Decadal analysis of stratospheric ozone depletion using data assimilation and Odin/SMR measurements

Kazutoshi Sagi

Department of Earth and Space Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2016
Decadal analysis of stratospheric ozone depletion using data assimilation and Odin/SMR measurements
Kazutoshi Sagi
ISBN: 978-91-7597-382-1

© Kazutoshi Sagi, 2016.

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie 4063
ISSN 0346-718X

Department of Earth and Space Sciences
Global Environmental Measurements and Modelling
Chalmers University of Technology
SE - 412 96 Gothenburg, Sweden
Phone + 46 (0)31 772 1000

Printed by Chalmers Reproservice
Chalmers University of Technology
Gothenburg, Sweden 2016

The text of this work is licensed under a Creative Commons Attribution 3.0 Unported License. All pictures and figures are property of their respective owners.

Cover: Correlation between assimilated NO and estimated ozone change inside the polar vortex at the potential temperature of 900K in the Arctic 2012/13 winter. More details of the picture can be seen in chapter 8.
Decadal analysis of stratospheric ozone depletion using data assimilation and Odin/SMR measurements
Kazutoshi Sagi
Department of Earth and Space Sciences
Chalmers University of Technology

Abstract
Serious ozone depletion has been measured every Antarctic spring since the early 80’s. This ozone depletion is considered to be a result of photo-chemical reactions and catalytic cycles resulting from anthropogenic halogen containing gases facilitated by the formation of Polar Stratospheric Clouds (PSCs). The reactive halogen species are released through heterogeneous reactions on the surface of the particles.

Arctic ozone depletion is, on the other hand, less severe and show larger variability than Antarctic loss because of the unstable and warmer condition. However, the Arctic stratosphere has been becoming colder and the Arctic ozone loss in the 2011 winter was comparable to Antarctic losses. Currently, global ozone is beleived to be recovering as a result of the Montreal Protocol (1987) on the control of ozone depleting substances (ODSs). Ozone depletion is directly/indirectly linked to the climate because the absorption of UV radiation changes the temperature field. It is therefore important to quantify the loss for future climate prediction.

The aim of this thesis is to quantify ozone depletions in several Arctic and Antarctic winters using ozone profiles measured by Odin/SMR and a data assimilation technique which is generally used in numerical weather prediction. The DIAMOND (Dynamical Isentropic Assimilation Model for Odin Data) is used in this thesis. A new vertical transport scheme was implemented into the DIAMOND model to account for the diabatic descent inside the polar vortex during the polar night. The new version of the DIAMOND model was examined for the specific northern winter (2009/2010) when SMILES (Superconducting Submillimeter-Wave Limb-Emission Sounder) observed stratospheric species as well as SMR. A decadal record of ozone depletion has been determined by comparing the assimilated fields to passively transported fields initialized by assimilation of SMR ozone data. Ozone retrieved from the emission line at 544 GHz has been demonstrated for use in ozone depletion studies in the thesis.

Two different chemical mechanisms, the Cl catalytic cycle with PSC formation and NOx related chemistry, can explain losses at different altitudes that occurred in the polar winters. This thesis also propose an alternative methodology for quantifying ozone depletion utilising assimilation information.

Keywords: remote sensing, stratosphere, data assimilation, limb sounding, ozone loss
The thesis is based on the following articles:


Paper in which I have participated but not appended to this thesis:

## Contents

### Chapter 1 – Introduction

1

### Chapter 2 – Stratospheric Ozone

2.1 Stratosphere .................................................. 4

2.2 Ozone chemistry ................................................ 6

- 2.2.1 Cl \(_x\) chemistry ........................................... 7

- 2.2.2 NO\(_x\) chemistry ........................................... 8

2.3 Ozone distribution .......................................... 10

2.4 Polar ozone depletion ...................................... 10

### Chapter 3 – Remote Sensing of Atmospheric Composition

3.1 Principles ....................................................... 13

- 3.1.1 Emission and absorption ................................. 14

- 3.1.2 Scattering ..................................................... 15

- 3.1.3 Radiative transfer equation ............................... 16

- 3.1.4 Retrieval, Inverse method and Optimal Estimation 17

  Optimal Estimation Method .................................. 17

  Error analysis .................................................... 18

3.2 Limb Sounding ............................................... 19

- 3.2.1 Odin/SMR .................................................... 19

- 3.2.2 SMILES .................................................... 21

### Chapter 4 – Transport Model

4.1 Modeling Eulerian tracer transport with low numerical diffusivity 24

  Eulerian transport scheme ................................... 24

  Ideal discretization of the tracer field ....................... 25

  The first order upstream scheme ............................... 25

  Second Order Moments ....................................... 26

### Chapter 5 – Data assimilation

5.1 Variation of Data assimilation technique .................. 29

- 5.1.1 Optimal Interpolation ................................. 29

- 5.1.2 Variational method ....................................... 30

- 5.1.3 Kalman filter ............................................. 31
CHAPTER 6 – ESTIMATION OF OZONE LOSS

6.1 Development of the estimation of ozone depletion based on data assimilation information ........................................ 34

CHAPTER 7 – SUMMARY OF PAPERS

7.1 Summary Paper A ..................................................... 37
7.2 Summary Paper B ..................................................... 38
7.3 Summary Paper C ..................................................... 39

CHAPTER 8 – OUTLOOK

Filling gaps in observations by using DA .......................... 41
Supplement the DIAMOND model with full chemistry ....... 43

BIBLIOGRAPHY

PAPER A .......................... 53
PAPER B .......................... 69
PAPER C .......................... 97
RELATED PAPER A .......... 121
Acknowledgments

Looking back my history of the past 5 years, I have got more grey hairs on my head, more wrinkles and also FEW kilogram of fat around my body (due to hot chocolates in almost every morning). Probably it is good because I may not need to show my id at systembolagets so often. During the great period, many people accompanied and helped me in different ways. Here I would like to express my appreciation to them.

First, I am deeply grateful to Prof. Donal Murtagh who provided helpful comments and suggestions not only research and study but also life in Sweden. In this thesis and appended papers, he fixed my broken English to proper English. However I did not ask him to check my English in the acknowledgements. So you can see how much he supported me and his patience for more than 5 years through the entire thesis comparing the acknowledgements. I would also like to express my gratitude to my friend Joachim Urban, who unfortunately passed away. Without their guidance and persistent help this thesis would not have been possible. I am also deeply grateful to my colleagues in the group, the department and friends who make my Swedish life more colourful. I really enjoyed the nice times during discussion of scientific, nature, computers and so on. I never forget the time with all of you.

I would also like to thank people in Japan. Thanks to Prof. Yasuko Kasai who dragged me into the complicated science. Without her support, I could not manage to get the opportunity to come Sweden. And I thank friends in Japan who I had nice dinners and beers together every time I came back Japan for holidays.

Finally I would like to thank my family for their moral support and warm encouragements from beyond the sea.
Chapter 1

Introduction

The ozone loss that occurs inside of the Antarctic polar vortex in the lower stratosphere in early spring is understood to be related to catalytic destruction by species such as Cl, Br and NO$_2$ and also the formation of Polar Stratospheric Clouds (PSCs) [Solomon, 1999]. This phenomenon has had large scientific and political impact because the ozone protects life on earth by absorbing harmful solar ultra-violet radiation. Currently ozone depleting substances (ODSs) are slowly decreasing in the stratosphere as a result of the Montreal Protocol (1987) on the control of ODSs. Many studies show that global ozone is also recovering since the end of the nineties [e.g., Jones et al., 2009, Tummon et al., 2015]. However the WMO assessment 2014 reported that the speed of the recovery is slower than it reported in the previous assessment.

Ozone loss also occurs in the northern hemisphere. Arctic ozone loss was considered to be less severe than the Antarctic ozone hole due to an unstable and warmer vortex limiting the formation of PSCs. However, in 2011 several groups reported unprecedented ozone depletion over the Arctic polar region approaching that of the Antarctic ozone hole [e.g. Arnone et al., 2012, Manney et al., 2011, Sinnhuber et al., 2011]. We still do not know if the loss in 2011 was a unique case or can be repeated in the future. Ozone depletion and climate change are indirectly linked. Several studies predicted that the stratospheric cooling induced by increasing atmospheric carbon dioxide enhances ozone depletion [Austin et al., 1992, Shindell et al., 1998]. In practice, the Arctic lower stratosphere has been getting colder over the past decades [WMO, 2011]. The linear dependence between the ozone depletion and the volume of air having a temperature below the threshold for polar stratospheric cloud (PSC) formation [Rex et al., 2006] implies that the Arctic stratospheric ozone depletion may become worse if the cooling trend continues.

It is still necessary to keep monitoring stratospheric ozone. Atmospheric ozone is observed from many platforms; ground based observations, ozone sondes, airborne instruments and satellite missions. The techniques employed by those platforms can be categorized into two main groups. In-situ measurements generally have superb precisions at specific measurement stations. Remote sensing products on the other
Introduction

hand allow us to cover a large area although they are often less precise. Quantifying ozone loss has been carried out using several methods based on different assumptions and instruments [e.g. Eichmann et al., 2002, Grooß and Müller, 2003, Rex et al., 2006, Singleton et al., 2007, Tilmes et al., 2004, Tsvetkova et al., 2007]. The methods, summarized in WMO [2007], are the ozone/tracer correlation method [e.g. Tilmes et al., 2004], the match method [e.g. Rex et al., 2006], the vortex-average method [e.g. Eichmann et al., 2002], the passive subtraction method [e.g. Singleton et al., 2007] and the Lagrangian transport calculation method [e.g. Grooß and Müller, 2007]. Indeed each method has different advantages and disadvantages, however, the comparison of different techniques shows reasonable agreement [e.g. Feng et al., 2011, Kuttippurath et al., 2010, Sonkaew et al., 2013, WMO, 2007].

The subject of this thesis is to study Arctic and Antarctic ozone depletion during the past decade. Rösevall [2007] presented a method which can map polar ozone depletion by assimilating satellite observational data into a two dimensional wind driven transport model. The Dynamical Isentropic Assimilation Model for Odin Data (DIAMOND) was developed for the estimation of polar ozone loss in his studies. In this thesis the DIAMOND model is used and optimized for analyzing ozone loss in both Northern and Southern hemispheres. In paper A, a new vertical transport scheme is implemented into DIAMOND to describe the continuous descent of air during the polar night inside the polar vortex. Paper A investigates ozone loss in the 2009/2010 Arctic winter by assimilating SMILES ozone and Odin/SMR ozone observations. Paper B extends the analysis of polar winters in both hemisphere when the SMR ozone version 2.1 data are available (from 2002 to 2013). For the first time, ozone retrieved from the emission line centred at 544 GHz in the SMR observed spectra is used in order to quantify ozone depletion in paper B. This provides better results at lower altitudes. A comparison between SMR 501 GHz ozone ver. 2.1 and SMR 544 GHz ozone ver. 2.0 has been performed in the paper B as well as comparison with previously published results. By using the results of the inter-annual ozone depletion over more than a decade in paper B, paper C accessed the relative year-to-year contributions of halogen-induced and NO$_x$-induced ozone depletions in the Northern hemisphere. This thesis also discusses a new methodology for ozone loss quantification with the re-use of assimilation information in chapter 6 .
Chapter 2

Stratospheric Ozone

The atmosphere is a layer of air trapped by the Earth’s gravity. The atmosphere, however is much thin compared to the radius of the Earth and has a mass of approximately $5 \times 10^{18}$ kg, it is essential to life and exhibits large variety of dynamical and/or meteorological phenomena. The pressure decreases exponentially with height. Almost 99% of the atmospheric mass settles below 30 km. Equation 2.1 gives a relation between the pressure $P$ at given altitude $z$ and surface pressure $P_0$ for an isothermal atmosphere.

$$P(z) = P_0 \cdot \exp\left(-\frac{z}{H}\right)$$ (2.1)

Here $H$ is a scale height of the atmosphere and is generally about 7 – 8 km on the Earth. The atmosphere commonly complies with the ideal gas law

$$PV = N \cdot k_B \cdot T,$$ (2.2)

where $V$ is the volume, $N$ is the number of molecules in that volume, $k_B$ is Boltzmann’s constant ($1.38 \times 10^{23}$ m$^2$kg s$^{-2}$ K$^{-1}$) and the absolute temperature in Kelvin.

The atmosphere can be divided into 4 layers by the vertical structure of temperature. Figure 2.1a shows the typical temperature profile. In the lowest layer called the troposphere, the temperature decreases with height. The thickness of the troposphere varies with latitudes and seasons. Generally, the tropopause which is the limitation of the troposphere is at about 15 km in the tropics and at 7 km in polar regions. The troposphere is characterized by strong turbulence and vertical mixing. The highest abundance of water vapour and atmospheric phenomena such as clouds and precipitation can be seen in this layer.

Above the tropopause the temperature gradient is reversed in the layer named the stratosphere. The stratosphere contains 90% of atmospheric ozone absorbing the ultra-violet (UV) radiation, which explains the positive lapse rate. Compared to the troposphere, the stratosphere is very dry and vertical mixing is hinder.

The next layer is the mesosphere. The temperature decreases with height in this layer due to the lack of ozone and radiative cooling to the space. A minimum temperature of less than 185 K is found in the summer hemisphere at the mesopause which
is placed at approximately 80 km. Above the mesopause the temperature increases to reach maximum values due to the photolysis of O$_2$ and N$_2$ as well as ionization. This last layer is called the **thermosphere**.

### 2.1 Stratosphere

The positive temperature gradient in the stratosphere means that warmer air lies over the cooler air, causing stable conditions. The stability of the atmosphere can be also categorized using the dry adiabatic lapse rate $\Gamma_d$ ($= -dT/dz = 9.8$ K km$^{-1}$).

$$-\frac{dT}{dz} > \Gamma_d \quad \text{unstable}$$

$$-\frac{dT}{dz} = \Gamma_d \quad \text{neutral} \quad \text{(2.3)}$$

$$-\frac{dT}{dz} < \Gamma_d \quad \text{stable}$$

An other useful tool when discussing atmospheric stability is potential temperature $\Theta$, expressed as below,

$$\Theta = T \left( \frac{P_0}{P} \right)^{\frac{R_d}{C_p}}, \quad \text{(2.4)}$$

where $R_d$ is the dry gas constant and $C_p$ is the heat capacity of the gas, respectively. $\Theta$ is described as the temperature that the dry air parcel would have if it moves
2.1. Stratosphere

Figure 2.2: Schematic view of dynamics of coupling between the stratosphere and the troposphere. Deep convection at the equator penetrates through the tropopause (heavy line), causing large-scale ascent of air. Mid-latitudes are the area where horizontal mixing exists. Near the pole the cold air descends. Source: Holton et al. [1995]

to the surface adiabatically. According to the description of potential temperature, \(dT/dz = -\Gamma_d\) gives \(d\Theta/dz = 0\) and when the atmosphere is in stable condition \(d\Theta/dz\) is positive. In the stratosphere potential temperature is always rising hence stable. Vertical motions are consequently small and stratospheric air tends to move on isentropic surfaces. Thus potential temperature is often used as a vertical coordinate in the stratosphere.

The stratosphere is denoted by large-scale ascent at the equator and large-scale descent in the polar regions. Figure 2.2 shows the schematic view of dynamical processes in the stratosphere. Exchange of air passing the tropopause rarely happens except in the Tropics due to deep convection and close to strong frontal systems.

Westerly winds typically occur in the winter hemisphere. The strongest wind is referred to as a jet stream at where the temperature has the largest gradient due to the polar night terminator. This jet stream is called the polar night jet and the area poleward of the jet is known as the polar vortex. The polar vortex isolates cold air from subtropical air.

The circulation in the stratosphere is weaker than in the troposphere and is driven by the breaking of planetary waves induced by topographical features at the surface. This circulation is named the Brewer-Dobson circulation (BDC). Planetary waves or Rosby waves propagate from the troposphere into the stratosphere where they dissipate and provide a westward torque. The main mechanism behind the BDC is the poleward transport produced by the westward torque of the Coriolis force. The distribution of long-lived tracer gases in the stratosphere is strongly influenced by the
Stratospheric Ozone

An other interesting phenomena in the stratosphere is a part of seasonal and semi-seasonal oscillations. The quasi-biennial oscillation (QBO) dominates the oscillation of the stratospheric equatorial zonal wind between easterlies and westerlies with a period of roughly 28 months. The alternating wind regimes are developed in the middle stratosphere and propagate downwards to the tropical tropopause. Currently gravity wave propagation is considered to be the main contributor to the QBO. More details can be found in Holton and Hakim [2012].

2.2 Ozone chemistry

Ozone is arguably the most important species in the stratosphere. Ozone molecules consist of three oxygen atoms and absorb the ultra violet (UV) radiation which is harmful to life. Figure 2.1b shows an ozone profile in number density and volume mixing ratio for the mid-latitudes between 10 to 120 km. The ozone layer lies between 20 to 40 km, while its peak in volume mixing ratio (8-11 ppmv) is located around 35 km. The number density peak of ozone is however situated at a lower altitude of about 15–25 km, since pressure decreases exponentially with increasing height. The absorption of UV light by ozone molecules leads to the negative lapse rate and increasing temperature in the stratosphere. The mean total ozone typically varies in a range of 300-400 DU\(^1\).

Chapman [1930] suggested a theory for ozone formation and destruction in the stratosphere. The following reactions in the theory are collectively called the Chapman mechanism.

\[
O_2 + h\nu (\lambda < 242 \text{ nm}) \rightarrow 2O \quad \text{(J1)}
\]

\[
O + O_2 + M \rightarrow O_3 + M \quad \text{(R1)}
\]

\[
O_3 + h\nu (\lambda = 242 - 336 \text{ nm}) \rightarrow O(^1D) + O_2 \quad \text{(J2)}
\]

\[
O(^1D) + M \rightarrow O + M \quad \text{(R2)}
\]

\[
O + O_3 + M \rightarrow 2O_2 + M \quad \text{(R3)}
\]

Here \(h\nu\) is the energy of a photon and \(M\) symbolizes a third body in the reaction, removing the excess energy due to collision. \(O(^1D)\) denotes atomic oxygen in an excited state. The binding energy of \(O_2\) (498 kJ/mol) is equal to the photon energy at 242 nm. The oxygen molecule is hence photolysed. The atomic oxygen provided from (J1) is in the ground-level triplet state \(O(^3P)\) and highly reactive due to its two unpaired electrons. The \(O(^3P)\) immediately reacts with a oxygen molecule and produces ozone in (R1). (R1) is significant in the area where the concentration of \(O\) is

---

\(^1\)The Dobson unit (DU) is a unit of measuring ozone concentration. One DU refers to the number of molecules of ozone that would be required to provide a layer with 0.01 millimeters thick at a standard atmosphere (273 K and 1013.25 hPa). 1DU = 2.6 \times 10^{20} \text{ molecules O}_3 \text{ m}^{-2}.

---
2.2. Ozone chemistry

high such as the upper stratosphere or low latitudes. Ozone is also photolysed by the UV radiation at wavelengths below 336 nm. Since (R1) and (J2) are fast reactions, (J2) is not a terminal sink of ozone. The excess energy acquired by M is in the form of thermal energy and heats up the stratosphere. (R2) is the actual reaction of the sink in the Chapman mechanism.

Observational studies showed that the Chapman mechanism and atmospheric transport alone do not correctly describe the real ozone distribution especially at mid-latitudes and in the polar stratosphere. Further reactions of catalytic species such as Br, Cl, NO and HO have to be considered to remove ozone from the stratosphere.

\[ X + O_3 \rightarrow O_2 + XO \]  \hspace{1cm} (R4)
\[ O + XO \rightarrow O_2 + X \]  \hspace{1cm} (R5)

Net: \[ O + O_3 \rightarrow 2O_2 \]

Here, X and XO represent a catalyst and an intermediate product, respectively. These reactions are usually faster than the Chapman reactions hence they constrain the ozone distribution. The following sub-sections introduce more details of the Cl and NO chemistry related to this thesis.

2.2.1 Cl\textsubscript{x} chemistry

Chlorofluorocarbons (CFCs) are anthropogenic gases. Concentrations of CFCs in the atmosphere were increasing during the 1970’s with a rate of 2 – 4% \( yr^{-1} \). CFC molecules are inactive in the troposphere and are transported into the stratosphere. The main source of chlorine in the stratosphere is photolysis of organic chlorine species (e.g. CFCs, HCFCs and CCl\textsubscript{4}). An example is the case of CFC-12 (CF\textsubscript{2}Cl\textsubscript{2}):

\[ CF_2Cl_2 + h\nu \rightarrow CF_2Cl + Cl. \] \hspace{1cm} (J3)

The reactions of the free Cl atoms and the chlorine monoxide radicals ClO lead to ozone destruction.

\[ Cl + O_3 \rightarrow ClO + O_2 \] \hspace{1cm} (R6)
\[ ClO + O \rightarrow Cl + O_2 \] \hspace{1cm} (R7)

Net: \[ O + O_3 \rightarrow 2O_2 \]

This cycle stops when Cl and ClO are converted to the reservoir species, HCl and ClNO\textsubscript{3} by reactions with CH\textsubscript{4} and NO\textsubscript{2}.

\[ Cl + CH_4 \rightarrow HCl + CH_3 \] \hspace{1cm} (R8)
\[ ClO + NO_2 + M \rightarrow ClNO_3 + M \] \hspace{1cm} (R9)
HCl has a lifetime of few weeks and is the most abundant inorganic stratospheric chlorine species. It constitutes more than 95% of the chlorine species above 50 km [WMO, 2011]. ClNO$_3$ is an important reservoir at altitudes below 30 km. The lifetime of ClNO$_3$ is on the order of a day. Hypochlorous acid HOCl is another chlorine reservoir.

\[
\begin{align*}
\text{Cl} + \text{O}_3 & \rightarrow \text{ClO} + \text{O}_2 \\
\text{OH} + \text{O}_3 & \rightarrow \text{HO}_2 + \text{O}_2 \\
\text{ClO} + \text{HO}_2 & \rightarrow \text{HOCl} + \text{O}_2 \\
\text{HOCl} + h\nu & \rightarrow \text{OH} + \text{Cl}
\end{align*}
\]

(R6) (R10) (R11) (J4)

\begin{align*}
\text{Net:} \quad 2\text{O}_3 & \rightarrow 3\text{O}_2
\end{align*}

The latter cycle is significant at lower altitudes (below 20 km) [Brasseur et al., 1999].

### 2.2.2 NO$_x$ chemistry

In the late 1960s many countries considered the use of a supersonic aircraft fleet flying in the stratosphere. Atmospheric chemists worked to assess the effects of such a fleet on the ozone layer. Nitric oxide (NO) is formed by oxidation of atmospheric N$_2$ at high temperatures. In the stratosphere NO reacts rapidly with ozone to produce NO$_2$, which then photolyses:

\[
\begin{align*}
\text{NO} + \text{O}_3 & \rightarrow \text{NO}_2 + \text{O}_2 \\
\text{NO}_2 + h\nu & \rightarrow \text{NO} + \text{O} \\
\text{O} + \text{O}_2 + \text{M} & \rightarrow \text{O}_3 + \text{M}
\end{align*}
\]

(R12) (J5) (R1)

This cycle occurs during daytime with a time scale of one minute and does not affect the ozone budget so is called a “null-cycle”. NO and NO$_2$ interconvert into each other. Odd nitrogen (NO$_x$) comprises the sum of nitrogen oxides (NO+NO$_2$+NO$_3$). The main source of odd nitrogen species and related nitrogen compounds in the stratosphere is the oxidation of N$_2$O transported from the troposphere by O($^1D$).

\[
\text{N}_2\text{O} + \text{O}(^1D) \rightarrow 2\text{NO} \quad \text{(R13)}
\]

(R13) accounts for only about 5% of the loss of N$_2$O in the stratosphere. The rest (95%) is conversion to N$_2$ by photolysis and alternative oxidation reactions by O($^1D$) [Jacob, 1999]. Another known but smaller source of stratospheric NO$_x$ is a polar winter descent of NO$_x$ created mainly by energetic particle precipitation (EPP) in the mesosphere and lower thermosphere [Barth, 2003]. NO lifetime is long enough to be transported down to the stratosphere by the meridional circulation without
photochemical destruction in polar night conditions [Brasseur and Solomon, 2005, Pérot et al., 2014]. Hence irregular events such as aurora and solar proton events (SPEs) provide NO\textsubscript{x}, which can influence the ozone concentration.

The nitrogen family has two cycles that destroy the stratospheric ozone. The first cycle involves both ozone and atomic oxygen and is thus dominant in the middle stratosphere [Brasseur and Solomon, 2005]:

\begin{align*}
\text{NO} + \text{O}_3 & \rightarrow \text{NO}_2 + \text{O}_2 \\
\text{NO}_2 + \text{O} & \rightarrow \text{NO} + \text{O}_2
\end{align*}

\begin{equation}
\text{Net : } \text{O} + \text{O}_3 \rightarrow 2\text{O}_2 \tag{R12}
\end{equation}

The other cycle does not require atomic oxygen and is more effective below 30 km:

\begin{align*}
\text{NO} + \text{O}_3 & \rightarrow \text{NO}_2 + \text{O}_2 \tag{R12} \\
\text{NO}_2 + \text{O}_3 & \rightarrow \text{NO}_3 + \text{O}_2 \tag{R15} \\
\text{NO}_3 + h\nu & \rightarrow \text{NO} + \text{O}_2 \tag{J6}
\end{align*}

\begin{equation}
\text{Net : } 2\text{O}_3 \rightarrow 3\text{O}_2 \tag{J6}
\end{equation}

The loss of NO\textsubscript{x} in daytime is the oxidation by OH.

\begin{equation}
\text{NO}_2 + \text{OH} + \text{M} \rightarrow \text{HNO}_3 + \text{M} \tag{R16}
\end{equation}

At nighttime nitrogen trioxide NO\textsubscript{3} reacts with NO\textsubscript{2} and is converted to N\textsubscript{2}O\textsubscript{5}.

\begin{equation}
\text{NO}_3 + \text{NO}_2 + \text{M} \rightarrow \text{N}_2\text{O}_5 + \text{M} \tag{R17}
\end{equation}

This formation of N\textsubscript{2}O\textsubscript{5} can occur only at night because during the daytime NO\textsubscript{3} coverts back to NO by (J6). HNO\textsubscript{3} and N\textsubscript{2}O\textsubscript{5} are non-radical species and have long lifetimes in the stratosphere. They are eventually converted back to NO\textsubscript{x}:

\begin{align*}
\text{HNO}_3 + h\nu & \rightarrow \text{NO}_2 + \text{OH} \tag{J7} \\
\text{HNO}_3 + \text{OH} & \rightarrow \text{NO}_3 + \text{H}_2\text{O} \tag{R18} \\
\text{N}_2\text{O}_5 + h\nu & \rightarrow \text{NO}_3 + \text{NO}_2 \tag{J8}
\end{align*}

Thus they are reservoirs for NO\textsubscript{x}.

Other minor reactions are not mentioned here. These reactions can be seen in Jacob [1999] and Brasseur and Solomon [2005].
2.3 Ozone distribution

Ozone concentrations in the stratosphere is a balance between photochemical production and removal related to the photolysis and the reactions with catalytic species. However, this balance is not local because ozone molecules produced at certain places are transported to other regions before the photochemical destruction. Figure 2.1b and 2.3 show a vertical ozone profile for the northern summer at mid-latitudes and the seasonal variation of zonal mean total ozone in DU, respectively. Stratospheric ozone is mainly produced in the tropics where the solar flux is maximum by (J1) and (R1). The ozone-rich air is then transported towards the poles by the BDC. As noted in 2.1, the BDC is driven by planetary waves from the troposphere. However, more waves propagate into the stratosphere in the northern hemisphere than in the southern hemisphere because of the irregular distribution of land and sea. The increased wave breaking in the surf zone makes the BDC stronger. Thus the amount of ozone transported towards each pole is not same due to the different strengths of the BDC and the maximum ozone is found at lower altitudes at higher latitudes in northern winter [Wallace and Hobbs, 2006, Wayne, 1991].

2.4 Polar ozone depletion

Severe large ozone depletion, well known as the "ozone hole", was discovered in the early 80’s. The ozone hole occurs over the Antarctica in early spring. The mechanism behind the loss is not explained by the gas phase reactions shown in previous sections alone. This is because that under the low solar exposure condition in polar spring the amount of atomic oxygen is small hence (R7) and (R14) are not efficient reactions.
Further studies revealed that exceptionally high ClO in this season and area plays a key role in ozone depletion [Jacob, 1999].

The following reactions give how the ClO-dimer cycle destroys ozone in the polar vortex.

\[
\begin{align*}
\text{ClO} + \text{ClO} + \text{M} & \rightarrow \text{ClOOCl} + \text{M} && \text{(R19)} \\
\text{ClOOCl} + h\nu & \rightarrow \text{ClOO} + \text{Cl} && \text{(J9)} \\
\text{ClOO} + \text{M} & \rightarrow \text{Cl} + \text{O}_2 + \text{M} && \text{(R20)} \\
2 \times (\text{Cl} + \text{O}_3) & \rightarrow 2 \times (\text{ClO} + \text{O}_2) && \text{(R21)}
\end{align*}
\]

Net : \(2\text{O}_3 \rightarrow 3\text{O}_2\)

Another catalytic cycle involving Br radicals is also important in the depletion:

\[
\begin{align*}
\text{ClO} + \text{BrO} & \rightarrow \text{Cl} + \text{Br} + \text{O}_2 && \text{(R22)} \\
( \text{or} \quad \text{ClO} + \text{BrO} & \rightarrow \text{BrCl} + \text{O}_2 && \text{(R23)} \\
\text{and} \quad \text{BrCl} + h\nu & \rightarrow \text{Cl} + \text{Br} ) && \text{(J10)} \\
\text{Cl} + \text{O}_3 & \rightarrow \text{ClO} + \text{O}_2 && \text{(R24)} \\
\text{Br} + \text{O}_3 & \rightarrow \text{BrO} + \text{O}_2 && \text{(R25)}
\end{align*}
\]

Net : \(2\text{O}_3 \rightarrow 3\text{O}_2\).

Ozone cannot be destroyed by these reactions during polar night and other periods of darkness because sunlight is required to complete each cycle.

The high ClO value is sustained by extremely cold temperature inside the vortex. Specifically, heterogeneous\(^2\) reactions on Polar stratospheric clouds (PSCs)\(^3\) produce ClO\(_x\). On PSCs particles, the Cl\(_x\) reservoirs HCl and ClNO\(_3\) are converted to active chlorine:

\[
\begin{align*}
\text{HCl} + \text{ClNO}_3 & \xrightarrow{\text{PSC}} \text{Cl}_2 + \text{HNO}_3 && \text{(R26)} \\
\text{Cl}_2 + h\nu & \rightarrow 2\text{Cl} && \text{(J11)}
\end{align*}
\]

Recent research demonstrates that low temperature are more critical than PSCs presence because (R26) also proceeds in the aqueous H\(_2\)SO\(_4\) aerosols present in the stratosphere at temperatures below 200 K [Jacob, 1999]. PSCs rather play an important role in ClO production.

---

\(^2\)Chemical reaction takes place in different phases (solid and gas, solid and liquid, two immiscible liquids and so on).

\(^3\)PSCs are categorized into three types by different threshold temperatures; nitric acid trihydrate (NAT) (type Ia) and ternary solution of HNO\(_3\),H\(_2\)SO\(_4\) and H\(_2\)O (type Ib) below 195 K and water ice (type II) below 188 K.
role in the ozone hole formation as a remover of HNO$_3$ from the stratosphere by sedimentation. This prevents the removal of ClO by (R9) and (R17).

Generally, loss of ozone is greater in the Antarctic pole winters than in the Arctic as a result of much colder temperatures. Since the southern hemisphere has less geographical variation than the northern hemisphere which means that the planetary waves are less active, the polar vortex is more symmetric and stable, which isolates cold air parcels.
Chapter 3

Remote Sensing of Atmospheric Composition

This chapter introduces the concept of measuring atmospheric composition using remote sensing. In situ measurements directly sample the atmosphere with high accuracy and precision. However their temporal and spatial coverage is poor. Atmospheric remote sensing from space is providing global measurement coverage. A short introduction of basic principles for remote sensing will be given first. This chapter also includes specific examples of the remote sensing technique related to the thesis.

3.1 Principles

Remote sensing applies electro-magnetic (EM) theory and radiative processes. All objects with non-zero temperature emit EM radiation. EM radiation is characterized by a spectrum of energy. Wavelength $\lambda$, frequency $\nu = c/\lambda$ or wave number $\kappa = 1/\lambda$ can be a variable of the spectrum (where $c$ is the speed of light$^1$). Here, to quantify the radiation, we consider an idealized physical body that absorbs entirely the incident electromagnetic radiation. This is called a "black body". The radiation from a black body obeys the Planck’s Law:

$$B_\nu = \frac{2h\nu^3}{c^2(e^{h\nu/k_BT} - 1)}$$  \hspace{1cm} (3.1)

where $k_B$ is Boltzmann’s constant$^2$ and $h$ is Planck’s constant$^3$. It can also be expressed in wavelength,

$$B_\lambda = \frac{2hc}{\lambda^5(e^{hc/\lambda k_BT} - 1)}$$  \hspace{1cm} (3.2)

$^1$\textit{c} = 3 $\times$ 10$^8$ [m s$^{-1}$]

$^2$k$_B$ = 1.38 $\times$ 10$^{-23}$ [J K$^{-1}$]

$^3$h = 6.63 $\times$ 10$^{-34}$ [J s]
The total energy flux $F$ emitted by the black body can be estimated by integration of Equation 3.2 over all wavelengths and a half sphere, giving $F = \sigma_s T^4$. $\sigma_s$ is a Stefan-Boltzmann’s constant\(^4\). This equation means that total energy emitted by a black body per unit surface area and time is simply proportional to temperature to the fourth power.

Local thermodynamic equilibrium (LTE) can be applied in the middle atmosphere (stratosphere and mesosphere). The energy population in gases will be modified by collisions. If the frequency of collision is higher than the life time of quantum states, the energy level populations dominantly depend on the local temperature and are given by Boltzmann’s distribution. This is the concept of LTE. Kirchhoff’s law refers to the relation between radiative emission and absorption in thermodynamic equilibrium.

$$a_\nu = \epsilon_\nu \quad (3.3)$$

Here $a_\nu$ is absorptivity and $\epsilon_\nu$ is emissivity of material, respectively. A corollary of Kirchhoff’s law is that an arbitrary body emits as much as it absorbs. Most objects are not perfect black bodies. The emitted radiation $L_\nu$ for a body with emissivity $\epsilon_\nu$ can be written using $B_\nu$,

$$L_\nu = \epsilon_\nu B_\nu \quad (3.4)$$

### 3.1.1 Emission and absorption

Once EM waves travel through a medium the photons will encounter molecules or atoms. When this happens the interactions include scattering and absorption by the particles or molecules. A molecule may undergo a transition to a higher energy level by absorbing EM radiation and it may drop to a lower level by emitting radiation. The bound change in internal energy, $\Delta E$, is related to the absorbed or emitted frequency by:

$$\Delta E = h \nu \quad (3.5)$$

Thus a molecule can emit/absorb radiant energy only if the wavelength of the radiation corresponds to the difference between two of its energy levels. Rotational transitions only require the lowest energy corresponding to wavelengths in the microwave and far-infrared (IR) regions. Transitions within vibrational states correspond to wavelengths in the far-IR to near-IR regions. Highly energetic waves, near-IR, visible and UV wavelengths refers to transitions between electronic states. Electronic transitions can involve rotational and vibrational transitions.

Discrete energy transitions are affected by line-broadening processes, pressure and Doppler effects. The absorption spectrum of a single line is written using its shape and strength

$$\gamma_a(\nu) = S f(\nu - \nu_0) \quad (3.6)$$

$$S = \int \gamma_a(\nu) \, d\nu \quad (3.7)$$

\(^4\)\(\sigma_s = 5.6698 \times 10^{-8} \text{ [W m}^{-2} \text{K}^{-4}]\)
3.1. Principles

here $\gamma_a(\nu)$ is absorption coefficient, $f(\nu - \nu_0)$ is the line shape function with the line centre wavenumber $\nu_0$, and $S$ is the line strength. The line shape is determined by the broadening mechanisms: 1) the natural line width from the uncertainty principle is negligibly small, 2) pressure broadening is dominating in the lower atmosphere and 3) Doppler broadening is important above the upper stratosphere. In the lower atmosphere, frequent collisions due to high pressure reduce upper state lifetimes, hence broadening the absorption lines. This is so-called pressure broadening. Pressure broadened lines have a Lorentzian line shape:

$$f(\nu - \nu_0) = \frac{\alpha/\pi}{(\nu - \nu_0)^2 + \alpha^2}$$

(3.8)

$$\alpha = \alpha_0 \left( \frac{P}{P_0} \right) \left( \frac{T_0}{T} \right)^n$$

(3.9)

Where $\alpha$ is the pressure broadened absorption line halfwidth at half maximum and $\alpha_0$ is the line width at a reference temperature $T_0$ and pressure $P_0$ (n is determined empirically). The point is that the line width is proportional to pressure. Thus, the lower the pressure with increasing altitude, the narrower the line becomes. On the other hand, molecules in a gas move with a certain velocity. This motion results in a Doppler shift in the absorbing frequency. Molecules with the random velocities exhibited a Maxwell-Boltzmann distribution, and the resulting line shape is a Gaussian. In this case the line shape and Doppler line halfwidth $\alpha_D$ are written by:

$$f(\nu - \nu_0) = \frac{1}{\alpha_D \sqrt{\pi}} \exp \left[ -\frac{(\nu - \nu_0)^2}{\alpha_D^2} \right]$$

(3.10)

$$\alpha_D = \nu_0 \sqrt{\frac{2k_B T}{mc^2}}$$

(3.11)

where $m$ is mass of molecules. The Doppler width is proportional to the wavenumber.

When the pressure broadened Lorentz halfwidth becomes comparable to the Doppler width, the clear relationship between line shape and altitude disappears. At these altitudes, line shapes become a convolution of the Lorentzian and Gaussian shape. The Voigt line shape gives a more realistic description of a line transition in the true atmosphere. However, there is no simple analytical function for the Voigt profiles, so various approximations are applied.

3.1.2 Scattering

Scattering is a physical process where incoming radiation is redistributed into all directions by the interaction with a particle. Two scattering processes are considered in the theory of scattering. In inelastic scattering the kinetic energy of an incident particle is not conserved: some of the energy of an incident particle can be used to excite the target in some manner (rotation, vibration, ionization etc.). On the
other hand the whole process of elastic scattering is conservative. This means that no exchange of energy happens during the interaction between molecules and atoms.

In this section we only consider elastic scattering. The characteristics such as the scattering attenuation $\gamma_s$ and the scattering angle distribution are determined by the size and refractive index of the interacting particle and wavelength. The scattering can be divided into 3 groups by the relationship between the particle’s diameter $d$ and wavelength.

$$\frac{\pi d}{\lambda} \begin{cases} 
\ll 1 & \text{(Rayleigh scattering)} \\
\approx 1 & \text{(Mie scattering)} \\
\gg 1 & \text{(Geometric scattering)} 
\end{cases} \quad (3.12)$$

For gas molecules in the UV/visible, the Rayleigh theory applies. The scattering cross-section in the scattering angle $\theta$ is given by:

$$\sigma_s = \frac{2\pi^5 d^6}{3 \lambda^4} \left( \frac{n^2 - 1}{n^2 + 2} \right)^2 (1 + \cos^2(\theta)) \quad (3.13)$$

where $n$ is the refractive index of the particle. More light is scattered in the backward and forward directions and the light is completely polarized at $\theta = 90^\circ$. Note that the $\lambda^{-4}$ dependence makes blue wavelengths much more efficiently scattered than red light. This is the reason why the sky is blue in daytime. Integration over $4\pi$ steradian and summed over all particles gives the Rayleigh scattering extinction coefficient $\gamma_R$, which is also proportional to $\lambda^{-4}$.

The Mie theory is applied to typical atmospheric aerosols. The Mie scattering cross section is more complex than that for Rayleigh scattering. The Mie scattering extinction $\gamma_M$ is generally proportional to $\lambda^{-1}$. Mie scattering is responsible for the white appearance of the clouds.

The last type of scattering is nonselective scattering. The wavelength-dependency of the nonselective scattering is infinitesimally small. A haze is a good example for nonselective scattering.

### 3.1.3 Radiative transfer equation

Incoming radiation passing through the atmosphere is attenuated due to scattering and absorption by aerosols and gases. The attenuation of the radiation depends on

1. the intensity of radiation,
2. the local concentration of the medium,
3. the effectiveness of the absorbers and scatterers.

The radiative transfer equations describes the intensity of the monochromatic radiation $L_\nu$ along a path $s$ in the medium.

$$\frac{dL_\nu}{ds} = -\gamma L_\nu + S_\nu \quad (3.14)$$
3.1. Principles

Here the first term on the right hand side indicates extinction of incident radiation \( L_\nu \) through the medium as a rate of the extinction coefficient \( \gamma = \gamma_a + \gamma_s \) and the second term describes the additional source radiation given as a source function \( S_\nu \).

For the case when scattering can be ignored such as at microwave frequencies and the cloud free atmosphere, the solution of Equation 3.14 is given as,

\[
L_\nu(s) = L_\nu(0) e^{-\tau_{\nu}(s,0)} + \int_0^s \gamma_a(s',\nu) B_\nu(T(s'))e^{-\tau_{\nu}(s',s') ds'}
\]

(3.15)

The opacity or optical thickness \( \tau_{\nu}(s_1, s_2) \) between two points \( s_1 \) and \( s_2 \) on the line-of-sight is defined as

\[
\tau_{\nu}(s, \nu) = \int_{s_1}^{s_2} \gamma_a(s, \nu) ds
\]

(3.16)

and the corresponding transmission \( \eta_{\nu}(s_1, s_2) \) is given by

\[
\eta_{\nu}(s) = e^{-\tau(s_1, s_2)}
\]

(3.17)

3.1.4 Retrieval, Inverse method and Optimal Estimation

In order to derive physical quantities such as temperature, pressure or ozone concentrations, solving the inverse problem is required. Often the inverse problem has no mathematically unique solution. This is called an ill-posed problem. To obtain a solution, an optimal estimation approach can be used in order to search for the appropriate constraint which determine the optimal solution from all possibilities which are consistent with the observation. In this section, the inverse method and optimal estimation solving for an atmospheric parameter is briefly introduced. Comprehensive details on this method are described in Rodgers [2000].

The forward model \( \mathbf{F} \), modeling radiative transfer and sensor characteristics, describes the measurement \( \mathbf{y} \) as a function of the true state parameters of the atmosphere \( \mathbf{x} \) and true model parameters \( \mathbf{b} \):

\[
\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \epsilon
\]

(3.18)

with a measurement error \( \epsilon \). The retrieved state vector \( \mathbf{\hat{x}} \) is given by the inverse model, \( \mathbf{I} \):

\[
\mathbf{\hat{x}} = \mathbf{I}(\mathbf{x}, \mathbf{b}, \mathbf{c}).
\]

(3.19)

In this equation, the forward model parameters \( \mathbf{b} \) are normally not perfectly known, for instance spectroscopic or instrumental parameters. \( \mathbf{c} \) stands for an additional parameters which constrain the retrieved state. For simplicity, we ignore \( \mathbf{b} \) from following paragraphs. All these quantities written in boldface mean the vectors.

**Optimal Estimation Method**

The Optimal Estimation Method (OEM) is employed in a linear retrieval algorithm. This method is based on Bayes’ theorem and combines statistical prior informations
on the variability of the searched parameters with measurements, using the associated errors as weights. Gaussian statistics are assumed for the measurement \( y \), the parameters \( x \) are to be retrieved. The estimated state vector \( \hat{x} \) is determined by minimizing the quantity \( \chi^2 \):

\[
\chi^2 = (y - F(x))^T S_y^{-1} (y - F(x)) + (x - x_a)^T S_a^{-1} (x - x_a)
\]

(3.20)

where \( S_y \) and \( S_a \) stand for the error covariance matrices for the measurement vector \( y \) and the a priori state vector \( x_a \). The first term in Eq. 3.20 determines how much measurement information is used, while the second term constrains the estimated state by the a priori state.

Normally, the forward model is not linear, thus a linearization by Taylor series expansion around a priori state \( x_a \) is required. Then Eq. 3.18 can be rearranged into,

\[
y = F(x_a) + K(x - x_a) + \epsilon
\]

(3.21)

The matrix \( K \) is the *differential weighting functions* of the forward model with respect to the state vector \( x \):

\[
K = \left. \frac{\partial F(x)}{\partial x} \right|_{x_a}
\]

(3.22)

Through the minimization of Eq. 3.20, the retrieved state can be obtained

\[
\hat{x} = x_a + (K^T S_y^{-1} + S_a^{-1})^{-1} K^T S_y^{-1} (y - K x_a)
\]

(3.23)

Since the problem is non-linear, this estimated \( \hat{x} \) is not accurate enough. Iterative processes are required to reach sufficient accuracy. If we apply a Newton iteration scheme, the optimal estimated state at iteration step \( i + 1 \) can be written as

\[
x_{i+1} = x_a + (K_i^T S_y^{-1} K_i + S_a^{-1})^{-1} K_i^T S_y^{-1} (y - F(x) + K_i (x_i - x_a))
\]

(3.24)

**Error analysis**

A linearization of the inverse function gives the contribution function matrix,

\[
D_y \equiv \frac{\partial \hat{x}}{\partial y} = (K^T S_y^{-1} K + S_a^{-1})^{-1} K^T S_y^{-1}
\]

(3.25)

The averaging kernel matrix, \( A \), represents the sensitivity of the retrieved state to the true state, which is given by

\[
A \equiv \frac{\partial \hat{x}}{\partial x} = D_y K
\]

(3.26)

Equ. 3.23 can be also written as

\[
\hat{x} = x_a + D_y (y - K x_a)
\]

(3.27)
The retrieved error can be obtained by rearranging Eq. 3.27

$$\hat{x} - x = D_y \epsilon + (A - I)(x - x_a)$$  (3.28)

Here $I$ is the identity matrix. The first term contributes the error due to statistical measurement noise and second term gives the smoothing (or null-space) error. Covariance matrices for each error are formulated, according to

$$S_M = D_y S_y D_y^{-1}$$  (3.29)

$$S_N = (A - I) S_a (A - I)^{-1}$$  (3.30)

where $S_M$ and $S_N$ are covariance matrices for the measurement and smoothing errors, respectively.

### 3.2 Limb Sounding

Understanding of stratospheric ozone and its chemistry has developed significantly since Dobson scientifically confirmed the existence of the ozone layer by constructing his first ground based UV spectrometer and monitoring of total ozone in the 1920’s [Dobson and Harrison, 1925]. A huge number of atmospheric remote sensing instruments and techniques have been designed for atmospheric chemistry. In general, remote sensing techniques are categorized into two main groups, i.e. passive and active. In addition, remote sensing is also subdivided on the basis of observation geometries, platforms and spectral coverage.

In the attached articles, ozone data observed from two limb sounding instruments SMR (Sub-Millimetre Radiometer) and SMILES (Superconducting subMillimeter-wave Limb Emission Sounder) were used. In the limb sounding technique, atmospheric constituents are derived from infrared or microwave radiation emitted by the atmosphere along the line of sight (LOS). The limb viewing geometry is illustrated in Figure 3.1. The limb sounding instruments obtain measurements at numerous tangent altitudes $z_0$. Most information in the emission originates near the tangent point $s_0$ because the LOS passes through comparatively narrow layers above $s_0$ due to the exponential decrease in pressure with height. The $90^\circ$ down looking nadir sounding geometry is for the case when the LOS crosses the surface and $s_0$ is found below the ground. Compared to the nadir geometry, limb viewing has advantages such as better vertical resolution and enhanced sensitivity to trace constituents. A disadvantage of limb viewing on the other hand are the higher probability of cloud interference. Both SMR and SMILES employ microwave emission detectors, which offers some insensitivity to clouds.

#### 3.2.1 Odin/SMR

Odin is a Swedish-led satellite mission in association with Canada, Finland and France. The Odin mission was designed for radio astronomy and limb sounding
Figure 3.1: A schematic diagram of a limb emission observation geometry. The instrument flies at the height of $z_r$ from the surface and receives emissions along the line of sight (LOS). The lowest altitude $z_0$ of the LOS is called the tangent altitude. The tangent point is defined by the position $s_0$ on the LOS.

of the Earth’s middle atmosphere [Murtagh et al., 2002, Nordh et al., 2003]. Odin was launched on 20 February 2001 into a sun-synchronous polar orbit with an inclination of $98^\circ$ at altitude of $\sim$600 km, and descending and ascending nodes at 6 and 18 hours local solar time respectively. Odin carries two different limb sounding instruments, OSIRIS (Optical Spectrograph/InfraRed Imaging System) and SMR (Sub-Millimetre Radiometer). The SMR instrument is made up of four tunable single-sideband Schottky-diode heterodyne microwave receivers which are connected to autocorrelation spectrometers [Frisk et al., 2003]. The ozone and $N_2O$ datasets used in paper A and B are products of the Chalmers processor version 2.1 [Urban et al., 2005a].

The SMR ozone profiles are retrieved from the emission line at 501.5 GHz. The altitude range is from approximately 17 km to 50 km with an altitude resolution of 2.5–3.5 km. An estimated single-profile precision is less than 20% ($\sim 1.5$ ppmv) [Urban et al., 2005a, 2006]. Validation studies of the SMR v2.1 ozone against balloon sonde measurements has been carried out by Jones et al. [2007] and Jégou et al. [2008].

The $N_2O$ profiles retrieved from the emission line at 502.3 GHz and cover altitudes in the range 12–60 km with a resolution of $\sim$1.5 km. The estimated systematic error is less than 12 ppbv [Urban et al., 2005a]. The validation of the $N_2O$ is reported by Urban et al. [2005b]. Other measurement comparisons with the Fourier Transform Spectrometer (FTS) onboard the Atmospheric Chemistry Experiment (ACE) and the Microwave Limb Sounder (MLS) on the Earth Observing System (EOS) Aura satellite are shown by Strong et al. [2008] and Lambert et al. [2007], respectively.
3.2. SMILES

SMILES (Superconducting subMillimeter-wave Limb Emission Sounder) is a passive atmospheric sensor attached to the Japanese Experiment Module (JEM) on board the International Space Station (ISS). SMILES was developed to measure vertical profiles of middle atmospheric species related to ozone chemistry by the Japan Aerospace Exploration Agency (JAXA) and the National Institute of Information and Communications Technology (NICT). SMILES was in operation from October 2009 until April 2010 providing atmospheric composition data typically within the latitude range of $38^\circ S - 65^\circ N$ [Kikuchi et al., 2010]. Approximately 1630 scans are obtained per day. Since the ISS has a non sun-synchronous orbit, the local time of SMILES measurement locations evolve over 24 h after 1 – 2 month. SMILES detects the ozone emission line at 625.371 GHz. An Acousto-Optical Spectrometer (AOS) with a bandwidth of 1.2 GHz with a resolution of 1.2 MHz produced the spectra. The superconductor-insulator-superconductor (SIS) mixers provided a low measurement noise; less than 0.7 K for a single AOS channel and a single spectrum. See Kikuchi et al. [2010] and Kasai et al. [2013] for further detail about the SMILES instrumentation.

Ozone data used in paper A were produced by the NICT level-2 chain version 2.1.5. The retrieval method described by Baron et al. [2011] was employed. The SMILES ozone profile covers altitudes from 16 to 90 km with a resolution of $\sim 3 - 4$ km and $\sim 6 - 10$ km for the stratosphere and mesosphere, respectively. The systematic error is lower than 0.3 ppmv in the stratosphere. The SMILES NICT ozone data was validated by Kasai et al. [2013].
Chapter 4

Transport Model

The model used in this thesis is called the DIAMOND (Dynamic Isentropic Assimilation Model for OdiN Data). This model is an off-line wind driven isentropic transport and assimilation model designed to simulate quasi-horizontal ozone transport in the lower stratosphere with low numerical diffusion [Rösevall et al., 2007a]. This chapter deals with the numerical transport model that computes the air flow for understanding the behavior of stratospheric chemical species especially ozone as described in chapter 2. The assimilation part of the model will be discussed in Chapter 5.

Two major numerical approaches are considered when solving the continuity equation of a fluid, the Lagrangian and the Eulerian. The idea of the Lagrangian method is to follow an arbitrary gas or air particle along its path and trace the change of its direction of motion, density and pressure along its way. On the other hand, the Eulerian approach focuses rather on the entire fluid field than on individual air parcels. Most numerical schemes can be divided into two classes based on Eulerian and semi-Lagrangian.

The Semi-Lagrangian scheme is an adapted Lagrangian approach. This scheme basically solves the transport problem based on Lagrangian methodology by capturing the behavior of a limited number of air parcels which initially coincide with a modeled area. Modeled concentrations at fixed time intervals are interpolated on the initial model grids. The advantages of semi-Lagrangian schemes are low computational cost and hence fast calculation time as time steps used in the schemes are not limited by stability conditions related to the grid spacing. However, the interpolation at each time interval does not conserve tracer concentrations. Therefore semi-Lagrangian schemes are not applied in this study.

An Eulerian scheme uses a fixed grid to divide space into small grid boxes. The transport problem is then solved numerically by estimating the tracer flow in and out of each box in a large number of discrete time-steps. The Eulerian scheme discretizes the continuity equation in time and space. Discretization errors in spatial differentiation causes numerical diffusion and oscillation. The Courant-Friedrichs-Lewy (CFL) condition is a necessary condition for stability while solving certain partial differential
equations. It means that the propagation speed of information must be faster than the propagation speed of the phenomenon. To achieve a fine grid size in the spacial domain under CFL condition, the time steps have to be shorter compared to mean flow. This can make Eulerian schemes quite time consuming, especially when implemented on global latitude/longitude grids that have very fine grid spacing near the poles. To achieve good numerical robustness, many schemes have been developed.

4.1 Modeling Eulerian tracer transport with low numerical diffusivity

The model used in the current studies employed the second order moments (SOM) tracer advection scheme of Prather [1986]. The SOM algorithm is regarded as sufficiently accurate for their many published scientific applications.

Eulerian transport scheme

For a flow field \((u, v, w)\), the advection equation of a (passive or active) tracer \(\psi\) in an Eulerian coordinate is:

\[
\frac{\partial \psi}{\partial t} = -u \frac{\partial \psi}{\partial x} - v \frac{\partial \psi}{\partial y} - w \frac{\partial \psi}{\partial z} \tag{4.1}
\]

We begin with the up-stream scheme. The \(\psi(x, t)\) donates the continuous tracer distribution function of an atmospheric constituent. To simplify the explanation, we consider 1-dimensional equations and describe the discretized physical state \(\psi\) at the center of the grid box \(x\) at the time \(t\) as below:

\[
\Psi^i_j \equiv \psi(x_j, t^i) \tag{4.2}
\]

\[
t^i \equiv i \Delta t \tag{4.3}
\]

\[
x_j \equiv jL. \tag{4.4}
\]

Here \(\Delta t\) is a temporal resolution (timestep) and \(L\) is the grid size. \(i\) and \(j\) denote the step number of time and grid, respectively. The Taylor expansion of \(\psi(x, t)\) around the grid box \(x_j\) is given as

\[
\psi(x, t) = \psi(x_j, t) + \psi'(x_j, t)(x - x_j) + \frac{1}{2} \psi''(x_j, t)(x - x_j)^2 + \frac{1}{6} \psi'''(x_j, t)(x - x_j)^3 + \ldots \tag{4.5}
\]

The discretization of \(\psi(x, t)\) in an individual grid box should be the mean of \(\psi_j(x, t)\),

\[
\Psi^i_j = \frac{1}{L} \int_{x_j-L/2}^{x_j+L/2} \psi_j(x, t) \, dx \tag{4.6}
\]

\[
= \psi(x_j, t) + \frac{L^2}{24} \psi''(x_j, t) + \frac{L^4}{384} \psi'''(x_j, t) + \ldots \tag{4.7}
\]

By the integration over one grid box, the odd terms in equation 4.5 are canceled out.
4.1. Modeling Eulerian tracer transport with low numerical diffusivity

Ideal discretization of the tracer field

The ideal discretization $\Psi(x, t)$ after one time step, which means $\psi(x, t)$ moves towards the point at the distance $u \Delta t$ without any numerical errors, will be given as

$$
\Psi_{j}^{i+1} = \frac{1}{L} \int_{x_{j} - L/2 - u_{j} \Delta t}^{x_{j} + L/2 - u_{j} \Delta t} \psi(x, t) \, dx
$$

(4.8)

Equation 4.8 can be converted under an assumption of positive $u_{j}$

$$
\Psi_{j}^{i+1} = \frac{1}{L} \int_{x_{j} - L/2 - u_{j} \Delta t}^{x_{j} + L/2} \psi_{j-1}(x, t) \, dx + \frac{1}{L} \int_{x_{j} - L/2}^{x_{j} + L/2 - u_{j} \Delta t} \psi_{j}(x, t) \, dx
$$

(4.9)

$$
= \left( \frac{u_{j} \Delta t}{L} \right) \psi(x_{j-1}, t) + \left( 1 - \frac{u_{j} \Delta t}{L} \right) \psi(x_{j}, t)
$$

$$
+ \frac{L}{2} \left( \frac{u_{j} \Delta t}{L} \right) \left( 1 - \frac{u_{j} \Delta t}{L} \right) \left( \psi'(x_{j-1}, t) - \psi'(x_{j}, t) \right)
$$

$$
+ \frac{L^2}{24} \left( \psi''(x_{j-1}, t) - \psi''(x_{j}, t) \right)
$$

$$
+ \frac{L^2}{24} \left( 1 - 2 \frac{u_{j} \Delta t}{L} \right)^3 \left( \psi'''(x_{j-1}, t) - \psi'''(x_{j}, t) \right) + \ldots
$$

(4.10)

The first order upstream scheme

An updated $\Psi_{j}^{i+1}$ by taking the first order upstream difference scheme is obtained as following expressions,

$$
\Psi_{j}^{i+1} = \begin{cases} 
\left( (L - |u_{j} \Delta t|) \Psi_{j}^{i} + |u_{j} \Delta t| \Psi_{j-1}^{i} \right) / L & \text{if } u_{j} \geq 0 \\
\left( (L - |u_{j} \Delta t|) \Psi_{j}^{i} + |u_{j} \Delta t| \Psi_{j+1}^{i} \right) / L & \text{if } u_{j} < 0 
\end{cases}
$$

(4.11)

Hence the expansion of $\Psi_{j}^{i+1}$ presented in equation 4.7 and the first order upstream scheme gives,

$$
\Psi_{j}^{i+1} = \left( \frac{u_{j} \Delta t}{L} \right) \psi(x_{j-1}, t) + \left( 1 - \frac{u_{j} \Delta t}{L} \right) \psi(x_{j}, t)
$$

$$
+ \frac{L^2}{24} \left( \frac{u_{j} \Delta t}{L} \right) \left( 1 - \frac{u_{j} \Delta t}{L} \right) \left( \psi''(x_{j-1}, t) - \psi''(x_{j}, t) \right)
$$

$$
+ \frac{L^4}{384} \left( \frac{u_{j} \Delta t}{L} \right) \left( 1 - \frac{u_{j} \Delta t}{L} \right) \left( \psi'''(x_{j-1}, t) - \psi'''(x_{j}, t) \right) + \ldots
$$

(4.12)

Comparing to the ideal case described in equation 4.10, some terms such as a first order derivative are not included in first order upstream scheme. The omitted terms correspond to and produce numerical diffusion with the diffusion constant $D = uL/2 \left( 1 - u \Delta t / L \right)$. A fine grid is needed to reduce the numerical diffusivity of the first order upstream scheme.
Second Order Moments

The basic idea of the SOM algorithm is that the quality of the transported field can be better preserved by tracking the moments of the sub-grid scale tracer distribution up to order 2. According to Prather [1986], the vector space of orthogonal polynomials can be expressed as the functions of $K_j(x)$. Each polynomial is expressed below,

\begin{align*}
K_0(x) &= 1 \\
K_x(x) &= x - \frac{L}{2} \\
K_{xx}(x) &= x^2 - Lx + \frac{L^2}{6}
\end{align*}

(4.13)

The tracer distribution in one grid can be approximated with ignoring moments of order 3 and higher,

\begin{equation}
\psi_j(x) = m_0K_0 + m_xK_x(x) + m_{xx}K_{xx}(x),
\end{equation}

(4.14) with coefficients $m_q$ represented in

\begin{equation}
m_q = \frac{\int_{-L/2}^{L/2} \psi_j(x)K_q(x) \, dx}{\int_{-L/2}^{L/2} K_q^2(x) \, dx}
\end{equation}

(4.15)

Note that $m_0$ is thus the average tracer concentration described in Equation 4.7 while $m_x$ and $m_{xx}$ denote the gradient of the tracer field in a grid box and the curvature, respectively. In SOM, the following quantities $S_q$ are also defined:

\begin{align*}
S_{0,j}^i &= \int_0^L \psi_j(x)K_0(x) \, dx = m_{0,j}^i L \\
S_{x,j}^i &= \frac{6}{L} \int_0^L \psi_j(x)K_x(x) \, dx = \frac{m_{x,j}^i L^2}{2} \\
S_{xx,j}^i &= \frac{30}{L^2} \int_0^L \psi_j(x)K_{xx}(x) \, dx = \frac{m_{xx,j}^i L^3}{6}.
\end{align*}

(4.16)

All of quantities have the units of tracer concentration times grid size. Thus these can be called moments of the tracer concentration. Considering the case where fluid in grid box $j - 1$ with width $\delta = u\Delta t$ is transported into the next box $j$ and an equal volume is flushed out from the box $j$ (see Figure 4.1). The discretized tracer concentration $\Psi_{j+1}^i$ is equal to $m_{0,j+1}^i$ and thus can be calculated from $S_{0,j}^i$. To determine the moment $S_{0,j+1}^i$, each grid cell is subdivided before transport into left and right boxes. Using superscripts L and R to refer to left and right sub-grids, moments are given as

\begin{align*}
S_{0,j}^{L,i} &= (1 - \alpha) \left[ S_{0,j}^i + \alpha S_{x,j}^i + \alpha(1 - 2\alpha) S_{xx,j}^i \right] \\
S_{x,j}^{L,i} &= (1 - \alpha)^2 \left( S_{x,j}^i + 3\alpha S_{xx,j}^i \right) \\
S_{xx,j}^{L,i} &= (1 - \alpha)^3 S_{xx,j}^i
\end{align*}

(4.17)
4.1. Modeling Eulerian tracer transport with low numerical diffusivity

**Figure 4.1:** Schematic picture of transportation of the tracer distribution $\psi(x)$. $\psi^i(x)$ in box($j-1$) at time $t^i$ moves to rightward by a distance $\delta = u \Delta t$. The new distribution $\psi^{i+1}(x)$ after one iteration can be a superposition of the moments of $\psi(x)$ in the left hand side (LHS) and right hand side (RHS) sub-boxes. Adapted from Hofmann and Morales Maqueda [2006].

and

\[
\begin{align*}
S^R_{0,j-1} &= \alpha \left[ S^i_{0,j-1} + (1 - \alpha)S^i_{x,j-1} + (1 - \alpha)(1 - 2\alpha)S^i_{xx,j-1} \right] \\
S^R_{x,j-1} &= \alpha^2 \left( S^i_{x,j-1} - 3(1 - \alpha)S^i_{xx,j-1} \right) \quad (4.18) \\
S^R_{xx,j-1} &= \alpha^3 S^i_{xx,j-1}
\end{align*}
\]

where $\alpha = \delta/L$. The resulting moments over the entire grid cell $j$ after $\Delta t$ will be calculated as the combination of moments of the left box and the right box,

\[
\begin{align*}
S^i_{0,j+1} &= S^R_{0,j+1} + S^L_{0,j} \\
S^i_{x,j+1} &= \alpha S^R_{x,j-1} + (1 - \alpha)S^L_{x,j} + 3 \left[(1 - \alpha)S^R_{0,j-1} - \alpha S^L_{0,j}\right] \quad (4.19) \\
S^i_{xx,j+1} &= \alpha^2 S^R_{xx,j-1} - (1 - \alpha)^2 S^L_{xx,j} \\
&\quad + 5 \left\{ \alpha(1 - \alpha)S^R_{x,j-1} - S^L_{x,j} \right\} + (1 - 2\alpha) \left[(1 - \alpha)S^R_{0,j-1} - \alpha S^L_{0,j}\right].
\end{align*}
\]

The SOM method achieves the transportation with very low numerical diffusion at the cost of its large memory demand. The SOM method requires storing not only
the mean tracer value \( \Psi \) but also the first- and second-order moments. In the two-dimensional case the number of moments per grid box increases to 6 because a cross moment i.e., \( S_{xy} \) is also needed.
Numerical models are powerful tools for studying chemical and atmospheric transport processes and monitoring ozone depletion. The result obtained from the models have been analyzed to investigate the relative roles of each process described in previous chapters. However, current model performance is still imperfect at reproducing the observed variability in the concentrations of ozone. This is due to uncertainties in chemical reactions, emissions, convective and diffusive transport and the complexity of physical processes in the model. These uncertainties make it hard to simulate detailed temporal and spatial structures in the stratospheric environment. Data assimilation (DA) is a technique that allows us to combine observational information and a numerical model in an optimal way. The basic concept of DA is to calculate a weighted average. The mathematical form of the weighted average is given as,

$$\hat{\Psi} = \frac{\sum_{i=1}^{n} \omega_i \Psi_i}{\sum_{i=1}^{n} \omega_i}$$  \hspace{1cm} (5.1)

Here $\hat{\Psi}$ and $\omega$ are the estimated states and the weights, respectively. Different DA approaches can be categorized by how to give the weights. There are two approaches to estimate weights in this chapter: these are the Minimum Variance (MV) and the Maximum of A Posteriori (MAP). The MV estimator results in the conditional mean of posterior probability density functions (PDF) while the MAP leads to the conditional mode instead.

### 5.1 Variation of Data assimilation technique

#### 5.1.1 Optimal Interpolation

Optimal Interpolation (OI) is a simple and basic method and based on the MV estimator. Following to Daley [1991] (chapter 4), we describe the discretized model in time as

$$\Psi_{t+1} = M(\Psi_t) + \epsilon_t$$  \hspace{1cm} (5.2)
where $\Psi$ is a true state vector, $M$ denotes the state transition model and $\epsilon_t^f$ is the process noise. At time $t$ an observation $y_t$ of the true state $\Psi_t$ is made according to

$$y_t = H_t(\Psi_t) + \epsilon_t^o$$

(5.3)

where $\epsilon^o$ is the observational noise with covariance $R$. $H_t$ is the observation model used to interpolate the state in the real atmosphere on the model grids to the positions of the observation. A model state $\Psi^b$ is also considered, which has the background error $\epsilon^b = \Psi^b - \Psi$ with covariance $B$. Under the assumption that the variations of the observation model $H$ are linear: for any $\Psi$ close enough to $\Psi^b$, an estimated (or analysis) value $\Psi^a$ is given by the following formulas:

$$\Psi^a = \Psi^b + \mathbf{K}(y - H(\Psi^b))$$

(5.4)

$$\mathbf{K} = (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1}$$

(5.5)

where $\mathbf{H}$ is a linear operator of $H$. The covariance for the estimated state can be calculated as

$$\mathbf{P}^a = \mathbf{B}(\mathbf{I} - \mathbf{K}^T) = (\mathbf{I} - \mathbf{K})\mathbf{B}$$

(5.6)

### 5.1.2 Variational method

The variational method is one of the major streams in DA. This method is based on the MAP. The three-dimensional analysis (3D-Var) and the four-dimensional analysis (4D-Var) are famous DA schemes using variational methods.

The analysis state of 3D-Var can be derived by minimizing the following cost function,

$$J = (y - H(\Psi^a_0))^T \mathbf{R}^{-1} (y - H(\Psi^a_0)) + (\Psi^a_0 - \Psi^b_0)^T \mathbf{B}^{-1} (\Psi^a_0 - \Psi^b_0)$$

(5.7)

OEM discussed in section 3.1.4 is also a version of the variational method for the one-dimensional case. Thus equations to estimate optimal states such as 5.7 are equivalent to equations 3.18-3.24 except for the notation\(^1\). Note that under the assumption of the Gaussian background and observation errors, the complete posterior PDF can be obtained and the MAP estimator is equivalent to the MV. For linear observation operators and Gaussian error statistics, 3D-Var and OI are therefore equivalent.

The 4D-Var is a simple generalization of the 3D-Var for four dimensional observations and currently used in many meteorological predictions by ECMWF, SMHI, Meteo-France, UK MetOffice, JMA etc. The cost function is extended to,

$$J = \sum_{n=0}^t (y_n - \tilde{H}_n(\Psi^a_0))^T \mathbf{R}^{-1} (y_n - \tilde{H}_n(\Psi^a_0)) + (\Psi^a_0 - \Psi^b_0)^T \mathbf{B}^{-1} (\Psi^a_0 - \Psi^b_0)$$

(5.8)

\(^1\) The notations in equations 3.18-3.24 can be converted for the 3D-Var as follows; $x \rightarrow \Psi$, $F(x) \rightarrow H(\Psi)$, $\mathbf{K} \rightarrow \mathbf{H}$, $\mathbf{S}_y \rightarrow \mathbf{R}$ and $\mathbf{S}_a \rightarrow \mathbf{B}$.
where $\hat{H}(\Psi_0) = H(M_t(M_{t-1}(\cdots M_1(\Psi_0))))$. By using tangent linear operators $\hat{H}$ of $H_t$ and $M$ of $M$, the gradient $g (\partial J/\partial \Psi_0)$ is

$$
g = \sum_{n=0}^t \hat{H}_n^T R^{-1}(y_n - \hat{H}_n(\Psi_0)) + B^{-1}(\Psi^0 - \Psi^b_0) \\
= \sum_{n=0}^t M_0^T M_1^T M_2^T \cdots M_{n-1}^T M_n^T H_n^T R^{-1}(y_n - \hat{H}_n(\Psi_0)) + B^{-1}(\Psi^a_0 - \Psi^b_0)$$

(5.10)

5.1. Variation of Data assimilation technique

Here $M_T$ is called as an adjoint of the transition model $M$. The first term in equation 5.10 indicates that the information of differences between the model and measurements propagates backward in time.

The variational methods compute optimal initial states by minimizing the cost function $J$. The 4D-Var scheme thus produces continuous analysis states which not only agree with observations but also following the processes in the model. Other advantage of 4D-Var is that it is available for complex nonlinear models. Furthermore, it is possible to enforce external constraints by inserting additional terms into cost function. On the other hand, the primary difficulty in implementing 4D-Var schemes is the need to develop an adjoint model for the system. In addition, the variational scheme requires several iterations to get optimal values. This fact makes the computational cost high.

5.1.3 Kalman filter

The Kalman filter (KF) is based on the MV estimator in the framework of a sequential DA as well as OI. The analysis equations of the linear KF are exactly the same with equations already described in the OI method. The weight $K$ is now also called "Kalman gain". The cycle of KF is performed with the following steps,

1. Initial forecast of the state by the model,
2. Kalman gain computation,
3. State calculation and error analysis,
4. State and error covariance forecast for the next time step.

The main difference between KF and OI is the updating of the error covariance matrix $B$ in step 4. The analysis error covariance $P^a$ is estimated from equation 5.6 and is also transported by the model. From equation 5.2, the forecast background error covariance matrix $B$ in the time step $\Delta t$ can be calculated as

$$B_{t+1} = M_t P^a_t M_t^T.$$  

(5.11)

Storing and transporting of full covariance matrices make the computational cost expensive. Moreover, model calculations by using a linearized matrix $M$ of the transportation may become numerically unstable [Ménard et al., 2000]. The DIAMOND
model adapted a less computationally expensive method suggested by [Ménard et al., 2000]. In this method the error covariance matrix $B$ would be replaced by an empirical correlation function $F(i,j)$ and the estimated error standard deviation in model grids $\sigma$,

$$B_{i,j} = \sigma_i \sigma_j F(i,j),$$

(5.12)

where $i$ and $j$ are indexes of the model grid. The correlation function $F$ is assumed to be a function of the geographical distance, $r(i,j)$, between each grid point and to be a Gaussian function tuned by an empirical parameter $L$ to fit to the transport model.

$$F(i,j) = \exp\left(\frac{-r^2(i,j)}{2L^2}\right)$$

(5.13)

The $\sigma$ is advected by the model described in chapter 4 along with the tracer $\Psi$. When the tracer is assimilated, the $\sigma$ should be updated by equation 5.6. Equation 5.6 can be written for the assimilated $\sigma$ as

$$(\sigma^a)^2 = (\sigma^b)^2 - \text{diag}(KB^T)$$

(5.14)

For the additional factors in error development such as missing chemical process in the model or imperfection of the transport scheme, a linear increment with time has been used to approximate the growth of uncertainty in the tracer fields [Rösevall et al., 2007a].

$$\sigma_{t+\Delta t} = \sigma_t + k_\sigma \cdot \Delta t$$

(5.15)

The error growth is terminated when $\sigma$ reaches the standard deviations expected in a climatology.
Chapter 6

Estimation of ozone loss

The subject of this thesis is to present a decadal series of ozone loss quantified by assimilating Odin/SMR data into a low diffusive transport and assimilation model DIAMOND. Assimilating a long record of Odin/SMR stratospheric ozone measurements is used to produce an ozone field in the transport model prior to the onset of ozone destruction in the polar vortex that occur when sunlight comes back to high latitudes in the early polar spring. Passively transported ozone fields without any further assimilation during the polar winter and spring can be useful as a reference fields describing the ozone layer that would have existed for the case of no polar ozone depletion occurring in the stratosphere. In this study, ozone depletion is subsequently mapped and quantified by subtracting the passively transported reference ozone fields (passive ozone) from ozone fields that have been continuously forced by assimilation of satellite data;

$$\Delta \Psi_n \equiv \Psi^n_a - \Psi^n_p,$$  \hspace{1cm} (6.1)

where $\Delta \Psi_n$, $\Psi^n_a$ and $\Psi^n_p$ represent chemical ozone change, active ozone and passive ozone at time $n$, respectively.

The passive ozone fields can alternatively be compared to individual satellite ozone profiles. And the vortex mean ozone loss can be calculated statistically as an averaging value of such comparisons. However satellite observations do not always evenly cover the entire polar vortex. This fact leads to that estimated ozone depletion may have local biases in measurements. On the other hand, assimilation of satellite datasets allows us to distribute the ozone mixing ratios at different locations in the stratosphere in a statistical way. Subtraction of passive ozone from active ozone is thus a beneficial method for visualizing ozone destruction. In addition, assimilating a large data set of satellite measurements can furthermore includes less random noise in measurements than using individual satellite measurements.
6.1 Development of the estimation of ozone depletion based on data assimilation information

The estimation of ozone loss described as above requires at least two calculations, i.e. active ozone fields and passive ozone fields. Currently a new less computationally intensive method to estimate ozone loss based on the data assimilation technique has been developed. The new method is described by the following equation.

\[ \Delta \psi_n = \sum_{p=0}^{n-1} M_{n,n-p}I(n-p) \]  \hspace{1cm} (6.2)

Here \( I(n) \) is an analysis increment at time \( n \), and \( M_{n,n-p} \) is a transport model from time \( n - p \) to \( n \). The analysis increment, which is given as the second-term on the right hand side in equation 5.4 in a Kalman filter, can be expressed as,

\[ I(n) = \psi_n^a - M_{n,n-1}\psi_{n-1}^a. \]  \hspace{1cm} (6.3)

The propagation of the analysis increment from \( n - 1 \) to time \( n \) is given as

\[ M_{n,n-1}I(n-1) = M_{n,n-1}\psi_{n-1}^a - M_{n,n-2}\psi_{n-1}^a, \]  \hspace{1cm} (6.4)

and similarly we get

\[ M_{n,n-2}I(n-2) = M_{n,n-2}\psi_{n-2}^a - M_{n,n-3}\psi_{n-3}^a, \]  \hspace{1cm} (6.5)

for different time periods. The right hand side of Equation 6.2 is a summation of equations 6.3, 6.4 and 6.5 up to \( n - (n - 1) \),

\[ \sum_{p=0}^{n-1} M_{n,n-p}I(n-p) = I(n) + M_{n,n-1}I(n-1) + M_{n,n-2}I(n-2) + \ldots + M_{n,1}I(1) \]

\[ = \psi_n^a - M_{n,0}\psi_0^a. \]  \hspace{1cm} (6.6)

Since the passive ozone is initialised with the initial analysis (time 0),

\[ \psi_n^p = M_{n,0}\psi_0^a. \]  \hspace{1cm} (6.7)

Thus equation 6.2 theoretically equals to equation 6.1. This also demonstrates that the the ozone loss estimation described in appended papers is the exact accumulated ozone change due to chemical reactions which are not included in a transport model. Since the new method is based on use of analysis increment, we call it "increment method" in the following paragraph.
6.1. Development of the estimation of ozone depletion based on data assimilation information

Figure 6.1: An example of the increment method compared with our standard ozone loss estimation. Panels (a) and (b) show the horizontal distribution of chemical ozone change on 30 March 2011 at a potential temperature of 500K derived using the standard method and the increment method described in equation 6.1 and 6.2, respectively. Odin/SMR 544 GHz ozone measurements were assimilated into a 2D advection model to obtain these maps. The analysis period of this experiment is from 1 Jan 2011 to 30 March 2011. Ozone loss is illustrated as blue areas, while the red area shows regions of production. The difference between (a) and (b) is presented in the panel (c).

Evaluation of the increment method has been performed using the DIAMOND model and Odin/SMR 544 GHz ozone. Figure 6.1 shows an example of a comparison between the previous and new methods. We have excellent agreement between two methods in the results of several experiments on the 2D advection model. However, these two ozone maps never perfectly match, since the numerical schemes of the active and passive ozone fields and the increment fields are different. For the case of a three months run in the 2D advection model, the maximum discrepancy between two ozone change calculations is approximately $\pm 0.2$ ppmv. Further development of this methodology requires some more tasks to be solved. For instance, in experiments examining Antarctic ozone loss with the 3D advection model, larger discrepancies occurred below 500K inside the polar vortex. This was because the increment method detected unexpected ozone enhancement at the height. We do not yet have a clear understanding of this. It may in part be due to larger numerical diffusion that can be generated by the additional vertical transport and attending boundary conditions. Further study is in progress.
Chapter 7

Summary of papers

7.1 Summary Paper A

Paper A studied ozone depletion for the 2009/2010 Arctic winter using measurements taken by the Odin/SMR and SMILES instruments. The object of the paper was to illustrate the applicability of the high sensitivity ozone measurements observed by SMILES for polar ozone loss assessment despite the lack of full geographical coverage.

To begin with, a new scheme for calculating the vertical transport was directly introduced into the DIAMOND model and tested for the long-lived species N$_2$O. As noted in Section 2.1, vertical motions can normally be ignored when using potential temperature as a vertical coordinate in the stratosphere. However, there is a strong cross isentropic descent inside the vortex due to radiative cooling of air masses during the polar night. In the previous work using the DIAMOND model the diabatic descent was taken into account by tracing N$_2$O [Rösevall et al., 2007a,b, 2008]. Rösevall et al. [2008] also used the diabatic heating rate Q [K/sec] derived from SLIMCAT 3d chemical transport model calculations [Chipperfield, 2006] to estimate contributions of the vertical transport and added the contributions on the ozone loss amount post-priori. In order to improve the methodology, the first-order upstream method was implemented because the general descent rate is small enough to satisfy the CFL condition. Model calculations with the new vertical transport scheme showed good performance when comparing modelled N$_2$O fields and assimilated N$_2$O fields.

The ozone loss for the 2009/2010 winter was quantified by subtracting modelled ozone from the assimilated ozone fields. Instrumental biases due to geographical issues such as the number of available measurements and the measurement precision appeared in assimilated ozone results as differences between SMILES and SMR. Paper A reported that the vortex mean of the SMILES assimilation tends to emphasize the contribution near the vortex edge, which results in larger ozone loss estimates for the SMILES assimilation for all periods. However assimilations from both instruments show clear similarities in the ozone distributions and their time evolution. Paper
A concluded that approximately 1 ppm (20%) ozone loss occurred and that two mechanisms were found to be behind the loss in the 2009/2010 winter. Chlorine catalytic ozone destruction chemistry (see section 2.4) was dominant below a potential temperature of 500K from the middle of January. Other chemical reactions related to NO$_x$ can be considered to induce depletions above the 600K potential temperature surface which were observed from the end of January.

### 7.2 Summary Paper B

Chemical ozone losses during the 2002–2013 winters were estimated by using Odin/SMR ozone measurements. For the first time, we have shown the inferred ozone loss derived from SMR ozone retrieved from a emission line centred at 544 GHz in this paper.

An internal comparison between the two datasets of 501 GHz and 544 GHz ozone has been carried out. The comparison showed a clear improvement in the results when using the 544 GHz ozone particularly in the early polar spring periods. Paper B also compared the loss in the 2004/2005 Arctic winter to the losses derived from other methods and instruments as presented in Sonkaew et al. [2013]. The comparison suggests that our ozone loss estimate is in agreement with other results within the error range although slightly smaller. 544 GHz ozone loss in the Arctic 2004/2005 winter is in good agreement with SCIAMACHY loss below 450K, while showing no loss around 550K where SCIAMACHY detected 0.5 ppmv loss. The comparison with Antarctic ozone depletions by Kuttippurath et al. [2015] shows agreement with MLS ozone loss within 0.1 ppmv, while our results were constantly 0.3 ppmv lower than Mimosa-Chim model calculations.

In the Northern hemisphere, our results show large inter-annual variability. Three types of chemical ozone losses are found to occur in cold, warm winters and winters intermediate between cold and warm. The cold type loss maximises in March below 500K as in the Southern hemisphere. For cold type loss, the maximum loss happened in 2010/2011 with a loss in volume mixing ratio of 2.1 ppmv at 450K. In the warm winters losses of 1.5 ppmv took place at 700K and are related to the occurrence of major sudden stratospheric warming (SSW) events in January. In the Southern hemisphere during the 2002 to 2012, chemical ozone losses generally started in mid-August and grew to 2.5 ppmv by the end of October. The vertical extent of this loss was 425 – 550K. Small inter-annual variability was seen. Antarctic winters had approximately 80 DU loss in the stratospheric column, except for 2002 when an unusual SSW occurred in the mid-September. In both hemisphere stratospheric ozone columns show a small increase, nevertheless the statistical confidence is not high enough to conclude that recovery is underway.
7.3 Summary Paper C

In paper B, we found different types of chemical ozone loss in the northern hemisphere characterised by different altitude ranges and periods of maximum peak loss. These are the halogen-induced loss in the lower stratosphere and the NO$_x$-induced loss in the mid-stratosphere. Konopka et al. [2007] showed that NO$_x$-induced loss is comparable to ClO$_x$-induced loss, in a case study of the SSW in 2002/03 winter. They concluded that meridional transport from lower latitudes is the main reason of the NO$_x$-induced loss in 2002/03 winter. This paper re-assessed and quantifies their conclusions over a much longer period having more major SSW events. In this study we have extended the analysis period for each assimilation experiment until the end of April to see how ozone loss occurs even in the weak polar vortex or remnants of the vortex after the final warming.

We focused on the loss in the 2010/2011 winter to investigate the halogen-induced loss. The very strong vortex during the winter perfectly separated ozone change distributions in the lower and middle stratosphere, which created the largest depletion in the last decade. We also found that similar tendencies are seen in the other relatively cold winters, such as 2004/05 and 2007/08. We selected the four winters 2003/2004, 2005/2006, 2008/2009 and 2012/2013 due to their prolonged zonal wind reversal and performed a composite calculation. Then we compared the composite with the loss of winter 2010/2011 as a representative of the halogen-induced case. Pronounced mid-stratospheric ozone losses are consistent with occurrences of such major SSW events, and their attendant large transport from lower latitudes.
This thesis quantifies ozone depletions in several Arctic and Antarctic winters using a long record of stratospheric ozone profiles measured by Odin/SMR and a data assimilation technique. According to paper B and C, ozone holes in the Southern hemisphere have continued to occur during the last decades, while large inter-annual variability in ozone loss is found in the Arctic lower and middle stratosphere. Hence it is still necessary to monitor stratospheric ozone depletion and its feedback on atmospheric processes.

During the past two decades, 1990-2010, many satellite missions based on limb observation have been launched. These satellite missions have provided vertically resolved profiles of many atmospheric constituents with an unprecedented spatial and temporal coverages. It allows us to use a large number of measurements of chemical species, both long-lived and short-lived, that are important to conduct scientific studies. At the time of writing the golden age of middle atmospheric observations has nearly ended. ENVI SAT and its instruments were lost in April 2012. JEM/SMILES failed in April 2010. The other missions (Odin OSIRIS and SMR, SCISAT-1 with ACE and MAESTRO, Aura with MLS and OMI, and TIMED with SABER) are getting older and are far beyond their expected lifetimes. The NASA OMPS and SAGE-III are short term missions, and provide limited set of species. Furthermore SAGE-III will not see over the poles.

**Filling gaps in observations by using DA**

In the future we may have to depend more upon numerical models than real measurements to study phenomena in the middle atmosphere. Model predictions however vary depending on initial settings, parameterisation and boundary conditions. DA can play a role model calculations by providing optimal use of the few available measurements and thus help to compensate for the lack of limb observations. This also allow us to assess the impact of chemical species, which have significant sources and sinks that are difficult to model accurately, on stratospheric ozone.

As described in section 2.2.2, EPP in the mesosphere and upper stratosphere can
generate NO which is rapidly converted into NO$_2$. NO$_x$ can descend into the mid polar stratosphere, reacting with other molecules along the way. Hence it is expected that mesospheric NO$_x$ can change the amount of stratospheric ozone. However no observations have found significant evidence for its impact on ozone loss. SMR observes NO, which is retrieved from an emission line centred at 551.7 GHz over an altitude range of 30 to 115 km with a vertical resolution of about 7 km [Sheese et al., 2013]. NO measurements from this frequency channel were available only one day per month before 2007, and approximately four days per month since 2007. I attempted DA of SMR NO as well as ozone in order to obtain a better interpolation in time and space constrained by transport in the atmosphere. Figure 8.1 presents correlation plots between assimilated NO and inferred ozone loss in 2010/2011 and 2012/2013 Arctic winters, respectively, at the potential temperature of 900K which is the lower altitude limit of the NO measurement. The two loss peaks at the right and centre in figure 8.1 can be explained with horizontal mixing of NO$_x$-rich air. Due to the major SSW in January 2013, losses are larger. The 2013 winter has an additional third peak in the loss around 90 days from the SSW, which is not seen in winter 2011. The period of this third peak in loss is consistent with when an unusually strong tongue of NO descent reached 40 km (∼900K) [Pérot et al., 2014]. However since the current version of the DIAMOND model does not include any chemistry the NO assimilation is highly uncertain as we do not correctly handle NO$_2$-NO conversion.
Supplement the DIAMOND model with full chemistry

In order to obtain correct estimates of ozone and related species, we have to consider ozone chemistry since stratospheric ozone interacts with many other species. Ozone chemistry is very complex (e.g. see in chapter 2). In future work a full chemistry scheme should be introduced into the passive transport scheme. This will allow us to predict local scale chemical conditions in the atmosphere in a manner similar to numerical weather predictions (NWPs). However, the chemical forecast is more challenging than NWPs [Lahoz and Errera, 2010]. The full scale chemistry would need several species to be assimilated and transported. The chemical time scales differ widely in the atmosphere. And the chemical equations are strict and their effects are clearly seen as strong correlation between species. Therefore chemical predictions require accurate specification of sources and sinks, parametrization of physical processes and radiation, and chemical initial conditions as well as dynamical parameters. The BASCOE (Belgian Assimilation System for Chemical ObsErVations) system based on 4D-Var is one of the DA system with full chemistry [Errera et al., 2008, Errera and Ménard, 2012]. However, that system requires considerable computing power. In this thesis, I showed a new methodology to extract chemical ozone change information, which can make an analysis speed up. The final goal of the DIAMOND is a minimal DA system with full chemistry and the increment method so that we can take an instant photo of the responses of chemical species to the atmosphere with ease.
BIBLIOGRAPHY


Bibliography


J Urban, N Lautié, E Le Flochmoën, C Jiménez, P Eriksson, J de La Noé, E Dupuy, M Ekström, L El Amraoui, U Frisk, D Murtagh, M Olberg, and P Ricaud. Odin/SMR limb observations of stratospheric trace gases: Level 2 processing of


