23), 3520–3537.

Lee, D. S., Pitari, G., Grewe, V., Gierens, K., Penner, J. E., Petzold, A., Prather, M. J., Schumann, U, Bai, A., Berntsen, T., Iachetti, D., Lim, L. L., and Sausen, R., 2010: Transport impacts on atmosphere and climate: Aviation, *Atmos. Environ.*, **44**, 4678–4734.

Liu, Y. 2009: Lagrangian studies of troposphere-to-stratosphere transport, *Ph.D. Thesis*, University of Cambridge.

Lorenz, D. J., and DeWeaver, E. T., 2007: The Tropopause Height and the Zonal Wind Response to Global Warming in the IPCC Scenario Integrations, *J. Geophys. Res.*, **112**, D10119, doi:10.1029/2006JD008087.

Lu, J., Vecchi, G. A., and Reichler, T., 2007: Expansion of the Hadley cell under global warming, *Geophys. Res. Lett.*, **34**, L06805, doi: 10.1029/2006GL028443.

Manabe, S., and Strickler, R. F., 1964: Thermal equilibrium of the atmosphere with a convective adjustment. *J. Atmos. Sci.*, **21**, 361–385.

Marquart, S., Sausen, R., Ponater, M, and Grewe, V., 2001: Estimate of the climate impact of cryoplanes, *Aerosp. Sci. Technol.*, (5), 73–84.

Marshall, G. J., 2003: Trends in the Southern Annual Mode from Observations and Reanalyses, *J. Clim.*, **16**, 4134–4143.

Martin, J. E., 2006: Mid-Latitude Atmospheric Dynamics, Wiley, 324 pp.

Maycock, A. C., Joshi, M. M., and Shine, K. P., 2010: The temperature response to stratospheric water vapour changes, *Q. J. R. Met. Soc.*, submitted.

Methven, J., 1997: Offline trajectories: Calculation and accuracy, Tech. Report 44, U.K. Univ. Global Atmos. Modelling Programme, Dept. of Meteorol., Univ. of Reading, Reading, U.K.

Methven, J., Berrisford, P., and Hoskins, B. J., 1999: A Lagrangian climatology for the North Atlantic, Tech. Report 9 , 99pp., Hadley Centre, Meteorol. Off., Bracknell, U.K.

Methven, J., Arnold, S.R., O'Connor, F.M., Barjat, H., Dewey, K., Kent, J. and Brough, N., 2003: Estimating photochemically produced ozone throughout a domain using flight data and a Lagrangian model, *J. Geophys. Res.*, **108**(D9), 4271, doi:10.1029/2002JD0029

Miake-Lye, R. C., Wood, E. C., Timko, M. T., Yu, Z., Herndon, S. C., Lee, B., Santoni, G., Knighton, W. B., 2010: Effects of Alternative Fuels on Hydrocarbon and Particle Emissions from Aircraft Engines, *Proceedings of the 2nd International Conference on Transport, Atmosphere and Climate (TAC-2)*, Sausen, R., van Velthoven, P. F. J., Brüning, C., and Blum, A., (Eds.).

Mitas, C. M., and Clement, A., 2005: Has the Hadley cell been strengthening in recent decades? *Geophys. Res. Lett.*, **32**, L03809. doi:10.1029/2004GL021765

Morris, G. A., Rosenfield, J. E., Schoeberl, M. R., and Jackman, C. H., 2003: Potential impact of subsonic and supersonic aircraft exhaust on water vapor in the lower stratosphere assessed via a trajectory model, *J. Geophys. Res.*, **108** (D3), 4103.

Mortlock, A., and Van Alstyne, R., 1998: Military, charter, unreported domestic traffic and general aviation, 1976, 1984, 1992, and 2015 emission scenarios. NASA/CR-1998-207639, The Boeing Company, Long Beach, CA.

Murcray, W. B., 1970: On the possibility of weather modification by aircraft contrails, *Monthly Weather Review*, **98**(10), 745–748.

Myhre, G., Kvalevåag, M., Rädel, G, Cook, J, Shine, K. P., Clark, H., Karcher, F., Markowicz, K., Kardas, A., Wolkenberg, P., Balkanski, Y., Ponater, M., Forster, P., Rap, A., and de Leon, R. R., 2009: Intercomparison of radiative forcing calculations of stratospheric water vapour and contrails, *Met. Zeitschrift*, **18**(6), 585–596.

Oinas, V., Lacis, A. A., Rind, D., Shindell, D. T., and Hansen, J. E., 2001: Radiative cooling by stratospheric water vapor: big differences in GCM results, *Geophys. Res. Lett.*, **28**(14), 2791–2794.

Petty, G. W., 1994: A fist course in atmospheric radiation. Sundog publishing, Wisconsin. 444pp.

Ponater, M., Brinkop, S., Sausen, R., and Schumann, U., 1996: Simulating the global atmospheric response to aircraft water emissions and contrails: a first approach using a GCM. *Ann. Geophysicae*, **14**, 941–960.

Ponater, M., Pechtl, S., Sausen, R., Schumann, U., and Hüttig, G., 2006: Potential of the cryoplane technology to reduce aircraft climate impact: A state-of-the-art assessment, *Atmos. Environ.*, **40**, 6928–6944.

Plumb, R. A., and Eluszkiewicz, J., 1999: The BrewerDobson Circulation: Dynamics of the Tropical Upwelling, *J. Atmos. Sci.*, **56**, 868–890.

QinetiQ, 2004: *ICAO engine exhaust emissions data bank*. `http://www.caa.co.uk/default.aspx?catid=702`, (last accessed 2 December 2008).

Quan, X-W., Diaz, H. F., and Hoerling, M. P., 2004: Changes in the Tropical Hadley Cell since 1950, *Advances in Global Change Research*, **21**, 85–120.

Rachner, M., 1998: *Die Stoffeigenschaften von Kerosin Jet A-1*, DLR-Mitteilung 98-01, DLR.

Randel, W. J., Wu, F., and Gaffen, D. J., 2000: Interannual variability of the tropical tropopause derived from radiosonde data and NCEP reanalyses, *J. Geophys. Res.*, **105**(D12), 15509–15524. doi:10.1029/2000JD900155

RCEP, 2002: The environmental effects of civil aviation in flight. Royal Commission on Environmental Pollution short report.

Reed, R. J., 1955: A study of a characteristic type of upper-level frontogenesis. *J. Meteorol.*, **183**, 653–684.

Reed, R. J., and Danielsen, E. F., 1959: Fronts in the vicinity of the tropopause, *Arch. Met. Geoph. Biokl. A.*, **11**, 1–17.

Reinking, R. F., 1968: Insolation reduction by contrails, *Weather*, **23**(4), 171–173.

Rind, D., Lonergan, P., and Shah, K., 1996: Climatic effect of water vapor release in the upper troposphere, *J. Geophys. Res.*, **101** (D23), 29,395–29,405.

Rogers, H., Teyssedre, H., Pitari, G., Grewe, V., Van Velthoven, P., and Sundet, J., 2002: Model intercomparison of the transport of aircraft-like emissions from sub- and supersonic aircraft, *Met. Zeitschrift*, **11** (3), 151–159.

Santer, B. D., Wehner, M. F., Wigley, T. M. L., Sausen, R., Meehl, G. A., Taylor, K. E., Ammann, C., Arblaster, J., Washington, W. M., Boyle, J. S., and Brüggemann W., 2003a: Contributions of Anthropogenic and Natural Forcing to Recent Tropopause Height Changes, *Science*, **301**(5632), 479–483. doi: 10.1126/science.1084123

Santer, B. D., Sausen, R., Wigley, T. M. L., Boyle, J. S., AchutaRao, K., Doutriaux, C., Hansen, J. E., Meehl, G. A., Roeckner, E., Ruedy, R., Schmidt, G., and Taylor, K. E., 2003: Behavior of tropopause height and atmospheric temperature in models, reanalyses, and observations: Decadal changes, *J. Geophys. Res.*, **108**(D1), 4002, doi:10.1029/2002JD002258

Santer, B. D., T. M. L. Wigley, A. J. Simmons, P. W. Kållberg, G. A. Kelly, S. M. Uppala, C. Ammann, J. S. Boyle, W. Brüggemann, C. Doutriaux, M. Fiorino, C. Mears, G. A. Meehl, R. Sausen, K. E. Taylor, W. M. Washington, M. F. Wehner, and F. J. Wentz, 2004: Identification of anthropogenic climate change using a second-generation reanalysis, *J. Geophys. Res.*, **109**, D21104, doi: 10.1029/2004JD005075.

Sausen, R., and Santer, B. D., 2003: Use of changes in tropopause height to detect human influences on climate, *Met. Zeitschrift*, **12**(3), 131–136.

Sausen, R., Isaksen, I., Grewe, W., Hauglustaine, D., Lee, D. S., Myhre, G., Köhler, M. O., Pitari, G., Schumann, U., Stordal, F., and Zerefos, C., 2005: Aviation radiative forcing in 2000: An update on IPCC (1999), *Met. Zeitschrift*, **14**(4), 555–561.

Schoeberl, M. R., Jackman, C. H., and Rosenfield, J. E., 1998: A Lagrangian estimate of aircraft effluent lifetime. *J. Geophys. Res.*, **103(D9)**, 10,817–10,825.

Schoeberl, M. R., and Morris, G. A., 2000: A Lagrangian simulation of supersonic and subsonic aircraft exhaust emissions. *J. Geophys. Res.*, **105(D9)**, 11,833–11,839.

Seidel, D. J., and W. J. Randel (2006) Variability and trends in the global tropopause estimated from radiosonde data, *J. Geophys. Res.*, **111**, D21101, doi: 10.1029/2006JD007363.

Seidel, D. J., and W. J. Randel (2007) Recent widening of the tropical belt: Evidence from tropopause observations, *J. Geophys. Res.*, **112**, D20113, doi: 10.1029/2007JD008861.

Seidel, D. J., Q. Fu, W. J. Randel, and T. J. Reichler (2008) Widening of the tropical belt in a changing climate, *Nature geoscience*, **1**, 21-24, doi:10.1038/ngeo.2007.38.

Seinfeld, J. H., and Pandis, S. N., 1998: *Atmospheric chemistry and physics: From air pollution to climate change.* John Wiley & Sons, 1326pp.

Shapiro, M. A., 1978: Further evidence of the mesosclae and turbulent structure of upper jet stream-frontal zonal systems, *Mon. Weather. Rev.*, **106**, 1100-1111.

Shapiro, M. A., 1980: Turbulent mixing within tropopause folds as a mechanism for the exchange of chemical constituents between the troposphere and stratosphere. *J. Atmos. Sci.*, **37**, 994–1004.

Shine, K. P., 2000: Radiative forcing of climate change, *Space Science Reviews*, **94**, 363–373.

Shine, K. P., Cook, J., Highwood, E. J., and Joshi, M. M., 2003: An alternative to radiative forcing for estimating the relative importance of climate change mechanisms, *Geophys. Res. Lett.*, **30**(20), 2047.

Simmons, A. J., Uppala, S., Dee, D., and Kobayashi, S., 2007: ERA-Interim: New ECMWF reanalysis products from 1989 onwards, *ECMWF Newsletter*, **110**, 25–35.

Simpson, G. C., 1928: Further studies in terrestrial radiation. *Mon. Weather Rev.*, **3**, 322–323.

Solomon, S., Qin, D., Manning, M., Marquis, M., Averyt, K., Tignor, M. M. B., Miller Jr., H. L., and Chen, Z., 2007: Climate Change 2007: The Physical Science Basis.

Cambridge University Press, 996 pp.

Solomon, S., Rosenlof, K. H., Portmann, R. W., Daniel, J. S., Davias, S. M., Sanford, T. J., and Plattner, G-K., 2010: Contributions of Stratospheric Water Vapor to Decadal Changes in the Rate of Global Warming, *Science*, **327**(12), 1219–1222.

Son, S-W., Polvani, L. M., Waugh, D. W., Birner, T., Akiyoshi, H., Garcia, R. R., Gettelman, A., Plummer, D. A., and Rozanov, E., 2009: The Impact of Stratospheric Ozone Recovery on Tropopause Height Trends. *J. Climate*, **22**, 429–445. doi: 10.1175/2008JCLI2215.1

Sprenger, M., and Wernli, H., 2003: A northern hemisphere climatology of cross-tropopause exchange for the ERA15 time period (1979–1993). *J. Geophys. Res.*, **108(D12)**, 8521, doi:10.1029/2002JD002636.

Sprenger, M., Wernli, H., and Bourqui, M., 2007: Stratosphere-troposphere exchange and its relation to potential vorticity streamers and cutoffs near the extratropical tropopause. *J. Atmos. Sci.*, **64**, 1587–1602.

Stohl, A., Bonasoni, P., Cristofanelli, P., Collins, W., Feichter, J., Frank, A., Forster, C., Gerasopoulos, E., Gäggeler, H., James, P., Kentarchos, T., Kromp-Kolb, H., Krüger, B., Land, C., Meloen, J., Papayannis, A., Priller, A., Seibert, P., Sprenger, M., Roelofs, G. J., Scheel, H. E., Schnabel, C., Siegmund, P., Tobler, L., Trickl, T., Wernli, H., Wirth, V., Zanis, P., and Zerefos, C., 2003: Stratosphere-troposphere exchange: A review, and what we have learned from STACCATO. *J. Geophys. Res.*, **108(D12)**, 8516, doi:10.1029/2002JD002490.

Stohl, A., Wernli, H., James, P., Bourqui, M., Forster, C., Liniger, M. A., Seibert, P., and Sprenger, M., 2003-STE: A new perspective of stratosphere-troposphere exchange. *Bull. Amer. Met. Soc.*, 1565–1573.

Trenberth, K. E., 1997: The definition of El Niño, *Bull. Amer. Met. Soc.*, **78**, 2771–2777.

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Da Costa Bechtold, V., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavae, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., Mcnally, A. P., Mahfouf, J. -F., Morcrette, J. -J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevie, D., Viterbo, P., and Woollen, J., 2005: The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc., 131*, 2961-3012, doi: 10.1256/qj.04.174

Uppala, S. M., Dee, D., Kobayashi, S., Berrisford, P., and Simmons, A., 2008: Towards a climate data assimilation system: status update of ERA-Interim, *ECMWF Newsletter*, **115**, 12–18.

Van Velthoven, P. F. J., and Kelder, H., 1996: Estimates of stratosphere-troposphere exchange: Sensitivity to model formulation and horizontal resolution, *J. Geophys. Res.*, **101**, 1429–1434.

Wernli, H., and Bourqui, M., 2002: A Lagrangian '1-year climatology' of (deep) cross-tropopause exchange in the extratropical Northern Hemisphere. *J. Geophys. Res.*, **107(D2)**, 4021, doi: 10.1029/2001JD000812.

Wielicki, B. A., Wong, T., Allan, R. P., Slingo, A., Kiehl, J. T., Soden, B. J., Gordon, C. J., Miller, A. J., Yang, S-K., Randall, D. A., Robertson, F., Susskind, J., and Jacobowitz, H., 2002: Evidence for Large Decadal Variability in the Tropical Mean Radiative Energy Budget, *Science*, **295**(5556), 841–844. doi: 10.1126/science.1065837

WMO, 1957: Definition of the tropopause. *WMO Bulletin*, **6**, World Meteorological Organisation, Geneva.

WMO, 1986: Atmospheric ozone, *Tech. Rep.*, **16**, World Meteorological Organisation, Geneva.

Zängl, G., and Hoinka, K. P., 2001: The Tropopause in the Polar Regions, *J. Clim.*, **14**(14), 3117–3139.

Acronyms and Abbreviations

AEDT	Aviation Environment Design Tool
AERO2k	An EC Fifth Framework Programme project, providing an emissions inventory for 2002
AEROCERT	Aircraft Environmental Impacts of Certification Criteria
AMOC	Air traffic flow management MOdelling Capability
ANCAT	Abatement of Nuisances Caused by Air Transport
AR4	IPCC Fourth Assessment Report
ATTILA	Atmospheric Tracer Transport In a LAgrangian model
BADA	Base of Aircraft DAta
CAA	Civil Aviation Authority
CAEP	Committee on Aviation Environmental Protection
CCM	Chemistry Climate Model
CHEM	Chemistry module that can be coupled to the ECHAM model
CTM	Chemistry Transport Model
DLR	Deutsches Zentrum für Luft und Raumfahrt
DOE	Department Of Environment
EC	European Commission
EC39A	ECHAM4.L39(DLR)/CHEM/ATTILA
ECHAM	ECMWF model, HAMburg version
ECHAM4.L39(DLR)	Version 4 of ECHAM with 39 vertical levels, developed at DLR
ECMWF	European Centre for Medium Range Weather Forecasts
ENSO	El Niño Southern Oscillation
ERA	ECMWF ReAnalysis
ES	Edwards and Slingo radiative transfer code
ETMS	Enhanced Traffic Management System
FAA	Federal Aviation Administration
FAST	Future Aviation Scenario Tool
FDH	Fixed Dynamical Heating
GCM	General Circulation Model
GEWEX	Global Energy and Water cycle EXperiment
GOES	Geostationary Operational Environmental Satellite
GOME	Global Ozone Monitoring Experiment
GSOP	Global Synthesis and Observations Panel

HIRS	High resolution Infrared Radiation Sounder
ICAO	International Civil Aviation Organisation
IEA	International Energy Authority
IFR	Instrument Flight Rules
IFS	Integrated Forecast System
INM	Integrated Noise Model
IPCC	Intergovernmental Panel on Climate Change
ISA	International Standard Atmosphere
ISSCP	International Satellite Cloud Climatology Project
JRA	Japanese ReAnalysis
LRC	Long Range Cruise
LTO	Landing and Take-Off
MODIS	Moderate Resolution Imaging Spectroradiometer
MSU	Microwave Sounding Unit
NAFC	North Atlantic Flight Corridor
NAO	North Atlantic Oscillation
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCEP	National Centres for Environmental Prediction
OAG	Official Airline Guide
OECD	Organisation for Economic Co-operation and Development
PCM	Parallel Climate Model
PIANO	Project Interactive ANalysis and Optimization
PNA	Pacific North American pattern
PV	Potential Vorticity
QGPV	Quasi Geostrophic Potential Vorticity
QUANTIFY	An EC Sixth Framework Programme project, providing an emissions inventory for 2000
RPK	Revenue Passenger Kilometres
SAM	Southern Annular Mode
SBUV	Solar Backscatter UltraViolet radiometer
SST	Sea Surface Temperature
STE	Stratosphere to Troposphere Exchange
STT	Stratosphere to Troposphere Transport
TOMS	Total Ozone Mapping Spectrometer
TRADEOFF	An EC Fifth Framework Programme project, providing an emissions inventory for 2000
VFR	Visual Flight Rules