

---

Electronic Thesis and Dissertation Repository

---

7-30-2019 10:00 AM

# Development of a 1-Dimensional Data Assimilation to Determine Temperature and Relative Humidity Combining Raman Lidar Backscatter Measurements And a Reanalysis Model

Shayamila N. Mahagammulla Gamage, *The University of Western Ontario*


Supervisor: Robert Sica, *The University of Western Ontario*

Co-Supervisor: Alexander Haefele, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Physics

© Shayamila N. Mahagammulla Gamage 2019

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>

 Part of the [Atmospheric Sciences Commons](#), [Meteorology Commons](#), [Statistical Methodology Commons](#), and the [Statistical, Nonlinear, and Soft Matter Physics Commons](#)

---

## Recommended Citation

Mahagammulla Gamage, Shayamila N., "Development of a 1-Dimensional Data Assimilation to Determine Temperature and Relative Humidity Combining Raman Lidar Backscatter Measurements And a Reanalysis Model" (2019). *Electronic Thesis and Dissertation Repository*. 6356.  
<https://ir.lib.uwo.ca/etd/6356>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

# Abstract

Water vapor is the most dominant greenhouse gas in Earth's atmosphere. It is highly variable and its variations strongly depend on changes in temperature. Atmospheric water vapor can be expressed as relative humidity (RH), the ratio of the partial pressure of water vapor in the mixture to the equilibrium vapor pressure of water over a flat surface of pure water at a given temperature. Liquid water can exist as super-cooled water for temperatures between 0°C to -38°C. Thus, RH can be measured either relative to water ( $RH_w$ ) or to ice ( $RH_i$ ).  $RH_i$  measurements are important in the upper tropospheric region, where the temperature is always less than 0°C, to study ice supersaturation (ISS) and its relation to the formation of cirrus clouds.

I present three studies all using a mathematical scheme called the optimal estimation method (OEM). The OEM is an inverse method that determines the most probable state consistent with the measurements and a priori knowledge. These studies use parts of a large set of existing measurements from the Raman Lidar for Meteorological Observations (RALMO) instrument located at the meteorological observatory in Payerne, Switzerland.

I first develop an OEM retrieval for temperature using RALMO's two pure rotational Raman (PRR) channel measurements. Retrieved temperatures show excellent agreement with coincident balloon-borne radiosonde measurements. A second OEM scheme is introduced to retrieve  $RH_w$  directly from RALMO measurements of back-scatter due to water vapor and nitrogen. I validate the OEM retrievals developed for temperature and  $RH_w$ . I then combine the OEM-retrieved temperature and  $RH_w$  with data from the European Centre for Medium-Range Weather Forecasts Re-analysis (ERA5) to compute a new and improved temperature and relative humidity product. The retrieval is enhanced by assimilating it with the ERA5 data. The quality of the  $RH_w$  retrievals from the assimilated OEM scheme greatly improves over retrievals which use less accurate a priori information.

Thirdly, I retrieve  $RH_i$  to detect ISS layers. I find the frequency of ISS layers in the free troposphere over Payerne to be about 27% using 82.5 hours of measurements.

**Keywords:** 1D Var Data Assimilation, Reanalysis, Optimal Estimation Method, ERA5, Raman lidar, Rotational Raman temperature, UTLS, Water vapor mixing ratio, Relative humidity, Ice supersaturation, Particle extinction

## Summary for Lay Audience

Water vapor is the most dominant greenhouse gas in Earth's atmosphere that is highly variable. Variations of the atmospheric water vapor strongly depend on changes in temperature. Accurate estimates of humidity and temperature and as well as the uncertainties associated are required for both weather and climate forecasting purposes. I present a new mathematical and statistical approach to estimate both atmospheric humidity and temperature using Raman lidar backscatter measurements. The new method provides full uncertainty budgets for each estimated temperature and relative humidity profile, that represent the errors due to instrumentation, estimation method and so on. I have also combined the Raman lidar measurements into the data from the ERA5 that is the latest major global reanalysis produced by European Centre for Medium-Range Weather Forecasts (ECMWF), to enhance the quality of the humidity and temperature estimates. My results show that the quality of the temperature and humidity retrievals are greatly improved and agree best with the measurements made by coincident radiosondes.

## Co-Authorship Statement

The entire thesis is written under the supervision of Dr. Robert Sica and Dr. Alexander Haeferle. Main research ideas are generated by Dr. Sica and Dr. Haeferle.

I was provided with raw Raman lidar measurements from Raman Lidar for Meteorological Observations (RALMO) and radiosonde measurements from MeteoSwiss located in Payerne, Switzerland. Both lidar and sonde measurements have been used in the work presented in Chapters 2,3, and 4.

The work presented in Chapter 2 was done in collaboration with Dr. Gianni Martucci. The traditional Raman lidar temperature analysis was performed and provided to me by Dr. Gianni Martucci. I was provided with the RALMO data processing codes and the OEM codes written by Dr. Sica and Dr. Haeferle in their water vapor and Rayleigh temperature studies. All the necessary MATLAB codes related to the OEM Raman temperature algorithm were written by me. I also conducted the OEM analysis for temperature retrievals and compared the results with the radiosonde and traditional Raman lidar temperatures.

The work done in Chapters 3 and 4 is based on the studies of OEM Raman temperature study by me and the OEM water vapor mixing ratio by Dr. Sica and Dr. Haeferle. In the work presented in both Chapters 3 and 4, I have used the ERA5 hourly reanalysis data obtained from the ECMWF data archive. I was responsible for providing the MATLAB codes for the OEM relative humidity retrievals, and I performed the OEM analysis.

## Acknowledgements

This work would not be possible without the guidance of my two supervisors Dr. Robert Sica and Dr. Alexander Haefele. I would thank both for supporting me with my work and providing me with opportunities to present my work in conferences and various other platforms. I would like to thank MeteoSwiss in Payerne (the federal Office of Meteorology and Climatology) in Switzerland for allowing me to use the data from their Raman lidar system. Understanding of the RALMO system and the detection of rotational Raman spectrum was possible thanks to Dr. Valentine Simeonov of Ecole Polytechnique Federale de Lausanne. I would like to thank Dr. Gianni Martucci for his invaluable instruction and inputs for my work. Also, I would like to thank everyone working at MeteoSwiss for their great support and motivation.

This work is not only supported by a MeteoSwiss and Western University but also by National Science and Engineering Research Council of Canada and the Canadian Space Agency under the Arctic Validation and Training for Atmospheric Research in Science (AVATARS) program.

I also thank the Western writing support center and Patricia Sica for their assistance in editing and proofreading all my papers and my thesis. I am thankful to my colleague Ghazal Farhani for being there for me for everything. I would also like to thank Dr. Emily McCullough, the first person whom showed me how the lidars work and how to make things work. Also, I am thankful for all the great inputs and support given by my purple crow lidar group members throughout the years of my PhD.

I am appreciative toward Dr. Blaine A. Chronik and Dr. Aaron Sigut who were on my academic committee and providing me with great advice and motivation to complete my thesis. I also thank Dr. Kanthi Kaluarachchi for being a motherly figure to me and guiding me in many ways to succeed. I thank all the staff at Physics department who were always helpful and supportive. I have no words to thank Rajitha, my roommate for tolerating me for all these years and I am surely indebted for her support.

Finally, I would like to thank my parents and my sisters. I am thankful for their unconditional love and support. Everything I achieve in my life including the PhD I would like to dedicate to my wonderful family.

# Contents

<b>Abstract</b>	<b>ii</b>
<b>Summary for Lay Audience</b>	<b>iii</b>
<b>Co-Authorship Statement</b>	<b>iv</b>
<b>Acknowledgements</b>	<b>v</b>
<b>List of Figures</b>	<b>x</b>
<b>List of Tables</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Overview . . . . .	1
1.2 Introduction to Earth’s atmosphere . . . . .	6
1.2.1 Temperature structure . . . . .	8
Troposphere . . . . .	8
Stratosphere . . . . .	10
Upper troposphere - lower stratosphere (UTLS) . . . . .	10
1.2.2 Atmospheric humidity . . . . .	12
Mixing ratio . . . . .	12
Specific humidity . . . . .	13
Vapor pressure . . . . .	13
Saturated vapor pressure . . . . .	13
Relative humidity . . . . .	15
Ice supersaturation (ISS) . . . . .	15

1.2.3	Instruments and techniques to measure relative humidity . . . . .	16
1.3	Lidars and atmospheric measurements . . . . .	18
	General description of a lidar . . . . .	18
	Atmospheric scattering related to lidar . . . . .	20
	Elastic scattering and Raman lidar equations . . . . .	24
1.4	Raman lidar for Meteorological Observations (RALMO) . . . . .	29
1.4.1	Pure rotational Raman (PRR) lidar . . . . .	31
1.4.2	Vibrational Raman/water vapor system . . . . .	31
1.5	Traditional Raman lidar algorithms . . . . .	32
1.5.1	Temperature measurements . . . . .	32
1.5.2	Water vapor mixing ratio . . . . .	33
1.5.3	Relative humidity . . . . .	35
1.6	Optimal estimation method (OEM) . . . . .	37
1.6.1	Advantages of implementing the OEM for lidar temperature and rela- tive humidity retrievals. . . . .	40
<b>List of Appendices</b>		<b>1</b>
<b>2 Application of the OEM for temperature retrievals</b>		<b>49</b>
2.1	Overview . . . . .	49
2.2	The Raman Lidar for Meteorological Observations . . . . .	51
2.3	The PRR lidar equation . . . . .	52
2.4	Application of the OEM for PRR temperature retrieval . . . . .	53
2.4.1	Brief review of the optimal estimation method . . . . .	53
2.4.2	The forward model for a PRR lidar . . . . .	54
2.4.3	Implementation of the RR temperature retrieval . . . . .	58
2.5	Results from the temperature retrieval . . . . .	61
2.5.1	Case 1: Nighttime with clear conditions . . . . .	61
2.5.2	Case 2: Daytime with clear conditions . . . . .	68
2.5.3	Case 3: Nighttime with cirrus cloud . . . . .	70
2.5.4	Case 4: Nighttime with lower level cloud . . . . .	74

2.6	Discussion . . . . .	76
2.6.1	Conclusion . . . . .	80
<b>3</b>	<b>Assimilated Raman lidar and ERA5 relative humidity retrievals</b>	<b>85</b>
3.1	Overview . . . . .	85
3.2	Data used in 1DVar reanalysis . . . . .	87
3.2.1	RAman lidar for Meteorological Observations (RALMO) . . . . .	87
3.2.2	ERA5 reanalysis data . . . . .	87
3.3	1D Var retrieval of relative humidity from ERA5-reRH . . . . .	89
3.3.1	Forward model . . . . .	89
3.3.2	Error covariance matrices . . . . .	93
	Measurement noise . . . . .	93
	<i>A priori</i> (background) relative humidity and temperature . . . . .	93
	Particle extinction and overlap . . . . .	96
	Background, lidar constants, and dead times . . . . .	96
	Forward model parameters . . . . .	97
3.3.3	Other 1D Var retrieval specifications . . . . .	98
3.3.4	Characterization of ERA5-reRH . . . . .	98
	Case 1: Nighttime, cloudy sky, 28 August 2012 2241-2311 UT . . . . .	98
	Case 2: Daytime, clear sky, 10 September 2011 1010 - 1040 UT . . . . .	100
3.4	1D Var retrieval of relative humidity from Raman lidar and U.S. standard at- mospheric model . . . . .	103
3.5	Results . . . . .	103
3.5.1	Validation of the reanalysis against radiosonde measurements . . . . .	103
	Nighttime . . . . .	104
	Daytime . . . . .	107
3.6	Discussion . . . . .	110
3.7	Conclusion . . . . .	113
<b>4</b>	<b>Detecting ice supersaturation layers over Payerne</b>	<b>118</b>
4.1	Overview . . . . .	118



4.2	Method . . . . .	120
4.3	Preliminary results . . . . .	121
4.4	Discussion and conclusions . . . . .	128
<b>5</b>	<b>Conclusions and future work</b>	<b>133</b>
5.1	Summary . . . . .	133
5.2	Conclusion . . . . .	136
5.3	Future work . . . . .	139
<b>A</b>	<b>Jacobians of the three types of PRR temperature forward models</b>	<b>140</b>
<b>B</b>	<b>Conventional method of Raman lidar relative humidity</b>	<b>141</b>
<b>C</b>	<b>1D Var retrieval of relative humidity from Raman lidar</b>	<b>145</b>
C.1	Case 1: Nighttime, clear sky, 09 September 2011 2200 - 2215 UT . . . . .	146
C.2	Case 2: Daytime, clear sky, 10 September 2011 1000 - 1015 UT . . . . .	148
C.3	Temperature, relative humidity and particle extinction time series . . . . .	154
C.4	Summary . . . . .	157
<b>D</b>	<b>Estimated bias and the covariance for relative humidity</b>	<b>158</b>
	<b>Curriculum Vitae</b>	<b>160</b>

# List of Figures

1.1	An atmospheric temperature profile based on the US Standard Atmosphere values showing the regions of the atmosphere based on temperature structure. . . .	9
1.2	The important processes in the UTLS region . . . . .	11
1.3	Variation with temperature of the saturation vapor pressure over a plane surface of pure water and the difference between the saturation vapor pressure over a plane surface of water to ice . . . . .	14
1.4	Schematic drawing of lidar system. . . . .	19
1.5	The different types of light scattering . . . . .	22
1.6	Raman backscatter spectrum of the atmosphere for an incident laser wavelength of 355 nm . . . . .	23
1.7	Optical diagram of the transceiver of RALMO (PB - Pellin Broka prism). Adapted from Dinov et al. (2010). . . . .	30
1.8	Schematic of the OEM retrieval algorithm . . . . .	38
2.1	Count rate for 30 min of coadded RALMO measurements from 2300 UT on 09 September 2011, a clear night . . . . .	62
2.2	Difference between the forward model and clear nighttime RALMO measurements on 09 September 2011 . . . . .	63
2.3	Averaging kernels and vertical resolution for temperature retrievals from the clear nighttime RALMO measurements on 09 September 2011 . . . . .	64
2.4	Retrieved temperature profile and the statistical uncertainty using the OEM from the clear nighttime RALMO measurements on 09 September 2011 in comparison with the coincident sonde measurements . . . . .	65

2.5	Random uncertainties and systematic uncertainties due to the forward model parameters for the temperature retrievals from the clear nighttime RALMO measurements on 09 September 2011 . . . . .	66
2.6	Retrieved geometrical overlap function from the clear nighttime RALMO measurements on 09 September 2011 . . . . .	67
2.7	Count rate for 30 min of clear coadded RALMO measurements from 1100 UT on 10 September 2011 . . . . .	68
2.8	Difference between the forward model and the clear daytime RALMO measurements on 10 September 201 . . . . .	69
2.9	Averaging kernels and vertical resolution for temperature retrievals from the clear daytime RALMO measurements on 10 September 2011 . . . . .	69
2.10	Retrieved temperature profile and the statistical uncertainty using the OEM from the clear daytime RALMO measurements on 10 September 2011 in comparison with the coincident sonde measurements . . . . .	70
2.11	Random uncertainties and systematic uncertainties due to the forward model parameters for the temperature retrievals from the clear daytime RALMO measurements on 10 September 2011 . . . . .	71
2.12	Count rate for 30 min of coadded RALMO measurements from 2300 UT on 05 July 2011 with the presence of a cirrus cloud . . . . .	72
2.13	Difference between the forward model and the nighttime RALMO measurements on 05 July 2011 with the presence of a cirrus cloud . . . . .	72
2.14	Averaging kernels and vertical resolution for temperature retrievals from measurements on 05 July 2011 with a cirrus cloud at 6 km height . . . . .	73
2.15	Retrieved temperature profile and the statistical uncertainty using the OEM from the RALMO measurements on 05 July 2011 with a cirrus cloud at 6 km height, in comparison with the coincident sonde measurements . . . . .	73
2.16	The OEM retrieved particle extinction, calculated Backscatter coefficient, and estimated lidar ratio profiles from the RALMO measurements on 05 July 2011 with of a cirrus cloud present at 6 km height . . . . .	74

2.17	Count rate for 30 min of coadded RALMO measurements from 2300 UT on 21 June 2011, which has a cloud base at an height about 4 km . . . . .	75
2.18	Difference between the forward model and the nighttime RALMO measurements on 21 June 2011 with the presence of a lower level cloud . . . . .	76
2.19	Averaging kernels and vertical resolution for temperature retrievals from measurements on 21 June 2011 with the presence of lower level cloud . . . . .	77
2.20	Retrieved temperature profile and the statistical uncertainty using the OEM from the RALMO measurements on 21 June 2011 with the presence of a lower level cloud, in comparison with the coincident sonde measurements . . . . .	77
2.21	The OEM retrieved particle extinction, calculated Backscatter coefficient, and estimated lidar ratio profiles from the RALMO measurements on 21 June 2011 with the presence of a lower level cloud . . . . .	78
3.1	Dates and times of the sondes were launched from Payerne, Switzerland that coincide with the ERA5 reanalysis data used to estimate the correlations of the temperature and relative humidity . . . . .	94
3.2	Temperature and relative humidity biases of the ERA5 reanalysis data . . . . .	95
3.3	Covariance and correlation matrices for temperature and relative humidity . . . . .	95
3.4	The OEM-retrieved temperature and relative humidity profiles, averaging kernels and uncertainty budgets from RALMO measurements on 28 August 2012 with 30 min temporal and 90 m vertical resolutions . . . . .	99
3.5	The OEM-retrieved temperature and relative humidity profiles, averaging kernels and uncertainty budgets from RALMO measurements on 10 September 2011 with 30 min temporal and 90 m vertical resolutions . . . . .	102
3.6	Nighttime temperature and relative humidity differences between bias corrected ERA5, RALMO, and ERA5-reRH, with sonde measurements for 14 nights . . . . .	104
3.7	Nighttime temperature biases of ERA5-Sonde, RALMO-Sonde, and ERA5-reRH-Sonde . . . . .	105
3.8	Daytime temperature and relative humidity differences between bias corrected ERA5, RALMO, and ERA5-reRH, with sonde measurements for 6 days . . . . .	108

3.9	Daytime temperature biases of ERA5-Sonde, RALMO-Sonde, and ERA5-reRH -Sonde . . . . .	109
4.1	Thickness of the 44 individual ISS layers detected over Payerne, Switzerland with 90 m vertical resolution. . . . .	122
4.2	The frequency of ISS layers that occur at temperatures below freezing point . .	124
4.3	Number of ISS layers occur at every 10 degree temperature ranges below freez- ing temperature. . . . .	124
4.4	ISS layers presented at different height layers in the atmosphere . . . . .	125
4.5	Number of ISS layers presence at clear and cloudy sky conditions . . . . .	126
4.6	Elastic backscatter signal, temperature, and relative humidity retrievals from the RALMO measurements made on 01 September 2011 from 1900-2200 UT .	127
A.1	Jacobians of the three different forward models . . . . .	140
C.1	Count rate for 15 min of RALMO measurements from 2200 UT on 09 Septem- ber 2011, a clear night . . . . .	146
C.2	Residuals between the forward model and clear nighttime RALMO measure- ments on 09 September 2011 . . . . .	147
C.3	Averaging kernels for temperature and relative humidity retrievals from the clear nighttime RALMO measurements on 09 September 2011 . . . . .	147
C.4	Vertical resolutions for temperature and relative humidity retrievals from the clear nighttime RALMO measurements on 09 September 2011 . . . . .	148
C.5	The OEM retrieved temperature and relative humidity in comparison with co- incident sonde measurements and <i>a priori</i> profiles from the clear nighttime RALMO measurements on 09 September 2011 . . . . .	149
C.6	Full uncertainty budgets for temperature and relative humidity retrievals from the clear nighttime RALMO measurements on 09 September 2011 . . . . .	150
C.7	Count rate for 15 min of RALMO measurements from 1000 UT on 10 Septem- ber 2011, a cloudy daytime . . . . .	151

C.8	Averaging kernels for temperature and relative humidity retrievals from the cloudy daytime RALMO measurements on 10 September 2011 . . . . .	152
C.9	The OEM retrieved temperature and relative humidity in comparison with coincident sonde measurements and <i>a priori</i> profiles from the cloudy daytime RALMO measurements on 10 September 2011 . . . . .	153
C.10	The OEM retrieved temperature profiles from RALMO measurements on 02 August 2011, from 1000-2359 UT with 15 min temporal and 90 m vertical resolutions. . . . .	154
C.11	The OEM retrieved particle extinction, RALMO back-scatter coefficient profile and estimated lidar ratio profile for 02 August 2011 . . . . .	155
D.1	Bias in the ERA5 relative humidity over ice measurements to the observational measurements from sonde . . . . .	158
D.2	Correlation and covariance of the bias corrected ERA5 relative humidity over ice	159
D.3	ERA5-reIce retrieved $RH_i$ from the RALMO measurements made from June-November, 2011 . . . . .	159

# List of Tables

1.1	Composition of Earth’s atmosphere . . . . .	6
1.2	Detected quantum lines from nitrogen and oxygen in PRR spectrum. . . . .	31
2.1	Return PRR wavelengths detected by the RALMO and the respective quantum lines from nitrogen and oxygen PRR spectrums. . . . .	56
2.2	Values and associated uncertainties for the OEM retrieval and forward model parameters. . . . .	57
2.3	Details of the 4 cases in different sky conditions we present to demonstrate the flexibility of our OEM temperature retrieval. . . . .	61
3.1	Values and associated uncertainties for the retrieval and forward model param- eters. . . . .	97
4.1	Dates and hours of measurements used in the ISS study . . . . .	121

# Chapter 1

## Introduction

### 1.1 Overview

Earth's climate is changing with time and in recent years, interest in better understanding the factors affecting the climate has increased. Only over the last two centuries has measuring the Earth atmospheric parameters such as temperature and humidity, and monitoring the changes of those parameters have become a scientific interest. Since the late 18<sup>th</sup> century, temperature measurements made by thermometers and other surface instruments have been available (Riehl et al., 1972). Balloon-borne sounding (radiosonde) measurements in the free atmosphere began after the Second World War (Riehl et al., 1972). Since the 1960s satellites have also been employed to measure and monitor the Earth's atmospheric parameters and the climate (Thies and Bendix, 2011). In the recent past, most atmospheric measurements were used primarily for weather forecasting purposes. Accuracy of the measurements plays a key role in both climate and weather predictions (Chahine, 1992; Palmer, 2000). Hence, improving measuring instrumentation and data analysis techniques has become a major interest in the scientific community.

Atmospheric water vapor is a fundamental parameter in the Earth's climate system and it is also with atmospheric temperature. Water vapor is known as the most significant greenhouse gas and plays a key role in thermodynamic and radiative processes in the atmosphere as well as in many other atmospheric processes (Wallace and Hobbs, 2006; Marshall and Plumb, 1989). The amount of water vapor is high in the atmospheric regions where the temperature is high.



High concentrations of water vapor in the atmosphere increase the absorption of long wave radiation, inducing warmer climate. The positive feedback of water vapor is by far the strongest feedback acting in the atmosphere (Held and Soden, 2000). Understanding the distribution and variability of the water vapor in the atmosphere along with the temperature variations allows for better understanding of the Earth's weather and climate systems.

The global radiosonde network provides most of the temperature and relative humidity information required for the forecast models. Even though there are thousands of stations where radiosondes are launched, the temporal resolution of the routine sonde measurements is rather low, with typically two radiosondes per day (Durre et al., 2006). Typically, radiosondes can take measurements up to about 30 km (a pressure altitude of about 11 hPa). However, it is also well known that the radiosonde relative humidity measurements are often not reliable in the upper troposphere (Leiterer et al., 1997; Nagel et al., 2001; Miloshevich et al., 2001; Noh et al., 2016; Ferreira et al., 2019). Among several other techniques available for improved water vapor measurements and temperature measurements, such as satellites and microwave radiometers, Raman lidar has become one of the potential tools that can provide water vapor and temperature measurements throughout the troposphere with high vertical and temporal resolutions (Whiteman et al., 1992; V  r  mes et al., 2016; Zuev et al., 2017). The Raman lidar technique uses the weak inelastic scattering of light by atmospheric water vapor, nitrogen and oxygen molecules. A typical Raman lidar system either measures temperature or water vapor as a function of height. Water vapor measurements from Raman lidars use the frequency-shifted backscattered radiation due to the excitation of the vibrational energy of the nitrogen and water vapor molecules to measure a mixing ratio (ratio of the number of water vapor molecules relative to the dry air molecules). However, for Raman lidar temperature measurements, the frequency shifted backscattered radiation, due to the rotational energies of the nitrogen and oxygen is considered. Among the three lidar techniques for temperature profiling (Rotational Raman, Rayleigh, and resonance fluorescence), Rotational Raman (RR) lidar has become the most efficient remote sensing technique for temperature profiling from the ground to the upper stratosphere. At lower altitudes, Mie scattering on aerosols prevents the use of the Rayleigh lidar method for temperature measurements (Alpers et al., 2004). Therefore, the RR spectra of atmospheric molecules are best to be used to measure the lower atmospheric temperatures. Par-

ticle extinction measurements are also possible with both vibrational-rotational and rotational Raman-scattered lidar signals. One can combine Raman lidar water vapor mixing ratio and temperature measurements to obtain a vertical profile of relative humidity. Measuring water vapor content in terms of relative humidity, where relative humidity is defined as the relation between the amount of water vapor present and the maximum amount that is physically possible at a given temperature, is important as it not only provides a measure of humidity but also a measure of temperature. Also, for temperatures below 0°C, relative humidity can be measured relative to water or ice. The necessity of measuring relative humidity over ice is crucial for atmospheric temperatures below -38°C as often liquid water does not exist beyond that temperature (except as super-cooled water). However, making direct measurements of water vapor pressure relative to ice are challenging. Using a variety of mathematical extrapolations such as Goff Gratch equation (List, 1984), Hyland and Wexler (Hyland, 1983), Magnus Teten (Murray, 1966), one can convert saturated vapor pressure measurements made with respect to water into saturated vapor pressure relative to ice. Hence, an estimation of relative humidity over ice ( $RH_i$ ) is possible in the atmospheric regions where the temperatures reach beyond -38°C. The  $RH_i$  measurements are important in the upper tropospheric region to study ice supersaturation (ISS) and formation of the cirrus clouds.

Generally, atmospheric humidity measurements are scientifically challenging to obtain due to their high variability. Thus, detecting ISS is difficult with lack of  $RH_i$  measurements. Various aircraft-based studies have shown the existence of frequent ISS in the UT (Krämer et al., 2009; Jensen et al., 2001; Gierens et al., 2000). An aircraft-based study by Jensen et al. (2005) has made extreme supersaturation measurements where  $RH_i$  reached up to 230% in clear sky conditions. Another study by Popp et al. (2007) using aircraft-based measurements showed high ISS of 230-250% in cloudy conditions. Due to the limited number of observations and their constrained temporal and spatial resolutions, it is difficult to understand the accuracy of  $RH_i$  measurements including the extreme observations made by aircraft. The main question that arises about the extreme observations is whether they are due to instrument artifacts or lack of knowledge of the physics of the ISS (Peter et al., 2006). In comparison to radiosonde and aircraft-based  $RH_i$  measurements, geostationary satellites provide a better set of global water vapor measurements. However, the spatial resolution of the satellite measurements is

poor compared to the other instruments. Even though traditional Raman lidar techniques do not measure direct  $RH_i$ , one can use the Raman lidar water vapor mixing ratio measurements together with ancillary temperature measurements to calculate  $RH_i$ . In addition to high temporal and spatial resolutions of the Raman lidar measurements, the measurements are made from a single ground-based location. Hence, using Raman lidar water vapor measurements has advantages for studying the climate impact of water vapor, supersaturation, and cloud formation.

The first known study of atmospheric ISS using Raman lidar measurements was made by Comstock et al. (2004). A year's worth of nighttime Raman lidar water vapor mixing ratio measurements calibrated against microwave radiometer water vapor measurements was used with radiosonde temperature measurements to estimate  $RH_i$ . That study focused on the frequency of high ISS in cirrus clouds. The results indicated that  $RH_i > 120\%$  frequently occurs at temperatures above  $-70^\circ\text{C}$ . The study by Comstock et al. (2004) does not provide for the uncertainty of the calculated  $RH_i$  measurements. A study by Immler et al. (2008) also used a combination of Raman lidar water vapor measurements with radiosonde temperature measurements to investigate cirrus, contrails, and ice supersaturated regions in high pressure systems at northern mid-latitudes. Raman lidar measurements made from August to September in 2000, in clear sky conditions (without low and mid-level clouds) were used to estimate  $RH_i$ . The results showed that the occurrence of cirrus and ISS are closely related. They observed frequent ice supersaturated regions in the uppermost troposphere (8 km to tropopause). Further investigations of optical depths, cirrus cloud classification, and contrails were also presented by Immler et al. (2008). Even though studies of ISS made using Raman lidar measurements are available, no study can be found that provides a direct retrieval of  $RH_i$  with retrieval uncertainties.

For the first time I present the application of the Optimal Estimation Method (OEM) mainly to retrieve temperature and relative humidity over water, and, as well, as over ice from the Raman lidar measurements. The OEM is an inverse method that has shown the potential to retrieve atmospheric aerosol, water vapor mixing ratio, ozone, and atmospheric temperature using lidar measurements (Povey et al., 2014; Sica and Haefele, 2016, 2015; Farhani et al., 2018; Mahagammulla Gamage et al., 2019). The OEM requires minimization of a cost function that measures the degree of fit of estimates of the atmospheric state to the measurements and to the *a priori* information (Palmer et al., 2000). In my work I have shown that the OEM

has several advantages over the traditional Raman lidar algorithms used to calculate temperature and water vapor mixing ratios. One of the important advantages of the OEM is that it can retrieve multiple other parameters such as overlap, particle extinction, lidar constant etc directly from the raw lidar measurements. The OEM also provides a full uncertainty budget including both random and systematic uncertainties on a profile-by-profile basis. The OEM provides estimates of retrievals of the atmospheric parameters such as temperature, humidity, and particle extinction with an estimate of a full uncertainty budget. Using the OEM-retrieved measurements and their uncertainty estimates in the weather and climate forecasts will provide better predictions with a well-defined uncertainty. The OEM-retrieved overlap functions, dead times, and lidar constants allow better understanding of the instrument used.

I applied an OEM analysis in three projects, all involving existing Raman lidar measurements from the Raman lidar measurements from the Meteoswiss/EPFL Raman lidar for Meteorological Observations (RALMO), located in Payerne, Switzerland. The OEM-retrieved temperature from the RALMO measurements in different sky conditions such as clear daytime, clear nighttime, low level clouds/aerosol, and cirrus cloud showed good agreement with the coincident sonde measurements and as well as better results than the traditional Raman temperature measurements especially in cloudy conditions. The relative humidity retrievals using the OEM use measurements from eight Raman channels in RALMO and allow direct retrievals of relative humidity. I further studied the time series of relative humidity over water and validated my results with coincident sonde measurements. The relative humidity over ice retrievals from the RALMO measurements were later used to investigate the frequency of ice super saturated event occurring above Payerne. Thus, I have successfully shown the OEM to retrieve relative humidity over water or ice from the Raman lidar measurements without the need of separate temperature and mixing ratio calculations.

Chapter 1 focuses on the Earth's atmosphere and atmospheric parameter measuring techniques. The first section of Chapter 1 gives a brief introduction to the Earth's atmosphere, temperature structure, and atmospheric humidity. Section 1.3 gives a brief introduction to lidars, atmospheric scattering and the lidar equation. Sections 1.4 and 1.5 introduce the RALMO lidar system and the traditional Raman lidar algorithm used to calculate relative humidity and temperature profiles. The Optimal Estimation Method (OEM) and some advantages of using it

over the traditional Raman lidar algorithms are given in Section 1.6.

The focus of Chapters 2 and 3 is to introduce and to describe the implementation of the Optimal Estimation Method (OEM) with Raman backscatter measurements to retrieve of temperature, relative humidity over water. Chapter 2 gives the results of the OEM-retrieved temperatures in four sets of conditions. Chapter 3 presents the assimilation of the set of Raman lidar measurements into European Centre for Medium-Range Weather Forecast Reanalysis (ERA5) data set. Chapter 4 of this thesis focuses on implementing the OEM scheme to retrieve relative humidity over ice to determine the occurrence of supersaturation events over Payerne, Switzerland. Chapter 5 summarizes and gives conclusions from all three projects and ideas for future work.

## 1.2 Introduction to Earth's atmosphere

Our Earth is enveloped by a relatively thin gaseous layer called the atmosphere, extending several thousands of kilometers above Earth's surface. Approximately 99% of the mass of the atmosphere is concentrated in the first 30 km from the surface (Wallace and Hobbs, 2006). In the early 1800s, John Dalton was able to recognize that the atmosphere is composed of several chemically distinct gases such as nitrogen and oxygen. He also determined the relative amounts of each gas found within the lower atmosphere. Later, in the 1920s, with the development of the spectrometer scientists were able to discover the atmospheric gases such as ozone and carbon dioxide that are very low in concentration. Table 1.1 shows the atmospheric gases and their composition (Wallace and Hobbs, 2006).

Table 1.1: Composition of Earth's atmosphere

Constituent	Molecular weight	Content (fraction of the total molecules)
Nitrogen (N <sub>2</sub> )	20.016	0.7808 (75.51% by mass)
Oxygen (O <sub>2</sub> )	32.00	0.2095 (23.14% by mass)
Argon (A)	39.94	0.0093 (1.28% by mass)
Water vapor (H <sub>2</sub> O)	18.02	0.5 (% by volume)
Carbon dioxide (CO <sub>2</sub> )	44.01	325 ppm
Neon (Ne)	20.18	18 ppm
Helium (He)	4.00	5 ppm
Krypton (Kr)	83.7	1 ppm
Hydrogen (H)	2.02	0.5 ppm
Ozone (O <sub>3</sub> )	48.00	0-12 ppm

A significant amount of atmospheric oxygen is produced by the photosynthesis reactions. Due to a variety of chemical reactions the oxygen in the atmosphere leads to the formation of an ozone ( $O_3$ ) layer in the upper atmosphere. Ozone filters the incoming solar radiation in the ultraviolet region (Wallace and Hobbs, 2006). The nitrogenous compounds from the metabolism of living organisms are returned to the atmosphere as nitrogen. Even though nitrogen is the most important component for all living beings, as DNA, RNA, and other proteins are made up of nitrogen, the atmospheric nitrogen is not in the usable form for most living beings. Lightning converts the atmospheric nitrogen to the usable molecules for life such as nitrate ions ( $NO_3^-$ ), ammonia ( $NH_3$ ), and urea [ $(NH_2)_2CO$ ] (Wallace and Hobbs, 2006). Two of the minor constituents of the atmosphere, water vapor and carbon dioxide play a key role in controlling the warming of the atmosphere. Atmospheric water vapor is the dominant greenhouse gas. It is highly variable and typically takes up to 0.5% of the volume of the Earth's atmosphere (Marshall and Plumb, 1989). The main source of the atmospheric water vapor is the evaporation from the ocean's surface. As moist air rises through the atmosphere, the air gets cooler and the water vapor in the air parcel condenses to form clouds. The water vapor condenses into water droplets to form clouds when it has a particle to condense upon. Such particles are called condensation nuclei and dust, pollen, sea salt, and black carbon are a few examples. Clouds that contain moisture are transported around the globe due to air currents and eventually the moisture returns to the ground as precipitation. Precipitation can occur in different forms such as rain, snow and hail. Once the water reaches the ground, a portion of it may evaporate back into the atmosphere and rest of the water may penetrate through the surface and become groundwater. Groundwater can either seep into the oceans, rivers, and streams, or it can be released back into the atmosphere through transpiration. Also, water runs over land into streams and lakes which eventually forms major rivers and carries all the water into the ocean. Then, again the water will be evaporated from the oceans. This process is called the hydrologic cycle of the Earth. The hydrologic cycle acts as an energy transfer and storage medium for the Earth's climate system. Both water vapor and carbon dioxide in the atmosphere absorb and emit infrared wavelengths of the solar spectrum (Marshall and Plumb, 1989; Wallace and Hobbs, 2006). The carbon dioxide concentration in the atmosphere is controlled by the processes such as photosynthesis, respiration, and exchange between the ocean

and the atmosphere (Marshall and Plumb, 1989; Wallace and Hobbs, 2006). Carbon dioxide is the most important of Earth's long-lived greenhouse gases. Compared to water vapor, carbon dioxide absorbs less heat but it stays in the atmosphere much longer. Even though, carbon dioxide is less abundant and less powerful than water vapor on a molecule per molecule basis, it absorbs wavelengths of thermal energy that water vapor does not. Thus, carbon dioxide adds to the greenhouse effect in a unique way and increases in atmospheric carbon dioxide are responsible for about two-thirds of the total energy imbalance that is causing global temperature to rise (Lindsey, 2018).

Today human activities are altering key dynamic balances in the atmosphere by increasing greenhouse gas levels in the lower atmosphere. This leads to raise the Earth's surface temperature by increasing the amount of heat radiated from the atmosphere back to the ground. Thus, leads to changing the Earth's climate. Two of the most important atmospheric parameters that alter the Earth's climate, temperature and humidity, are discussed in the following sections.

### **1.2.1 Temperature structure**

Temperature is a key parameter of the state of the atmosphere that varies greatly both vertically and horizontally. As the vertical structure of the temperature is qualitatively similar everywhere, it is often used in characterizing and understanding the atmosphere. The vertical temperature structure mostly depends on atmospheric pressure, humidity, and the effects of solar radiation. The atmosphere can be divided into 4 main layers based on its temperature, as shown in Fig. 1.1. Starting from the surface the layers are troposphere, stratosphere, mesosphere, and thermosphere. For the purposes of this work I only consider the lowermost two regions of the atmosphere.

#### **Troposphere**

The troposphere is the lowest part of the atmosphere that contains about 75% of all of the atmospheric air molecules including almost all of the water vapor and dust particles in Earth's atmosphere (Riehl et al., 1972). Thus, most of the weather such as clouds, rain, snow, wind, heat occurs in the troposphere. In this part of the atmosphere the temperature decreases with

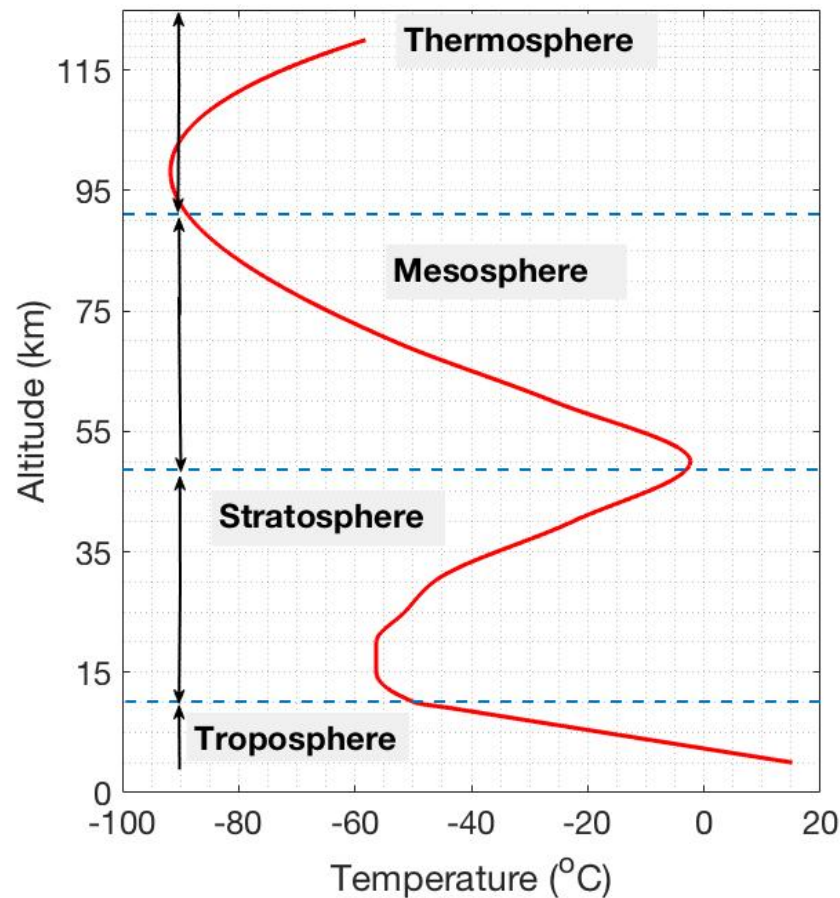


Figure 1.1: An atmospheric temperature profile based on the US Standard Atmosphere values showing the regions of the atmosphere based on temperature structure.

height. The troposphere is warmest near Earth's surface as it is heated from the below and coldest at its top, where it meets up with the layer above (the stratosphere) at a boundary region called the tropopause. The tropopause is lowest in the poles, where it is 7-10 km above the Earth's surface and it is highest (17-18 km) near the equator. The tropopause is also an inversion layer, where the air temperature starts to decrease with height. The inversion layers hold warmer air above cooler air, allowing only a little mixing between the troposphere and the stratosphere. The solar radiation that streams through the atmosphere heats the Earth's surface, and the surface radiates the heat back into the atmosphere warming the air molecules near the surface. Due to the pressure difference the warm air moves upwards and starts to expand. When air expands it cools. So, air higher up in the troposphere is cooler than the air lower down. The



cooler air tends to form clouds, rain, and snow. This continuous process of the rising of warm air and sinking or condensation of cool air makes the troposphere a layer with well mixed air (Riehl et al., 1972; Wallace and Hobbs, 2006; Marshall and Plumb, 1989).

The concept "standard atmosphere" is used by atmospheric scientists to describe an average atmosphere with no variations caused by weather, latitude, season and so on. In the standard atmosphere model, the bottom of the atmosphere is considered to be at sea level and the temperature at the sea level is  $15^{\circ}\text{C}$ . The temperature at the top of the troposphere is given as  $-57^{\circ}\text{C}$ . The rate at which the temperature changes with the height is called the lapse rate and in the troposphere the lapse rate is about 6.5 degrees per kilometer (Atmosphere, 1976).

### **Stratosphere**

The stratosphere extends upwards from the tropopause to about 50 km. It contains most of the ozone in the atmosphere. The temperature in the stratosphere increases with height due to the absorption of the ultraviolet (UV) radiation by the ozone. This means the stratosphere has a negative lapse rate. The temperature inversion in the stratosphere suppresses convection and damps out the vertical motions of tropospheric air. Therefore, it is significantly less turbulent than the troposphere. Hence, almost all commercial airliners cruise inside the stratosphere. Stratospheric temperatures are also affected by the seasonal changes, reaching particularly low temperatures in winter (Riehl et al., 1972; Wallace and Hobbs, 2006).

There are strong interactions among radiative, dynamical, and chemical processes, in the stratosphere. This leads to more rapid horizontal mixing of gaseous components as compared to the vertical mixing in the stratosphere. Due to the complexity of the stratospheric circulations the existing global climate models and stratospheric chemistry-climate models fail to simulate or produce temperature trends or temperature profiles that match with the observations in the stratosphere (Solomon et al., 2010). Therefore, measuring accurate stratospheric temperatures is of particular interest in the climate science community.

### **Upper troposphere - lower stratosphere (UTLS)**

The upper troposphere-lower stratosphere (UTLS) plays a key role in radiative forcing and chemistry-climate coupling of the Earth's atmosphere. The UTLS is the region between the

upper troposphere and the lower stratosphere, from roughly 5 to 22 km in height. The coupling between dynamics, chemistry, and radiation is found to be strong in the UTLS. The understanding of the chemical and dynamical behavior of the tropopause region, and its long-term variability has become one of the main interests in atmospheric science. The primary motivation is to understand the coupled processes in UTLS to identify their role in climate change, which is a necessity for improving weather and climate model simulations of this region (Wallace and Hobbs, 2006).

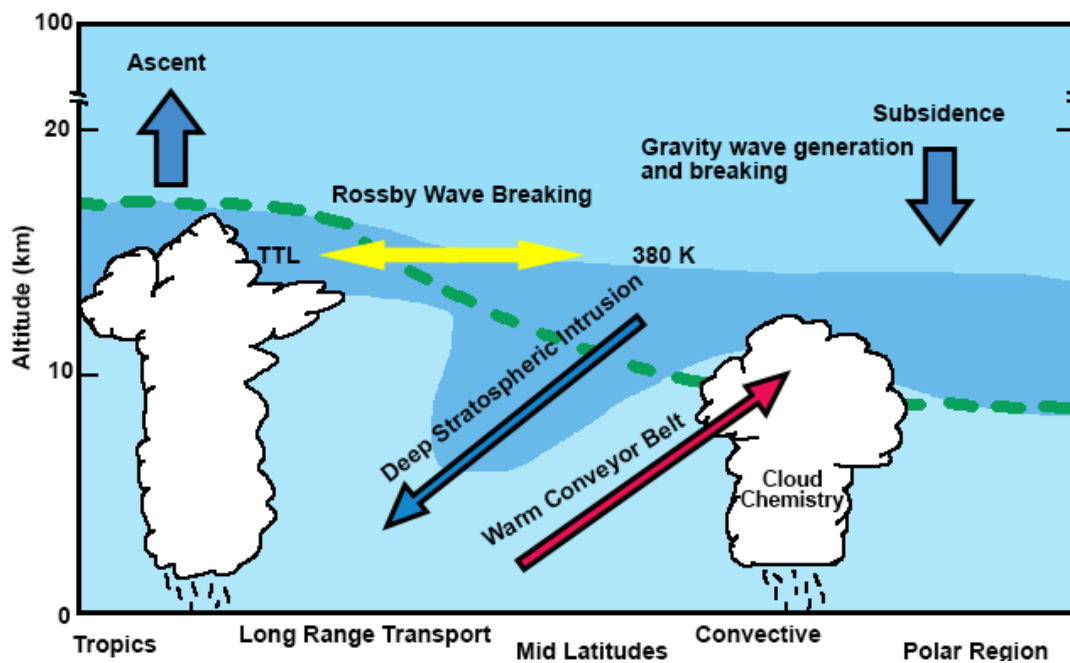


Figure 1.2: This figure shows the important processes such as coupling dynamics, chemistry, and cloud microphysics in the UTLS region. The dashed green line represents the time average tropopause. In the tropics, maximum outflow from deep convection occurs near 12-14 km, while the cold point tropopause occurs near 17 km. The intervening region has characteristics intermediate between the troposphere and stratosphere, and is termed the tropical transition layer (TTL). Extratropical stratosphere-troposphere exchange occurs in tropopause folds and intrusions linked with synoptic weather systems; these events transport stratospheric air high in ozone into the troposphere. Transport above the subtropical jet couples the TTL and the extratropical lower stratosphere. In addition, synoptic scale uplift and deep convection brings near-surface emissions of gases and particles into the upper troposphere, where they can strongly influence global-scale chemistry. Gravity wave generation and breaking also contribute to the mixing of chemical constituents in the UTLS. (Adapted from Pan et al. (2010); Stohl et al. (2003)).

The UTLS processes depend crucially on the distribution of greenhouse gases (GHGs), such as ozone and water vapor, as well as aerosols and clouds. The Fig. 1.2 illustrates some features of the UTLS region.

Each of the three projects for this thesis provides a means to more accurately measure temperature and relative humidity in the UTLS, and contributes to our understanding of the important coupling processes in the UTLS region.

### 1.2.2 Atmospheric humidity

Water can be found in all three phases in Earth's atmosphere - solid, liquid, and vapor. The release of latent heat of water vapor when it condenses into the liquid or solid phase, together with the Earth's rotation, drives the large-scale circulation of the atmosphere. Water vapor is the most dominant greenhouse gas in the Earth's atmosphere that is highly variable and poorly understood. Even relatively small changes in the atmospheric water vapor can play a significant role in the Earth's atmospheric energy balance (Wallace and Hobbs, 2006).

The amount of the water vapor in the atmosphere can be expressed in many different ways. In this section the various measures of variable atmospheric water vapor content are presented.

#### Mixing ratio

The mixing ratio is the ratio of mass of the water vapor  $m_{wv}$  to the mass of the dry air  $m_d$  in a certain volume of air and is expressed as

$$w \equiv \frac{m_{wv}}{m_d}. \quad (1.1)$$

The mixing ratio is often expressed in the units of grams of water vapor to kilograms of the dry air.

### Specific humidity

The specific humidity,  $q$  is a measure of the mass of the water vapor with respect to the mass of air per unit volume defined as,

$$q = \frac{\rho_{wv}}{\rho}, \quad (1.2)$$

where  $\rho = \rho_d + \rho_{wv}$  is the total mass of the air (dry air plus water vapor) per unit volume. The specific humidity is conserved in the absence of mixing or of condensation. In other words, specific humidity does not vary as the temperature or pressure of an air parcel changes if the moisture in the air parcel is not removed or added. The stability of the specific humidity makes it a useful parameter to identify the properties of a moving air mass.

### Vapor pressure

The vapor pressure is the partial pressure of the atmospheric water vapor. Vapor pressure can be expressed as,

$$e = \frac{w}{0.622 + w} P \quad (1.3)$$

where  $P$  is the total air pressure, assuming water vapor behaves as an ideal gas.

### Saturated vapor pressure

The saturation vapor pressure is the partial pressure of the water vapor in equilibrium with a plane surface of pure water. It solely depends on the temperature. If the temperature is increased, the molecules will have more internal energy and will vibrate faster. Thus, more molecules will be able to break free and evaporate from the liquid surface. As a result, the vapor pressure to increase. For the temperatures below  $-38^\circ\text{C}$  water will no longer be found in the liquid form. Therefore, the saturation will occur over the plane surface of ice instead of water (Gierens et al., 2000). As the bonds between adjacent molecules are stronger in an ice surface than they are in a liquid surface, at the same temperature fewer molecules will escape from an ice surface than from a liquid surface. Therefore, the saturation vapor pressure over ice is lower than the saturation vapor pressure over liquid water (shown in Fig. 1.3).

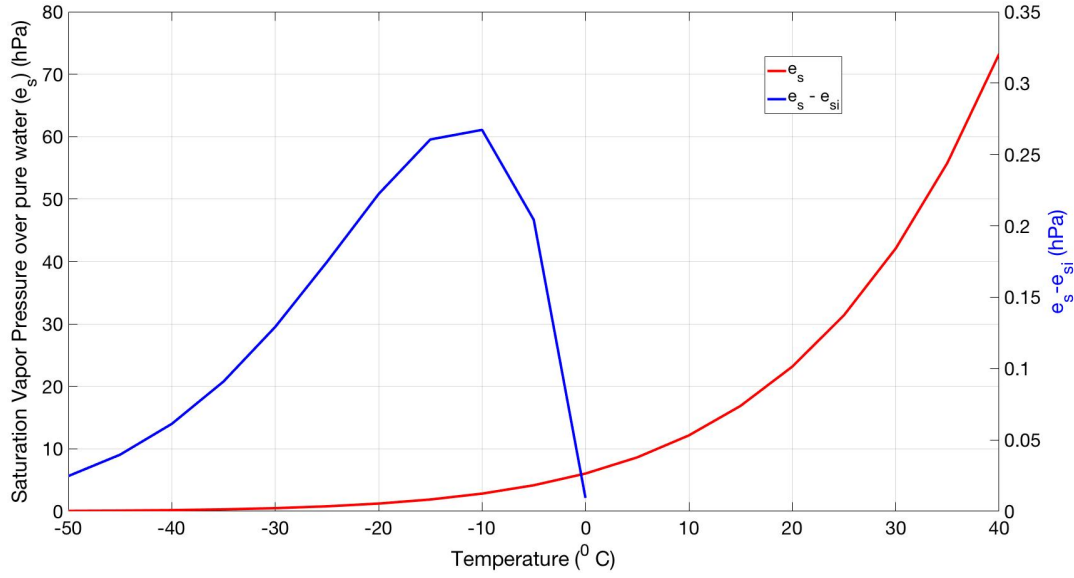


Figure 1.3: Variation with temperature of the saturation vapor pressure  $e_s$  over a plane surface of pure water (red line) and the difference between  $e_s$  and the saturation vapor pressure over a plane surface of ice  $e_{s,i}$  (blue line) (Adapted from Wallace and Hobbs (2006)).

There are several existing formulations that allow the estimation of saturation vapor pressure over both water and ice. For this thesis, I use the saturation vapor pressure equations from (Hyland, 1983).

The saturation vapor pressure over liquid water ( $e_{s,w}$ ) for temperatures below  $0^\circ$  is given by:

$$\begin{aligned} \log e_{s,w} = & \frac{-0.58002206 \times 10^4}{T} + 0.13914993 \times 10^1 - 0.48640239 \times 10^{-1} \times T \\ & + 0.41764768 \times 10^{-4} \times T^2 - 0.14452093 \times 10^{-7} \times T^3 \\ & + 0.65459673 \times 10^1 \times \log(T). \end{aligned} \quad (1.4)$$

The saturation vapor pressure over liquid ice ( $e_{s,i}$ ) for ice is given by (Hyland, 1983):

$$\begin{aligned} \log e_{s,i} = & \frac{-0.56745359 \times 10^4}{T} + 0.63925247 \times 10^1 - 0.96778430 \times 10^{-2} \times T \\ & + 0.62215701 \times 10^{-6} \times T^2 + 0.20747825 \times 10^{-8} \times T^3 \\ & - 0.94840240 \times 10^{-12} \times T^4 + 0.41635019 \times 10^1 \times \log(T). \end{aligned} \quad (1.5)$$

### Relative humidity

Relative humidity is expressed as a ratio of the vapor pressure to the saturation vapor pressure. Thus, relative humidity can be estimated over water or ice. The general formulation of the relative humidity is

$$RH = \frac{e}{e_s} \times 100. \quad (1.6)$$

where  $e_s$  is the saturated vapor pressure over water or ice. Relative humidity is the most popular scale in both meteorology and atmospheric science as it has certain advantages over absolute measures of water vapor concentration. It is usually in the range of 0% (completely dry) and 100% (saturation) where an absolute scale must be much wider because the concentration of water molecules decreases from the ground to the tropopause by roughly a factor of 10 000 (Gierens et al., 2012). However, there is nothing to forbid relative humidity from exceeding 100%. Such conditions where the relative humidity exceeds 100% are called supersaturation. Supersaturation is an unstable state if there is no condensed phase. Another advantage is that cloud formation is controlled by relative humidity, not by absolute humidity (Gierens et al., 2012).

### Ice supersaturation (ISS)

Ice supersaturation is a frequent atmospheric phenomenon that occurs mainly in the UTLS region where the temperature reaches below the freezing point. In the upper troposphere, temperatures easily go below  $-40^\circ\text{C}$  (Gierens et al., 2012). Thus, the super-cooled pure water droplets would freeze spontaneously. However, if the water droplets are not pure (contain aerosol particles, etc.) or are highly diluted they will not freeze spontaneously even at temperatures below  $-40^\circ\text{C}$ . In order for the droplets to freeze, they need to gain more water molecules from the ambient air. That requires ambient water vapor concentration corresponding to relative humidity with respect to ice to be more than 145%. Ice formation in these relatively high supersaturation levels are called ISS. ISS can last for hours or even days and the thickness of such formations can be from 100 m to 3-4 km. The horizontal extensions of the ISS layers are not well known yet. ISS is often found within cirrus clouds (ice clouds) and contrails, but also can be formed even in cloud free regions. ISS regions are known to have small effects on

atmospheric radiation flow (Gierens et al., 2012). However, it was found that as soon as thin cirrus clouds form within ISS regions, the radiation effects double by two orders of magnitude (Fusina et al., 2007). Thus, ISS plays an important role in cloud formation and as well as in Earth's weather and climate.

### 1.2.3 Instruments and techniques to measure relative humidity

As discussed in the Overview in Section 1.1, measurements from Raman lidars, typically ground-based, provide an excellent way to obtain vertical temperature and relative humidity profiles, and I use the Raman lidar technique in this thesis. Details will be discussed in the following sections. In this section I will provide a brief discussion on other humidity measuring instruments such as radiosondes, microwave radiometers, and weather satellites.

Radiosonde is the most common atmospheric humidity measuring technique and has been in use since the late 1930s (DuBois, 2002). Radiosondes play an important role in providing long-term high-quality time series of climatology trends of various parameters. There are about 1,300 radiosonde launch sites all over the globe and most countries share data with the rest of the world through international agreements. Almost all routine radiosondes are launched 45 minutes before the official observation times of 0000 UTC and 1200 UTC (Seidel et al., 2011). This provides an instantaneous snapshot of the atmosphere. The radiosonde is a small package consisting of multiple instruments including sensors that measure pressure, temperature, and relative humidity, and a GPS that returns position information. The radiosonde is suspended below a large balloon inflated with hydrogen or helium gas and as the balloon rises at about  $5 \text{ ms}^{-1}$  the radio transmitter sends the measurements obtained by the sensors to the ground station. Radiosondes can provide continuous and detailed profiles from the ground to altitudes of 30 km and above. The hygistor is the relative humidity sensor employed in the radiosondes located at a place where the outside air passes can reach. A hygistor consists of a glass slide or plastic strip covered with a moisture sensitive film of lithium chloride (LiCl) and a binder; metal strips are located along the edges. The electrical resistance of the LiCl changes with a change in the atmospheric humidity. The hygistor on most radiosondes is designed to record the ambient relative humidity with respect to water in the range from 15% to 100%.

Hygristors do not provide accurate measurements of relative humidity at temperatures below the freezing point. When the radiosonde passes through a cloud or a layer of ice, the humidity sensor can freeze and it will then either not provide any measurements or measure the humidity inaccurately. To prevent water condensing on the sensors during the ascent of the balloon, some radiosonde products now occupy two sensor elements that include heating of the of a sensor elements (Hopkins). However, the radiosonde technique is still not developed enough to measure relative humidity over ice in regions below freezing. In Chapters 3 and 4 of this thesis I use the radiosonde measured relative humidity profiles to compare the relative humidity retrieved from the Raman lidar measurements. For more details of the radiosonde humidity measurements and their accuracy refer to Miloshevich et al. (2001); Peixoto and Oort (1996); Sapucci et al. (2005); Bock et al. (2013), and Dirksen et al. (2014).

Weather satellites are outfitted with various types of sensors that can measure atmospheric water vapor from space (Jones et al., 2009). The operational meteorological observational systems use nadir sounding (down-looking) instruments to measure the tropospheric water vapor. Vertical resolution of the nadir microwave humidity sounders is typically several kilometers and the horizontal resolution is about 10-15 km at nadir. Nadir observations cover approximately the entire globe (mostly on polar orbits) and the observations are made in both day and nighttime. Some of the limitations of nadir observations are: limited altitude information from pressure broadening, only sensitive to region with largest water abundance (troposphere), and sensitivity to tropospheric clouds (large ice particles or water drops) (Urban, 2013). Satellites use the limb sounding technique to measure the trace amounts of water vapor in the upper troposphere and throughout the entire middle atmosphere with resolution of typically only a few kilometres. The limb observations in the troposphere and lowermost stratosphere are often limited by the water, cloud, or aerosol absorption depending on wavelength. Further details of satellite instruments measuring atmospheric water vapour and observation techniques are available in Urban (2013).

Microwave radiometer is another instrument that is used to measure atmospheric water vapor. It can be used as a ground-based instrument to measure humidity from the surface and it can also be used in satellites to measure humidity from space. Microwave radiometers measure thermal emissions from the downwelling brightness temperature in the atmosphere.



Most commercial microwave radiometers operate in the 2060 GHz frequency range, where the atmospheric thermal emissions are influenced by atmospheric temperature and humidity. A microwave radiometer consists of an antenna system, microwave radio-frequency components and a signal processing unit. When appropriate detection frequencies are used, the emission of microwave radiation from the atmospheric trace gases of liquid water and of ice crystals can be measured. The emissivity of the substances and their radiative temperature depends on the substance concentration, pressure, and temperature in the atmosphere. Thus, each substance can be estimated by measuring its radiative temperature in appropriate frequency bands. The absorption band of water vapour is located between the frequencies of 20 and 30 GHz. Thus, measurements of the radiative temperatures in water vapor absorption band allow us to estimate the integrated water vapour (IWV) content and a profile of absolute humidity. More details of microwave radiometers are given in Löhnert et al. (2009); Hewison (2007), and Liljegren et al. (2005).

Raman lidars are another potential tool that provides atmospheric humidity measurements with high spatial and temporal resolution. In this thesis I use the Raman lidar measurements to retrieve relative humidity and details will be discussed in Section 1.3.

HATPRO Radiometers, dropsondes, and aircraft based instruments are a few other atmospheric humidity measuring techniques. Further details of these instruments and comparisons of measurements made by each instrument are available in Soden and Lanzante (1996), Wallace and Hobbs (2006), and Buehler et al. (2004).

## 1.3 Lidars and atmospheric measurements

### General description of a lidar

The lidar (light detection and ranging) technique is an active remote sensing technique. It is one of the tools that has the ability to measure the atmosphere at ambient conditions with high temporal and spatial resolution, and the potential of covering large altitude ranges in the atmosphere. A typical lidar system (shown in Fig. 1.4) consists of a transmitting system (laser) and a receiving system (telescope, optical analyzer, etc). Lidars emit light pulses into the atmosphere and the light is scattered by the atmospheric molecules in all directions. A portion of the light

that is scattered back is collected by the lidar telescope. Using the lidar's photon detecting system the amount of backscattered light can be measured as a function of altitude. Depending on the type of interaction processes of the emitted light with the atmospheric constituents, different atmospheric parameters and conditions such as temperature, humidity, ozone, particulates, cloud and wind can be detected (Weitkamp, 2006). To improve the precision of the detection

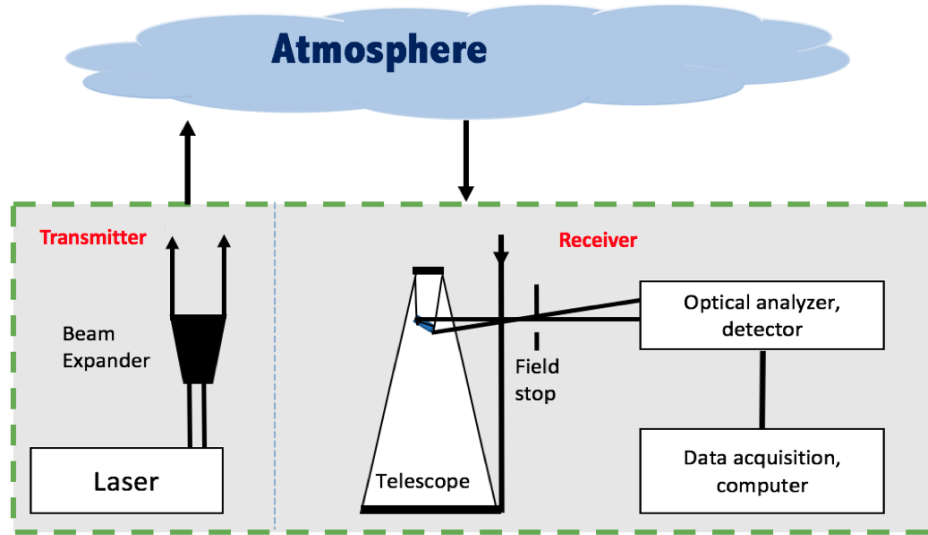


Figure 1.4: Schematic drawing of lidar system.

the lasers used in the lidars are monochromatic (i.e. one colour /one wavelength), collimated (i.e. very small divergence over large distances), intense (i.e. lots of energy in a small area) and polarized (i.e. energy is aligned in one direction). There are different types of lidars depending on the laser used and the type of scattering studied. Different types atmospheric scattering are discussed in the next section. The main 5 basic lidar techniques are, elastic-backscatter lidar, differential absorption lidar (DIAL), Raman lidar, fluorescence lidar, and Doppler lidar. For all three projects in this thesis I am interested in the measurements from Raman lidars and atmospheric Raman scattering processes.

As shown in Fig. 1.4, the lidar receiving system consists of a telescope that collects all the backscattered light, an optical analyzer where the signal is spectrally separated, amplified and transformed to an electrical signal, and a data acquisition system where the signal is digitized

and stored in a computer unit. The diameter of the telescope depends on the purpose of the lidar and it can vary from 0.1 m to a few meters. The optical analysis of the backscattered light is often done before the detection. In most cases interference filters are placed in front of the detector to allow only a certain pass-band around the wavelength of interest enter the detector (Kovalev and Eichinger, 2004).

Signal detection is typically realized with photomultiplier tubes (PMTs). PMTs are very sensitive and is used when the backscatter signal is weak ( $<10$  MHz). PMTs store only the number of photon counts per time interval after emission of the laser pulse. The efficiency of a PMT to measure and record pulses depends on the time taken up by all components of the signal processing. The number of photons detected by the PMTs can be measured using a photon counting system. There are two types of counting systems, nonparalyzable systems that require a fixed recovery time and paralyzable systems that don't. The fixed recovery time of a counting system is called the dead time of the system. In a paralyzable system if the time gap between two signals reaching the detector is larger than dead time then the event will be recorded. Thus, the observed rate is equal to the rate at which time intervals occur that exceed dead time. The main difference between the two types of counting systems is that nonparalyzable detector systems are not effected if the signal is not processed whereas paralyzable detector systems are effected (Wandinger, 2005a).

In lidars for strong backscatter signals typically the measurements in lower altitudes are recorded using analog recorders where the average current produced by the photo pulses is measured. Then analog-to-digital signal conversion and digital signal processing need to be performed (Wandinger, 2005a).

### **Atmospheric scattering related to lidar**

Atmospheric scattering is the redirection of the original path of electromagnetic (EM) radiation (i.e. an incident light ray) traveling through the atmosphere due to particles or gas molecules present in the atmosphere. As the EM wave interacts with the atmospheric particles or gas molecules the electron's cloud within the atmospheric constituent starts to oscillate periodically with the same frequency as the incident EM wave. This oscillation later induces a dipole moment on the atmospheric constituent. The oscillating induced dipole moment creates a source

of EM radiation, thereby resulting in scattered light. The atmospheric scattering depends on the wavelength of the incident radiation and also on the size of the atmospheric constituent that scatters the radiation. In this section, I will briefly discuss three main atmospheric scattering processes: Mie scattering, Rayleigh scattering, and Raman scattering.

Rayleigh scattering is the most dominant scattering mechanism in the upper atmosphere and it is a type of an elastic scattering, meaning Rayleigh scattered radiation has the same frequency as the incident radiation. Rayleigh scattering occurs when particles are very small compared to the wavelength of the radiation. These could be particles such as small specks of dust or nitrogen and oxygen molecules (Kovalev and Eichinger, 2004).

The intensity of Rayleigh scattering varies inversely with the fourth power of the wavelength ( $\lambda^{-4}$ ). Rayleigh scattering is the phenomenon behind the sky appearing blue during the day and red/orange during sunrise and sunset. As sunlight travels through the atmosphere, the shorter wavelengths (i.e. blue) of the visible spectrum scatter more than the longer wavelengths. At sunrise and sunset, the light has to travel farther through the atmosphere than daytime. Hence, the scattering of the shorter wavelengths is more complete and this allows a greater proportion of the longer wavelengths of the visible spectrum to penetrate through the atmosphere (Kovalev and Eichinger, 2004).

For the atmosphere, assuming that the scattering is from a spherically symmetric molecule the total Rayleigh scattering cross-section can be approximated (Kovalev and Eichinger, 2004):

$$\sigma \simeq \frac{32\pi^3}{3\lambda^4} \left( \frac{n-1}{N} \right)^2 \quad (1.7)$$

where  $N$  is the number density of the molecules and  $n$  is the refractive index.

When using Rayleigh scattering for diagnostics usually the scattered light is collected over a limited solid angle. Therefore, it is useful to define a differential cross-section for Rayleigh scattering such that

$$\frac{\partial \sigma}{\partial \Omega} \simeq \frac{4\pi^2}{\lambda^4} \left( \frac{n-1}{N} \right)^2 \sin^2 \phi, \quad (1.8)$$

where  $\phi$  is the scattering angle and  $\Omega$  is the solid angle (Kovalev and Eichinger, 2004).

Mie scattering is an elastic scattering that occurs due to the particles that are about the size of the wavelength of the incident radiation. Atmospheric aerosols, smoke particles, water

vapor molecules, and other large atmospheric molecules cause Mie scattering. The Mie signal is proportional to the square of the particle diameter and is much stronger than Rayleigh scattering.

Raman scattering is an inelastic process where monochromatic light or photons is scattered off a molecule at a different wavelength than the incident. Thus, the scattered photons energy will be different than the incident photon. The change in energy is proportional to the rotational or vibrational energy levels of the atom or the molecule that the photon scattered off (Kovalev and Eichinger, 2004). As shown in Fig. 1.5, the scattered photon can either gain energy from

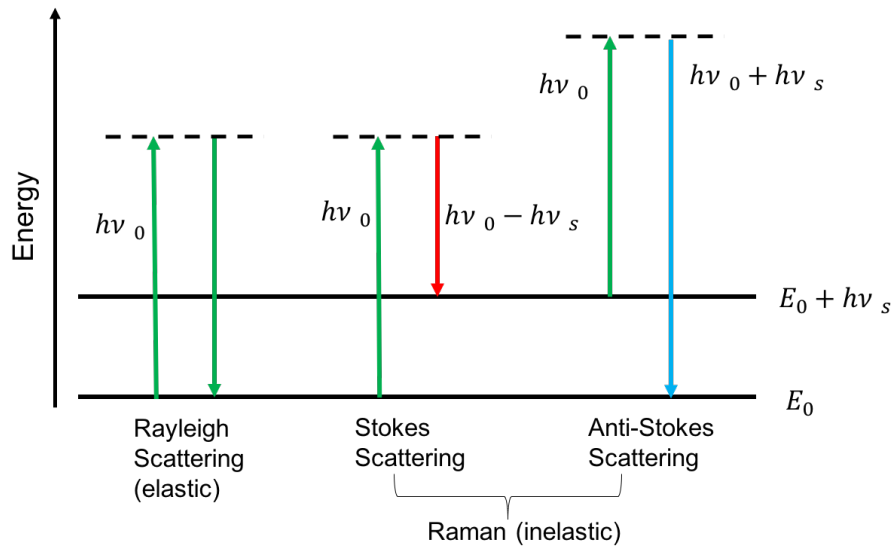


Figure 1.5: The different types of light scattering: Rayleigh scattering (no exchange of energy: incident and scattered photons have the same energy), Stokes Raman scattering (atom or molecule absorbs energy: scattered photon has less energy than the incident photon) and anti-Stokes Raman scattering (atom or molecule loses energy: scattered photon has more energy than the incident photon).

the interaction and shift to a higher frequency (red-shift) or lose energy from the interaction and shift to a lower frequency (blue-shift). The processes of gaining and losing energy are known as Stokes and anti-Stokes shift respectively. The frequency shift for the scattering molecules is given by,

$$\Delta\tilde{\nu} = \tilde{\nu}_i - \tilde{\nu}_s = \frac{\Delta E}{hc} \quad (1.9)$$

where  $\tilde{\nu}_i$  is the incident frequency,  $\tilde{\nu}_s$  is the scattering frequency,  $\Delta E$  is the energy difference

between the molecular energy levels involved,  $h$  is the Planck's constant, and  $c$  is the speed of light in a vacuum.

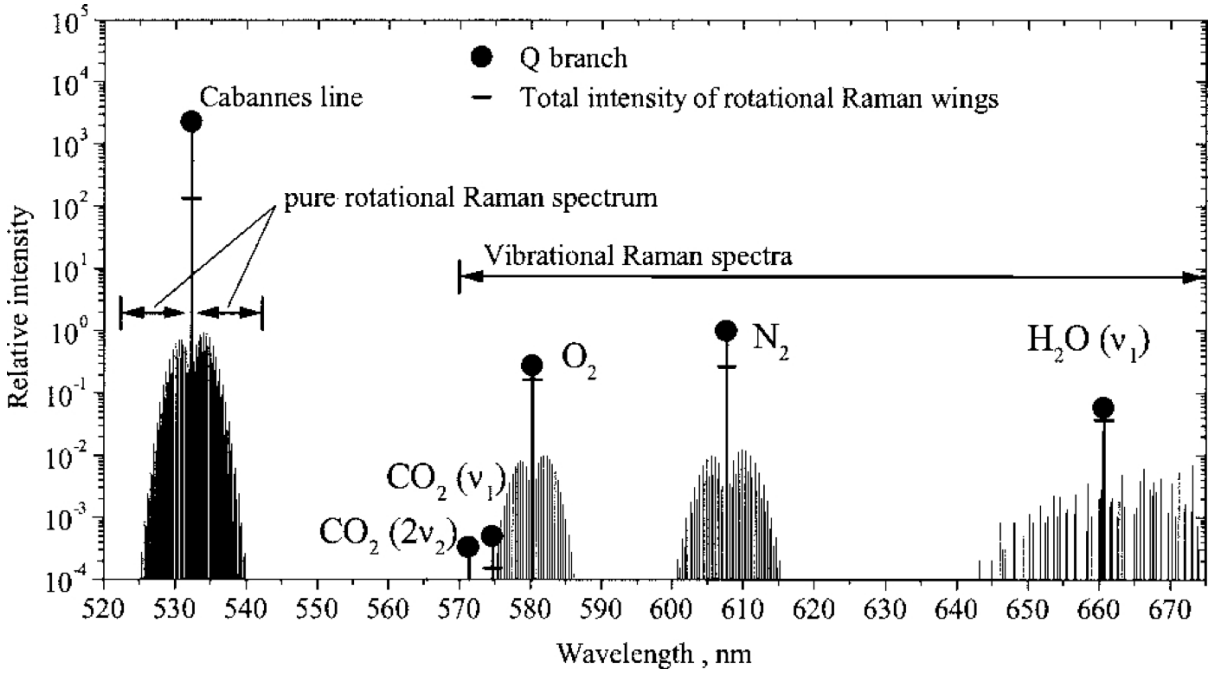


Figure 1.6: Raman backscatter spectrum of the atmosphere for an incident laser wavelength of 355 nm, normal pressure, a temperature of 300 K, an  $N_2$  and  $O_2$  content of 0.781 and 0.209, respectively, and a water-vapor mixing ratio of 10 g/kg. The curves for liquid water and ice are arbitrarily scaled. (Figure 9.2 in Wandinger (2005b))

Since the scattering targets of interest are mostly diatomic molecules in the atmosphere such as nitrogen and oxygen, I expect that there will be an energy shift due to both the vibrational and rotational transitions as shown in Fig. 1.6. The vibrational energy transition levels for diatomic molecules can be approximated using the model of a freely rotating harmonic oscillator:

$$E_{vib,v} = hc_0 \tilde{\nu}_{vib} (v + 1/2), \quad \text{for } v = 0, 1, 2, \dots \quad (1.10)$$

where  $\nu_{vib}$  is the specific vibrational wave-number of the molecule and  $v$  is the vibrational quantum number (Behrendt, 2005). For the energy levels for the rotational quantum numbers is approximated by,

$$E_{rot,J,v} = hc_0 [B_v J(J+1) - D_v J^2(J+1)^2], \quad \text{for } J = 0, 1, 2, \dots \quad (1.11)$$

where  $J$  is the rotational quantum number that belongs to each vibrational level. The specific rotational constant and centrifugal distortion or stretching constant of the molecule are  $B_v$  and  $D_v$  respectively. The two constants  $B_v$  and  $D_v$  depend on the actual vibrational state of the molecule. For more details of the constants refer to Weber (1979).

Equations 1.10 and 1.11 can be applied into Eq. 1.9 to find the frequency shift of a certain molecule due to Raman scattering. Thus, it is not possible to obtain every frequency transition in vibrational or rotational Raman scattering. The selection rules for vibrational and rotational transitions, are

$$\begin{aligned}\Delta v &= 0, \pm 1 \quad (\text{vibrational}) \\ \Delta J &= 0, \pm 2 \quad (\text{rotational}),\end{aligned}\tag{1.12}$$

where  $\Delta J$  is the difference between higher rotational quantum number to the lower quantum number, which is independent of the initial or the final quantum state of the molecule. The difference between the vibrational quantum numbers of the final and the initial states is  $\Delta v$  (Behrendt, 2005).

The intensity of the observed Raman line depends on the cross-section of the corresponding rotational scattering process. The Raman differential cross-section for backscattered (scattering angle is  $180^\circ$ ) photons is approximated by

$$\left(\frac{d\sigma}{d\Omega}\right)_{RR} = k_v \tilde{\nu}_1^4 \left(\frac{7}{60} \gamma^2\right)\tag{1.13}$$

where  $k_v = \frac{\pi^2}{\epsilon_0^2}$  with  $\epsilon_0$  is the permittivity of vacuum and  $\gamma$  is the anisotropy of the polarizability (Placzek, 1934). Simplified forms of Raman rotational differential cross-sections for Stokes and anti-Stokes branches are given in Section 1.3.

### Elastic scattering and Raman lidar equations

In the following sections I will use Raman lidar measurements to retrieve both atmospheric temperature and humidity. I will also use elastic lidar measurements to estimate the particle extinction in the atmosphere. In this section I will explain the general forms of lidar equation that explains the relationship between the number of collected backscattered photons to the atmospheric and instrumental parameters. The observed backscattered photocount rate at

particular altitude can be represented as:

$$N_{obs}(z) = N_0 \frac{ct}{2} \Lambda \eta \frac{O(z)}{z^2} \beta(z, \lambda) e^{-(\tau_{\lambda_0} + \tau_{\lambda})} + B(z), \quad (1.14)$$

where  $N_0$  is the average rate of photons emitted by the laser,  $B(z)$  is the background photon count rate,  $c$  is the speed of light,  $t$  is the temporal resolution of the laser,  $O(z)$  is the geometrical function,  $\Lambda$  is the area of the telescope,  $\eta$  is the detection efficiency of the lidar system,  $\beta(z, \lambda)$  is the volume backscatter cross-section for the particle that is scattering the radiation and the two terms  $\tau_{\lambda_0}$  and  $\tau_{\lambda_R}$  are scattering the atmospheric transmission.

The function  $O(z)$  shows that the geometrical effect of the laser emitting and receiving systems of the lidar. The geometric setup of the emitter and the receiving optics determines the capability of the detection of backscatter signals at distances close to the lidar. At short distances the laser beam can not be fully imaged onto the detector. Thus, only a portion of the scattered signal is received. The signals received by the lidar vary with the height and depends on the laser beam diameter, shape, and divergence, the receiver field of view, and the location of emitter and receiver optical axes relative to each other. This variation is represented by the function  $O(z)$  and its value is zero at the lidar and becomes unity when the laser beam is completely imaged onto the detector (Wandinger, 2005a).

For different lidar systems the background photon count rate either can be a constant or a function of altitude. The term  $\frac{ct}{2}$  together represents the effective pulse length of the laser or often referred to as the vertical resolution of the system. I define the lidar constant  $C$  as a product of all the system specific constants,

$$C = N_0 \frac{ct}{2} \Lambda \eta \quad (1.15)$$

The differential cross-section indicates the amount of radiation back-scattered scattered by a particular particle within the solid angle  $d\pi$  (Wandinger, 2005a).

$$\beta(z, \lambda) = \sum_j n_j(z) \frac{d\sigma_{j,sca}(\lambda, \pi)}{d\Omega}. \quad (1.16)$$

The volume backscatter cross-section for the particle that is scattering the radiation,  $\beta(z, \lambda)$



can be represented as the sum of the product of particle number density  $n_j(z)$  and the differential scattering cross-section of the particular particle  $\frac{d\sigma_{j,sca}(\lambda, \pi)}{d\Omega}$  at a particular altitude. The volume backscatter cross-section is measured in the SI unit of per meter per steradian.

The optical depth at the transmitted wavelength  $\tau_{\lambda_0}$  and at the return wavelength  $\tau_{\lambda_R}$  together express the fraction of laser radiation attenuated by the atmosphere that is shown in the term  $e^{-(\tau_{\lambda_0} + \tau_{\lambda_R})}$  in Eq. 1.14.

The general lidar equation given in Eq. 1.14 changes depending on the type of the scattering that each lidar system measures. The two main forms of the lidar equation are elastic and the Raman lidar scattering equations. In elastic scattering the emitted and the returned wavelengths are the same. Hence, the volume backscatter cross-section will only expanded into two terms,

$$\beta(z, \lambda) = \beta_{mol}(z, \lambda_0) + \beta_{aer}(z, \lambda_0) \quad (1.17)$$

where  $\beta_{mol}(z, \lambda_0)$ ,  $\beta_{aer}(z, \lambda_0)$  are the volume backscatter cross-section due to molecular scattering and due to aerosol scattering at the laser's wavelength  $\lambda_0$ , respectively.

The other term that changes in the lidar equation for elastic scattering is the atmospheric transmission. The two terms  $\tau_{\lambda_0}$  and  $\tau_{\lambda_R}$  are equal to each other as  $\lambda_0 = \lambda$ . Hence, the optical depth terms for elastic transmission take the simple form:

$$\tau_{\lambda_0} + \tau_{\lambda_R} = 2 \int_0^z [\alpha_{mol}(r, \lambda_0) + \alpha_{aer}(r, \lambda_0)] dr \quad (1.18)$$

where  $\alpha_{mol}(r, \lambda_0)$  and  $\alpha_{aer}(r, \lambda_0)$  are the extinction coefficient due to molecular and aerosol absorption and scattering respectively. Eq. 1.18 is evaluated from the surface to the altitude  $z$ .

For Raman scattering lidars the laser wavelengths  $\lambda_0$  do not equal to the wavelength detected at the telescope. Therefore, the volume backscatter cross-sections need to be evaluated at both wavelengths.

$$\beta(z, \lambda) = \beta(z, \lambda_0, \lambda_R) \quad (1.19)$$

where  $\beta(z, \lambda_0, \lambda_R)$  is the volume backscatter cross-section due to molecular scattering at the Raman shifted wavelength  $\lambda_R$  detected at the telescope.

The atmospheric transmission for Raman scattering also needs to be evaluated at both  $\lambda_0$

and  $\lambda_R$  wavelengths. Therefore, the optical depth terms for Raman transmission is

$$\tau_{\lambda_0} + \tau_{\lambda} = \int_0^z [\alpha_{mol}(r, \lambda_0) + \alpha_{mol}(r, \lambda_R) + \alpha_{aer}(r, \lambda_0) + \alpha_{aer}(r, \lambda_R)] dr \quad (1.20)$$

where  $\alpha_{mol}(r, \lambda_0)$  and  $\alpha_{mol}(r, \lambda_R)$  are the molecular extinction coefficients at  $\lambda_0$  and  $\lambda_R$  wavelengths respectively.  $\alpha_{aer}(r, \lambda_0)$  and  $\alpha_{aer}(r, \lambda_R)$  are the extinction coefficients due to aerosol scattering at  $\lambda_0$  and  $\lambda_R$  wavelengths respectively. Both aerosol and molecular extinction coefficients are measured in the SI units of inverse meters.

The molecular and aerosol extinction terms occurring in the atmospheric transmission are related:

$$\alpha_{aer} = LR \cdot \beta_{aer} = LR \cdot \beta_{mol} \cdot (\mathfrak{R}_{\beta} - 1) \quad (1.21)$$

where  $LR$  is the lidar ratio and  $\mathfrak{R}_{\beta}$  is the backscatter ratio.

$$\mathfrak{R}_{\beta} = \frac{(\beta_{mol} + \beta_{aer})}{\beta_{mol}}. \quad (1.22)$$

The lidar ratio is different for different types of particles. For molecules, the lidar ratio is  $\frac{8\pi}{3}$  sr assuming no absorption (Cattrall et al., 2005). The lidar ratio for aerosols depends on the size and shape of the particle, the wavelength scattered, and the refractive index of the aerosol being scattered. For visible wavelength the aerosol lidar ratio can vary from 20 sr to 100 sr (Ansmann et al., 1992). The lidar ratio for clouds is usually smaller for clouds than aerosols and can vary between 6 sr and 60 sr depending on the cloud type and altitude (Reichardt, 1999; Giannakaki et al., 2007; Chen et al., 2002).

For Raman lidar systems that detect pure rotational Raman wavelengths, the atmospheric transmissions are approximately the same as the atmospheric transmission for elastic scattering as given in Eq. 1.18. However, the volume backscatter coefficient needs to be evaluated at each return wavelength for pure rotational Raman scattering.

The differential backscatter cross-section for single lines of the PRR spectrum is expressed

by (Behrendt, 2005):

$$\left(\frac{d\sigma}{d\Omega}\right)_{\Pi}^{RR,i}(J) = \frac{112.\Pi^4 g_i(J)hcB_{0,i}(v_o + \Delta v_i(J))^4 \zeta_i^2}{(2I_i + 1)^2 kT} \times X(J) \exp\left(\frac{-E_{rot,i}(J)}{kT}\right) \quad (1.23)$$

where for the Stokes branch

$$X(J) = \frac{(J+1)(J+2)}{2J+3} \quad \text{for } J = 0, 1, 2, \dots \quad (1.24)$$

and for the anti-Stokes branch

$$X(J) = \frac{J(J-1)}{2J-1}. \quad \text{for } J = 2, 3, 4, \dots \quad (1.25)$$

Here  $g_i(J)$  is the statistical weighting factor, which depends on the nuclear spin,  $I_i$ , for each atmospheric constituent,  $h$  is Planck's constant,  $c$  is the velocity of light,  $k$  is Boltzmann's constant,  $B_{0,i}$  is the ground-state rotational constant,  $v_o$  is the frequency of the incident light, and  $\zeta_i$  is the anisotropy of the molecular polarizability.

The rotational energy,  $E_{rot,i}(J)$  for each Stokes and anti-Stokes branch is estimated based on the assumption of a homonuclear diatomic molecule in the quantum state  $J$  for nitrogen and oxygen molecules with no electronic momentum coupled to the scattering (Behrendt, 2005).

Accounting for all changes in the lidar equation (Eq. 1.14) for elastic and Raman scattering,

$$N_{elastic}(z, \lambda_0) = C \frac{O(z)}{z^2} [\beta_{mol}(z, \lambda_0) + \beta_{aer}(z, \lambda_0)] e^{-2 \int_0^z [\alpha_{mol}(r, \lambda_0) + \alpha_{aer}(r, \lambda_0)] dr} + B(z), \quad (1.26)$$

$$N_{Raman}(z, \lambda_0, \lambda_R) = C \frac{O(z)}{z^2} \beta(z, \lambda_0, \lambda_R) e^{-\int_0^z [\alpha_{mol}(r, \lambda_0) + \alpha_{mol}(r, \lambda_R) + \alpha_{aer}(r, \lambda_0) + \alpha_{aer}(r, \lambda_R)] dr} + B(z), \quad (1.27)$$

where  $N_{elastic}(z, \lambda_0)$  and  $N_{Raman}(z, \lambda_0, \lambda_R)$  are the observed photon count rates for elastic scattering and Raman scattering.

The lidar equation given in Eq. 1.14 presumes that the count rates linearly depend on the

number of received photons that are valid for analog channels and for digital channels when the count rates are below 1 MHz. For the digital channels at higher count rates, the true ( $N_{tr}$ ) and observed ( $N_{obs}$ ) counts are related by dead time  $\gamma$  (Donovan et al., 1993).

For a non-paralyzable counting system (Kovalev and Eichinger, 2004):

$$N_{obs} = \frac{N_{tr}}{1 + N_{tr}\gamma}. \quad (1.28)$$

For a paralyzable counting system (Kovalev and Eichinger, 2004):

$$N_{obs} = N_{tr} \exp(-\gamma N_{tr}) \quad (1.29)$$

## 1.4 Raman lidar for Meteorological Observations (RALMO)

The Raman scattering cross-sections are several orders of magnitude smaller than those for elastic scattering. Thus, Raman lidars are the best and widely used lidar technique to measure lower atmospheric temperature and humidity. Raman lidars detect rotational Raman bands that contain information on the temperature and the vibrational Raman bands that contain water vapor information in the scattering volumes.

The main objective of this thesis is to develop a new technique to accurately retrieve atmospheric temperature and relative humidity from Raman lidar measurements. Therefore, I use Raman lidar measurements from the Meteoswiss/EPFL Raman lidar for Meteorological Observations (RALMO), located in Payerne, Switzerland (46°48' N, 6°56' E). RALMO is a fully automated lidar, operating since 2008 with an average data availability of 50%. The transmitting system of the RALMO consists of a frequency tripled, Q-switched Nd:YAG laser of 354.7 nm supplying up to 400 mJ emission energy at a 30 Hz repetition rate with a 8 ns pulse duration. The lidar telescope receiver consists of four 30 cm in diameter telescopes that are tightly arranged around a 15x beam expander and a near range optical fiber, located off of the optical axis of one of the four telescopes.

The RALMO consists of two polychromators that allow water vapor, temperature, and aerosol profiling. All five telescopes of the RALMO are coupled by fiber optics to the wa-



### 1.4.1 Pure rotational Raman (PRR) lidar

RALMO detects the returns of the Raman-shifted backscatter arising from rotational energy state transitions of nitrogen and oxygen molecules due to the excitation at the laser wavelength at 354.7 nm in analog and photon counting mode. RALMO consists of two PRR channels whose measurements are used to calculate the atmospheric temperature as a function of height. The high quantum number channel (JH) of RALMO is assigned to the backscattered signals from the energy exchange that occurs in the high quantum states for both the Stokes (355.77-356.37 nm) and anti-Stokes (353.07-353.67 nm) branches. The low quantum number channel (JL) is assigned to the signals from the energy exchange occurring in the low quantum states in the Stokes (355.17-355.76 nm) and anti-Stokes (353.67-354.25 nm) branches. Using the equations given by Herzberg (2013) the quantum numbers detected by each PRR channel for both nitrogen and oxygen molecules can be calculated. Table 1.2 shows the quantum states detected by each PRR channel.

Table 1.2: Detected quantum lines from nitrogen and oxygen in PRR spectrum.

PRR channel	Nitrogen quantum lines (Stokes and anti-Stokes)	Oxygen quantum lines (Stokes and anti-Stokes)
JL	3,4,5,7,8,9	5,7,9,11,13
JH	10,11,12,13,14,15	15,17,19,21

### 1.4.2 Vibrational Raman/water vapor system

The second polychromator in the RALMO is designed to isolate the rotational-vibrational Raman signals of nitrogen and water vapor (wavelengths of 386.7 and 407.5 nm, respectively). The optical signals are detected by photomultipliers and acquired by a transient recorder. These measurements are used to estimate the atmospheric mixing ratio as a function of height.

A detailed description of RALMO is given by Dinoev et al. (2010), and the instruments validation is given in (Brocard et al., 2013).

## 1.5 Traditional Raman lidar algorithms to calculate temperature and relative humidity

This section presents the application of traditional Raman lidar algorithms to estimate temperature, water vapor mixing ratio, and relative humidity using the Raman lidar measurements.

### 1.5.1 Temperature measurements

The traditional method for calculating temperature from lidar PRR measurements, originally proposed by Cooney (1972), can be broken down into four steps.

1. Correction of the raw measurements for effects including dead time, background, signal-induced noise, gluing, etc., and then smoothing of the corrected PRR lidar signals.
2. Determination of the coefficients of the calibration function using a reference temperature profile (from a co-located radiosonde, for example) and the previously smoothed lidar measurements.
3. Calculation of temperature using the inversion of the calibration function.
4. Estimation of the random and systematic temperature uncertainties.

To determine temperature in this method, the ratio  $Q(T, z)$  is formed:

$$Q(T, z) = \frac{N_{RR2}(T, z)}{N_{RR1}(T, z)} \quad (1.30)$$

where  $N_{RR1}$  and  $N_{RR2}$  are the corrected PRR signals, each of which has an opposite temperature dependence. Assuming both channels have the same geometrical overlap function and atmospheric extinction, the temperature dependence of the ratio  $Q$  is:

$$Q(T, z) = \frac{\left( \sum_{i=O_2, N_2} \sum_{J_i} \tau_{RR2}(J_i) \eta_i \left( \frac{d\sigma}{d\Omega} \right)_\pi^{RR,i} (J_i) \right)}{\left( \sum_{i=O_2, N_2} \sum_{J_i} \tau_{RR1}(J_i) \eta_i \left( \frac{d\sigma}{d\Omega} \right)_\pi^{RR,i} (J_i) \right)}. \quad (1.31)$$

The differential backscatter coefficients for the above equation need to be estimated depending on which branch of the PRR spectrum is used. The transmissions of the lidar receiver for each rotational Raman line ( $J_i$ ) are  $\tau_{RR1}(J_i)$  and  $\tau_{RR2}(J_i)$ .

The ratio  $Q(T, z)$  can be expressed as a function of temperature by a simple analytic formula (Arshinov et al., 1983) that is also known as the calibration function (Behrendt, 2005):

$$Q(T, z) = \exp\left(a - \frac{b}{T}\right) \quad (1.32)$$

where  $a$  and  $b$  are the calibration coefficients. Behrendt (2005) shows the uncertainties due to the choice of calibration function, such as using a second-order polynomial. He concluded that when measurements which cover a large range of temperature ( $> 50$  K) are available for calibration, an extended version of Eq. 1.32 with a second-order temperature term is the best-suited calibration function. Otherwise, Eq. 1.32 with two calibration coefficients should be used to avoid large extrapolation uncertainties. In practice, temperature measurements from radiosondes flown coincidentally with the lidar measurements are used to estimate the calibration coefficients. The frequency of the re-calibration differs for individual lidar systems. Equation 1.32 is widely used in many lidar systems that detect single or multiple RR lines in each RR channel (Arshinov et al., 1983; Behrendt and Reichardt, 2000). The frequency of the calibration differs for individual lidar systems. Eq. 1.32 is also valid for lidar systems that retrieve multiple RR lines in each RR channel (Behrendt and Reichardt, 2000).

Further studies of approximation of calibration functions and effect on the temperature calculations can be found in Cooney (1972); Arshinov et al. (1983); Di Girolamo et al. (2004); Behrendt (2005), and Zuev et al. (2017).

## 1.5.2 Water vapor mixing ratio

The traditional Raman method for water vapor mixing ratio measurements uses two Raman signals, one of which is the backscatter signal from the water vapor ( $N_{obs,wv}$ , 406 nm), and the backscatter signal from the nitrogen ( $N_{obs,n2}$ , 386 nm) which is considered as the reference signal. Usually, a single line from the Stokes vibration-rotation of the water vapor and nitrogen spectra is detected by each channel. The water vapor mixing ratio ( $m_{wv}(z)$ ) relative to dry



air can be obtained from the signal ratio of water vapor and nitrogen measurements, that are corrected for non-linearity and background and then rearranging the resulting equation:

$$m_{wv}(z) = \frac{n_{wv}(z)}{n_{n2}(z)} = K_{wv} \frac{(N_{obs,wv}(z) - B_{wv}) \exp\left(-\int_0^z \alpha_{\lambda_{wv}}(z) dz\right)}{(N_{obs,n2}(z) - B_{n2}) \exp\left(-\int_0^z \alpha_{\lambda_{n2}}(z) dz\right)}. \quad (1.33)$$

$K_{wv}$  is the water vapor calibration constant that depends on the instrumental transmission, detector deficiencies, Raman scattering cross-sections of water vapor and nitrogen and the ratio of the number densities of nitrogen to the dry air. The calibration constant can be determined by comparison with *in situ* water vapor measurements. Hence, the lidar water vapor mixing ratio profiles depend on the accuracy of the measurements made by the second instrument. Often, the relative humidity measurements (later converted to water vapor mixing ratios for comparison) from the coincident radiosonde are used to determine the water vapor calibration constant. Accuracy of the water vapor calibration constant depends on the sonde measurements, and improved calibration techniques are given in the literature (Whiteman et al., 1992; Ferrare et al., 1995).

In Eq. 1.33 the overlap functions for water vapor and nitrogen channels are assumed to be identical, which is reasonable because both channels are connected to the same receiving system in the RALMO. The atmospheric transmission at water vapor and nitrogen wavelengths, due to Rayleigh scattering that is required for the calibration process, can be estimated using the equations given in Nicolet (1984).

The traditional Raman lidar temperature algorithm uses the ratio of two corrected pure rotational Raman (PRR) signals (Cooney, 1972). Corrected PRR signals can be obtained by applying, if required, a non-linear correction to the raw lidar measurements and by removing the background signal. The traditional retrieval algorithm also requires a determination of a lidar calibration function and the calibration function coefficients using ancillary measurements. Primary importance is in the calibration of the lidar returns to allow absolute temperature measurements. In the traditional method, the ratio of the corrected photocounts from the two PRR channels is fit to a set of corresponding temperature data points usually obtained from radiosondes. The calibration function is an approximation of the relationship of the signal ratio and temperature and depends on two or more coefficients. Calibration errors exceeding 0.5 K

can arise if the calibration data do not cover a sufficient temperature range (Behrendt, 2005). Based on the determined lidar calibration function, temperature, and absolute and relative uncertainties can be estimated. Studies by Di Girolamo et al. (2004); Gerasimov and Zuev (2016); Zuev et al. (2017) have shown application of various types of calibration functions and their uncertainties to estimate temperature from the Raman lidar measurements.

The Raman lidar technique for water vapor mixing ratio calculation also requires a ratio of corrected measurements from water vapor and nitrogen signals. A calibration constant that relates to the molecular and optical properties of both scattering species and systematic corrections is required. Similar to the traditional Raman lidar temperature algorithm, the water vapor calibration constant can also be determined using ancillary measurements. Also, to obtain relative humidity profile one has to combine the calculated water vapor mixing ratio and temperature profiles. The uncertainty of the relative humidity profiles can be calculated by error propagation and only a statistical uncertainty can be estimated without any effort. The first attempt to calculate relative humidity by combining Raman water vapor and temperature measurements was made by Mattis et al. (2002) and the relative uncertainty of the calculated relative humidity was between 5 to %, the uncertainty of the temperatures being the dominating source of uncertainty. A 1 to 2 K temperature accuracy is required for reducing the relative uncertainty in the relative humidity to, on average, less than 10%. The Raman lidar measurements used to calculate temperature and water vapor mixing ratios need to be corrected for non-linearity and background.

### 1.5.3 Relative humidity

Raman lidar measurements are not sensitive to the relative humidity but to densities of water vapor molecules in the atmosphere (mixing ratio). Thus, direct determination of relative humidity profiles from the Raman lidar measurements is not possible. One has to calculate temperature from pure-Rotational Raman (PRR) lidar channels and water vapor mixing ratios from water vapor Raman lidar channels simultaneously to calculate relative humidity. A first attempt to calculate relative humidity over water from Raman lidar measurements by combining Raman temperature and water vapor mixing ratio is presented by Mattis et al. (2002).

As discussed in Section 1.2.2 relative humidity can be expressed as the ratio of vapor pressure to saturated vapor pressure over water or ice (Eq. 1.6). Also, as shown in Eq. 1.3 vapor pressure is a function of mixing ratio where the ratio of mass of water vapor to mass of dry air is considered. However, as given in Eq. 1.33, Raman lidar measures mixing ratio relative to the nitrogen. Therefore, I can approximate the lidar mixing ratio  $m_{wv}$  using

$$m_{wv} = \frac{n_{wv}}{n_{n2}} \approx \frac{n_{wv}}{0.78n_{dry}}. \quad (1.34)$$

Hence, vapor pressure  $e$  can be expressed as,

$$e = \frac{m_{wv}}{0.797 + m_{wv}} P. \quad (1.35)$$

The pressure  $P(z)$  required to find the water vapor pressure, can be obtained either from routine sonde measurements or can be estimated using the hydro-static equilibrium or even from standard atmospheric models. Hydro-static equilibrium describes the complete balance between the force of gravity and the atmospheric pressure:

$$P_{HSEQ}(z) = P_0 \exp \left[ - \int_{z_0}^{z_{top}} \frac{M(z)g(z)}{R^*T(z)} dz \right]. \quad (1.36)$$

where  $P_{HSEQ}(z)$  is the pressure,  $M(z)$  is the height-dependent mean molecular mass,  $g(z)$  is the acceleration due to gravity,  $P_0$  is the pressure at the maximum altitude, and  $R^*$  is the universal gas constant.

Relative humidity over water can be obtained by applying the saturated vapor pressure over water (Eq. 1.4) and the vapor pressure from Eq. 1.35 in Eq. 1.6. Similarly, relative humidity over ice can be obtained by applying saturated vapor pressure over ice (log inverse of Eq. 1.5) and the vapor pressure from Eq. 1.35 in Eq. 1.6.

One of the advantage of calculating relative humidity over water from the Raman lidars is that those can be compared with the radiosonde relative humidity profiles that are measured over water. The majority of the existing records of relative humidity for temperatures below 0°C are expressed relative to water. For the studies such as ice supersaturation it is required to consider relative humidity over ice for temperatures below 0°C (Tables, 1951).

## 1.6 Optimal estimation method (OEM)

The OEM is an inverse method that uses the measurements  $y$  to estimate the state (retrieval) variables  $x$  of a system via a forward model. The forward model  $F$  contains all the physical and instrumental factors describing the measurements including the atmospheric state and the instrument. The forward model can include model parameters  $b$  which are assumed and not retrieved, and their effect on the retrieved quantities uncertainty can be calculated. A short summary of the method is given below. For more details refer to Rodgers book on inverse modelling (Rodgers, 2000). The measurements are related to the forward model by:

$$y = F(x, b) + \epsilon \quad (1.37)$$

where  $\epsilon$  is the difference between the measurements and the forward model, that is the measurement uncertainty. In the presence of measurement uncertainties, a statistical estimate for the state vector can be obtained from an estimate of the *a priori* state vector ( $x_a$ ).

Under the assumption that all parameters have Gaussian probability density functions Bayes theorem, to determine the probability that the system has a state  $x_a$ , given the measurements  $y$ ,  $P(x_a|y)$ ,

$$-2 \ln P(x|y) = [y - F(x, b)]^T S_y^{-1} [y - F(x, b)] + [x - x_a]^T S_a^{-1} [x - x_a], \quad (1.38)$$

where  $S_y$  is the measurement covariance that describes the random measurement uncertainty and  $S_a$  is the *a priori* covariance. The quantity  $-2 \ln P(x_a|y)$  normalized by the length of  $y$  is referred to the cost of the OEM retrieval. It measures the goodness of fit for a solution, and for good models the cost is on the order of unity. The first term in the cost function is similar to the root-mean square deviation (RMSD) of the modelled measurement to the observed, weighted by the error on that observation. The second term in the cost function represents the deviation of the state from the *a priori*, weighted by its expected distribution.

The maximum probability state  $\hat{x}$  then solves,

$$\nabla_x (-2 \ln P(x|y)) = -K^T S_y^{-1} [y - F(\hat{x}, b)] + S_a^{-1} [\hat{x} - x_a] = 0 \quad (1.39)$$

where  $K = \nabla_x F(\hat{x}, b)$  is the Jacobian that describes the sensitivity of the forward model to the state.

For linear forward models Eq. 1.24 can be solved with Newton iteration. For quasi-linear forward models, it can be solved using the Gauss-Newton iteration. For the forward models that are nonlinear the Marquardt-Levenberg method can be used:

$$x_{i+1} = x_i + [(1 + \Gamma_i)S_a^{-1} + K_i^T S_y^{-1} K_i]^{-1} \{K_i^T S_y^{-1} [y - F(x_i, b)] - S_a^{-1} [x_i - x_a]\} \quad (1.40)$$

where  $\Gamma_i$  is chosen at each iteration to minimize the cost function. An initial guess for  $\Gamma$  needs to be set up and the factor by which it is increased or decreased is a matter for the experiment in each particular case. In my work I use the factor by which  $\Gamma$  increases or decreases at each step to be 10 that is also the value suggested by Marquardt.

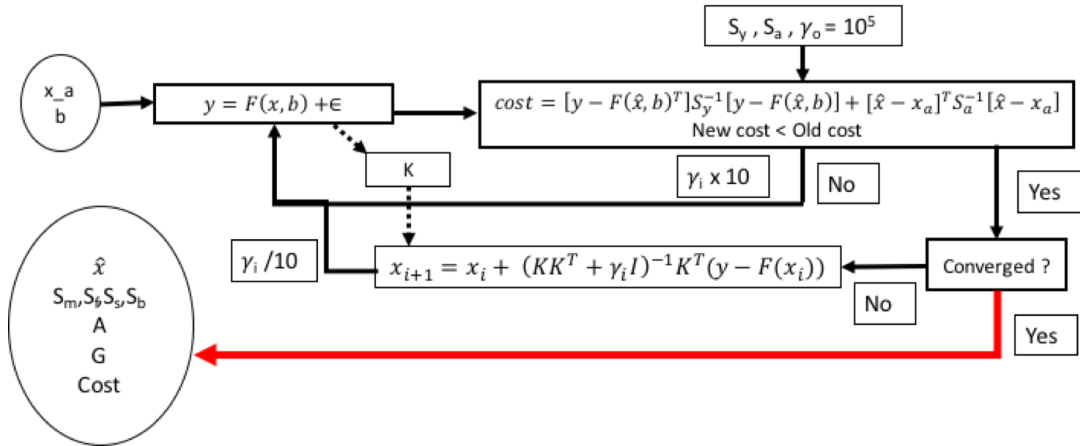


Figure 1.8: Schematic of the optimal estimation method retrieval algorithm.

The schematic of the OEM algorithm is shown in Fig. 1.8. If the cost function increases after an iteration,  $\Gamma_i$  is increased by a factor of ten. Otherwise, it is reduced by a factor of ten. I consider only 15 maximum iteration steps in my OEM algorithm and the OEM algorithm is considered to be converged when the cost is decreased and the change in the state is less than some fraction of the predicted error,

$$(x_{i+1} - x_i)_j < \epsilon \sqrt{\hat{S}_{jj}} \quad \forall j. \quad (1.41)$$

When the cost function reaches its minimum, the optimum retrieval parameters  $\hat{x}$  can be obtained using the following expression:

$$\hat{x} = (I - A)x_a + Ax + (\hat{K}^T S_y^{-1} \hat{K} + S_a^{-1})^{-1} \hat{K}^T S_y^{-1} \epsilon \quad (1.42)$$

where  $G$  is the gain matrix. It indicates the sensitivity of the retrieval to the measurement:

$$G = \frac{\partial \hat{x}}{\partial y}. \quad (1.43)$$

Once  $K$  and  $G$  are found, the averaging kernels (or weighting functions)  $A$  can be calculated. The averaging kernel matrix describes the sensitivity of the retrieved state to the true state, while the width of each element defines the resolution of the retrieval. The product of the gain and the Jacobian is the averaging kernel:

$$A = \frac{\partial \hat{x}}{\partial x} = GK. \quad (1.44)$$

For a perfect retrieval,  $A$  is the unit matrix. In reality, rows of  $A$  are generally peaked functions and the width of each element defines the spatial resolution of the retrieval. The area of the averaging kernel, or the sum of its elements, can also be regarded as the response of the retrieval. The response is found to be approximately unity at the levels where the retrieval is accurate, and in general can be considered as a rough measure of the fraction of the information that comes from the measurements, rather than from the *a priori* profile.

The uncertainty budget is determined from the measurement and model parameter covariance matrices (Rodgers, 2000). The total covariance  $S_{total}$  is:

$$S_{total} = S_m + S_F. \quad (1.45)$$

where  $S_m$  is the retrieval covariance due to measurement noise and  $S_F$  is the retrieval covariance due to the forward model parameter uncertainty. The retrieval covariance due to measurement noise  $S_m$  is

$$S_m = GS_y G^T. \quad (1.46)$$

The retrieval covariance due to the forward model parameters  $S_F$  is

$$S_F = GK_b S_b K_b^T G^T, \quad (1.47)$$

where  $K_b$  and  $S_b$  are the forward model parameter Jacobian and covariance matrices respectively. More details of the OEM are given in the Rodger's textbook (Rodgers, 2000).

### 1.6.1 Advantages of implementing the OEM for lidar temperature and relative humidity retrievals.

The OEM offers several significant advantages over the traditional Raman lidar retrievals:

- The OEM uses raw lidar measurements without the need of gluing or corrections for background or non-linearities.
- The OEM allows the retrieval of multiple parameters from the same measurements. For an example temperature, overlap function, and particle extinction can be retrieved from the Raman PRR measurements.
- The OEM provides a complete uncertainty budget, including random and systematic uncertainties due to model parameters.
- No post or pre-filtering of the retrieval is required and the height resolution of the retrieved profile is the full width at half maximum of the averaging kernels, which is computed at each height bin.
- The OEM is fast computationally (on the order of the traditional method), and is extremely fast relative to some other inversion methods such as the grid-search technique.
- The OEM can be applied at any required height or time resolution, e.g., nightly averaged profiles that are co-added in height or individual profiles at high temporal-spatial resolution.
- The OEM is flexible and can be extended with further measurements in the forward model and/or additional retrieval parameters.

## Bibliography

- Alpers, M., Eixmann, R., Fricke-Begemann, C., Gerding, M., and Höffner, J.: Temperature lidar measurements from 1 to 105 km altitude using resonance, Rayleigh, and Rotational Raman scattering, *Atmospheric Chemistry and Physics*, 4, 793–800, 2004.
- Ansmann, A., Wandinger, U., Riebesell, M., Weitkamp, C., and Michaelis, W.: Independent measurement of extinction and backscatter profiles in cirrus clouds by using a combined Raman elastic-backscatter lidar, *Applied optics*, 31, 7113–7131, 1992.
- Arshinov, Y. F., Bobrovnikov, S., Zuev, V. E., and Mitev, V.: Atmospheric temperature measurements using a pure rotational Raman lidar, *Applied Optics*, 22, 2984–2990, 1983.
- Atmosphere, U. S.: National oceanic and atmospheric administration, National Aeronautics and Space Administration, United States Air Force, Washington, DC, 1976.
- Behrendt, A.: Temperature measurements with lidar, in: *Lidar*, pp. 273–305, Springer, 2005.
- Behrendt, A. and Reichardt, J.: Atmospheric temperature profiling in the presence of clouds with a pure rotational Raman lidar by use of an interference-filter-based polychromator, *Applied Optics*, 39, 1372–1378, 2000.
- Bock, O., Bosser, P., Bourcy, T., David, L., Goutail, F., Hoareau, C., Keckhut, P., Legain, D., Pazmino, A., Pelon, J., et al.: Accuracy assessment of water vapour measurements from in situ and remote sensing techniques during the DEMEVAP 2011 campaign at OHP, *Atmospheric Measurement Techniques*, 6, 2777–2802, 2013.
- Brocard, E., Philipona, R., Haeferle, A., Romanens, G., Mueller, A., Ruffieux, D., Simeonov, V., and Calpini, B.: Raman Lidar for Meteorological Observations, RALMO-Part 2: Validation of water vapor measurements, *Atmospheric Measurement Techniques*, 6, 2013.
- Buehler, S., Kuvatov, M., John, V., Leiterer, U., and Dier, H.: Comparison of microwave satellite humidity data and radiosonde profiles: A case study, *Journal of Geophysical Research: Atmospheres*, 109, 2004.



- Cattrall, C., Reagan, J., Thome, K., and Dubovik, O.: Variability of aerosol and spectral lidar and backscatter and extinction ratios of key aerosol types derived from selected Aerosol Robotic Network locations, *Journal of Geophysical Research: Atmospheres*, 110, 2005.
- Chahine, M. T.: The hydrological cycle and its influence on climate, *Nature*, 359, 373, 1992.
- Chen, W.-N., Chiang, C.-W., and Nee, J.-B.: Lidar ratio and depolarization ratio for cirrus clouds, *Applied Optics*, 41, 6470–6476, 2002.
- Comstock, J. M., Ackerman, T. P., and Turner, D. D.: Evidence of high ice supersaturation in cirrus clouds using ARM Raman lidar measurements, *Geophysical research letters*, 31, 2004.
- Cooney, J.: Measurement of atmospheric temperature profiles by Raman backscatter, *Journal of applied meteorology*, 11, 108–112, 1972.
- Di Girolamo, P., Marchese, R., Whiteman, D., and Demoz, B.: Rotational Raman Lidar measurements of atmospheric temperature in the UV, *Geophysical Research Letters*, 31, 2004.
- Dinoyev, T., Simeonov, V., Calpini, B., and Parlange, M.: Monitoring of Eyjafjallajökull ash layer evolution over payerne Switzerland with a Raman lidar, *Proceedings of the TECO*, 2010.
- Dirksen, R., Sommer, M., Immler, F., Hurst, D., Kivi, R., and Vömel, H.: Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde, *Atmospheric Measurement Techniques*, 7, 4463–4490, 2014.
- Donovan, D., Whiteway, J., and Carswell, A. I.: Correction for nonlinear photon-counting effects in lidar systems, *Applied optics*, 32, 6742–6753, 1993.
- DuBois, J. L.: Invention and Development of the Radiosonde with a Catalog of Upper-Atmospheric Telemetering Probes in the National Museum of American History, Smithsonian Institution, *Smithsonian Studies in History and Technology*, 2002.
- Durre, I., Vose, R. S., and Wuertz, D. B.: Overview of the integrated global radiosonde archive, *Journal of Climate*, 19, 53–68, 2006.

- Farhani, G., Sica, R. J., Godin-Beekmann, S., and Haeferle, A.: Optimal Estimation Method Retrievals of Stratospheric Ozone Profiles from a DIAL Lidar, *Atmospheric Measurement Techniques Discussions*, 2018, 1–22, <https://doi.org/10.5194/amt-2018-310>, 2018.
- Ferrare, R., Melfi, S., Whiteman, D., Evans, K., Schmidlin, F., and Starr, D. O.: A comparison of water vapor measurements made by Raman lidar and radiosondes, *Journal of Atmospheric and Oceanic Technology*, 12, 1177–1195, 1995.
- Ferreira, A. P., Nieto, R., and Gimeno, L.: Completeness of radiosonde humidity observations based on the Integrated Global Radiosonde Archive, *Earth System Science Data*, 11, 603–627, 2019.
- Fusina, F., Spichtinger, P., and Lohmann, U.: Impact of ice supersaturated regions and thin cirrus on radiation in the midlatitudes, *Journal of Geophysical Research: Atmospheres*, 112, 2007.
- Gerasimov, V. V. and Zuev, V. V.: Analytical calibration functions for the pure rotational Raman lidar technique, *Optics express*, 24, 5136–5151, 2016.
- Giannakaki, E., Balis, D., Amiridis, V., and Kazadzis, S.: Optical and geometrical characteristics of cirrus clouds over a Southern European lidar station, *Atmospheric Chemistry and Physics*, 7, 5519–5530, 2007.
- Gierens, K., Schumann, U., Helten, M., Smit, H., and Wang, P.-H.: Ice-supersaturated regions and subvisible cirrus in the northern midlatitude upper troposphere, *Journal of Geophysical Research: Atmospheres*, 105, 22 743–22 753, 2000.
- Gierens, K., Spichtinger, P., and Schumann, U.: Ice supersaturation, in: *Atmospheric Physics*, pp. 135–150, Springer, 2012.
- Held, I. M. and Soden, B. J.: Water vapor feedback and global warming, *Annual review of energy and the environment*, 25, 441–475, 2000.
- Herzberg, G.: *Molecular spectra and molecular structure*, vol. 1, Read Books Ltd, 2013.

- Hewison, T. J.: 1D-VAR retrieval of temperature and humidity profiles from a ground-based microwave radiometer, *IEEE transactions on geoscience and remote sensing*, 45, 2163–2168, 2007.
- Hopkins, E. J.: An Upper Air Probe, URL <http://www.aos.wisc.edu/~hopkins/wx-inst/wxi-raob.htm>.
- Hyland, R.: Formulations for Thermodynamic Properties of the Saturated Phases Of H<sub>2</sub>O from 173.15 K to 473.15 K, *Ashrae transactions*, A, 2, 500–513, 1983.
- Immler, F., Treffeisen, R., Engelbart, D., Krüger, K., and Schrems, O.: Cirrus, contrails, and ice supersaturated regions in high pressure systems at northern mid latitudes, *Atmospheric Chemistry and Physics*, 8, 1689–1699, 2008.
- Jensen, E., Smith, J., Pfister, L., Pittman, J., Weinstock, E., Sayres, D., Herman, R., Troy, R., Rosenlof, K., Thompson, T., et al.: Ice supersaturations exceeding 100% at the cold tropical tropopause: implications for cirrus formation and dehydration, *Atmospheric Chemistry and Physics*, 5, 851–862, 2005.
- Jensen, E. J., Toon, O. B., Vay, S. A., Ovarlez, J., May, R., Bui, T., Twohy, C. H., Gandrud, B. W., Pueschel, R. F., and Schumann, U.: Prevalence of ice-supersaturated regions in the upper troposphere: Implications for optically thin ice cloud formation, *Journal of Geophysical Research: Atmospheres*, 106, 17 253–17 266, 2001.
- Jones, A., Urban, J., Murtagh, D. P., Eriksson, P., Brohede, S., Haley, C., Degenstein, D., Bourassa, A., Savigny, C. v., Sonkaew, T., et al.: Evolution of stratospheric ozone and water vapour time series studied with satellite measurements, *Atmospheric Chemistry and Physics*, 9, 6055–6075, 2009.
- Kovalev, V. A. and Eichinger, W. E.: *Elastic lidar: theory, practice, and analysis methods*, John Wiley & Sons, 2004.
- Krämer, M., Schiller, C., Afchine, A., Bauer, R., Gensch, I., Mangold, A., Schlicht, S., Spelten, N., Sitnikov, N., Borrmann, S., et al.: Ice supersaturations and cirrus cloud crystal numbers, *Atmospheric Chemistry and Physics*, 9, 3505–3522, 2009.

- Leiterer, U., Dier, H., and Naeber, T.: Improvements in radiosonde humidity profiles using RS80/RS90 radiosondes of Vaisala, *Contributions to atmospheric physics*, 70, 319–336, 1997.
- Liljegren, J. C., Boukabara, S.-A., Cady-Pereira, K., and Clough, S. A.: The effect of the half-width of the 22-GHz water vapor line on retrievals of temperature and water vapor profiles with a 12-channel microwave radiometer, *IEEE Transactions on Geoscience and Remote Sensing*, 43, 1102–1108, 2005.
- Lindsey, R.: Climate change: atmospheric carbon dioxide, National Oceanographic and Atmospheric Administration, News & Features. August, 2018.
- List, R.: Smithsonian Meteorological Tables. 6th rev version, 1984.
- Löhnert, U., Turner, D., and Crewell, S.: Ground-based temperature and humidity profiling using spectral infrared and microwave observations. Part I: Simulated retrieval performance in clear-sky conditions, *Journal of Applied Meteorology and Climatology*, 48, 1017–1032, 2009.
- Mahagammulla Gamage, S., Sica, R. J., Martucci, G., and Haeferle, A.: Retrieval of Temperature From a Multiple Channel Pure Rotational Raman-Scatter Lidar Using an Optimal Estimation Method, *Atmospheric Measurement Techniques Discussions*, 2019, 1–34, <https://doi.org/10.5194/amt-2019-107>, URL <https://www.atmos-meas-tech-discuss.net/amt-2019-107/>, 2019.
- Marshall, J. and Plumb, R. A.: *Atmosphere, ocean and climate dynamics: an introductory text*, vol. 43, Academic Press, 1989.
- Mattis, I., Ansmann, A., Althausen, D., Jaenisch, V., Wandinger, U., Müller, D., Arshinov, Y. F., Bobrovnikov, S. M., and Serikov, I. B.: Relative-humidity profiling in the troposphere with a Raman lidar, *Applied optics*, 41, 6451–6462, 2002.
- Miloshevich, L. M., Vömel, H., Paukkunen, A., Heymsfield, A. J., and Oltmans, S. J.: Characterization and correction of relative humidity measurements from Vaisala RS80-A radioson-

- des at cold temperatures, *Journal of Atmospheric and Oceanic Technology*, 18, 135–156, 2001.
- Murray, F. W.: On the computation of saturation vapor pressure., Tech. rep., Rand Corp Santa Monica Calif, 1966.
- Nagel, D., Leiterer, U., Dier, H., Kats, A., Reichardt, J., and Behrendt, A.: High accuracy humidity measurements using the standardized frequency method with a research upper-air sounding system, *Meteorologische Zeitschrift*, 10, 395–405, 2001.
- Nicolet, M.: On the molecular scattering in the terrestrial atmosphere: An empirical formula for its calculation in the homosphere, *Planetary and Space Science*, 32, 1467–1468, 1984.
- Noh, Y.-C., Sohn, B.-J., Kim, Y., Joo, S., and Bell, W.: Evaluation of temperature and humidity profiles of Unified Model and ECMWF analyses using GRUAN radiosonde observations, *Atmosphere*, 7, 94, 2016.
- Palmer, P. I., Barnett, J., Eyre, J., and Healy, S.: A nonlinear optimal estimation inverse method for radio occultation measurements of temperature, humidity, and surface pressure, *Journal of Geophysical Research: Atmospheres*, 105, 17 513–17 526, 2000.
- Palmer, T. N.: Predicting uncertainty in forecasts of weather and climate, *Reports on progress in Physics*, 63, 71, 2000.
- Pan, L. L., Bowman, K. P., Atlas, E. L., Wofsy, S. C., Zhang, F., Bresch, J. F., Ridley, B. A., Pittman, J. V., Homeyer, C. R., Romashkin, P., et al.: The stratosphere–troposphere analyses of regional transport 2008 experiment, *Bulletin of the American Meteorological Society*, 91, 327–342, 2010.
- Peixoto, J. and Oort, A. H.: The climatology of relative humidity in the atmosphere, *Journal of climate*, 9, 3443–3463, 1996.
- Peter, T., Marcolli, C., Spichtinger, P., Corti, T., Baker, M. B., and Koop, T.: When dry air is too humid, *Science*, 314, 1399–1402, 2006.
- Placzek, G.: Rayleigh-streuung und Raman-effekt, vol. 2, Akad. Verlag-Ges., 1934.

- Popp, P., Marcy, T., Watts, L., Gao, R., Fahey, D., Weinstock, E., Smith, J., Herman, R., Troy, R., Webster, C., et al.: Condensed-phase nitric acid in a tropical subvisible cirrus cloud, *Geophysical Research Letters*, 34, 2007.
- Povey, A., Grainger, R., Peters, D., and Agnew, J.: Retrieval of aerosol backscatter, extinction, and lidar ratio from Raman lidar with optimal estimation, *Atmospheric Measurement Techniques*, 7, 757–776, 2014.
- Reichardt, J.: Optical and geometrical properties of northern midlatitude cirrus clouds observed with a UV Raman lidar, *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 24, 255–260, 1999.
- Riehl, H. et al.: *Introduction to the Atmosphere*, 1972.
- Rodgers, C. D.: *Inverse methods for atmospheric sounding: theory and practice*, vol. 2, World scientific, 2000.
- Sapucci, L. F., Machado, L. A., da Silveira, R. B., Fisch, G., and Monico, J. F.: Analysis of relative humidity sensors at the WMO radiosonde intercomparison experiment in Brazil, *Journal of Atmospheric and Oceanic Technology*, 22, 664–678, 2005.
- Seidel, D. J., Sun, B., Pettey, M., and Reale, A.: Global radiosonde balloon drift statistics, *Journal of Geophysical Research: Atmospheres*, 116, 2011.
- Sica, R. and Haeferle, A.: Retrieval of temperature from a multiple-channel Rayleigh-scatter lidar using an optimal estimation method, *Applied optics*, 54, 1872–1889, 2015.
- Sica, R. and Haeferle, A.: Retrieval of water vapor mixing ratio from a multiple channel Raman-scatter lidar using an optimal estimation method, *Applied optics*, 55, 763–777, 2016.
- Soden, B. J. and Lanzante, J. R.: An assessment of satellite and radiosonde climatologies of upper-tropospheric water vapor, *Journal of Climate*, 9, 1235–1250, 1996.
- Solomon, S., Rosenlof, K. H., Portmann, R. W., Daniel, J. S., Davis, S. M., Sanford, T. J., and Plattner, G.-K.: Contributions of stratospheric water vapor to decadal changes in the rate of global warming, *Science*, 327, 1219–1223, 2010.

- Stohl, A., Wernli, H., James, P., Bourqui, M., Forster, C., Liniger, M. A., Seibert, P., and Sprenger, M.: A new perspective of stratosphere–troposphere exchange, *Bulletin of the American Meteorological Society*, 84, 1565–1574, 2003.
- Tables, S. M.: *RJ List*, 6th rev. ed., Washington, DC, 1951.
- Thies, B. and Bendix, J.: Satellite based remote sensing of weather and climate: recent achievements and future perspectives, *Meteorological Applications*, 18, 262–295, 2011.
- Urban, J.: Satellite sensors measuring atmospheric water vapour, in: *Monitoring Atmospheric Water Vapour*, pp. 175–214, Springer, 2013.
- Vérèmes, H., Payen, G., Keckhut, P., Duflot, V., Baray, J.-L., Cammas, J.-P., de Bellevue, J. L., Evan, S., Posny, F., Gabarrot, F., et al.: A Raman lidar at Maïdo Observatory (Reunion Island) to measure water vapor in the troposphere and lower stratosphere: calibration and validation, *Atmospheric Measurement Techniques*, pp. Under–Review, 2016.
- Wallace, J. M. and Hobbs, P. V.: *Atmospheric science: an introductory survey*, vol. 92, Elsevier, 2006.
- Wandinger, U.: Introduction to lidar, in: *Lidar*, pp. 1–18, Springer, 2005a.
- Wandinger, U.: Raman lidar, in: *Lidar*, pp. 241–271, Springer, 2005b.
- Weber, A.: High-resolution rotational Raman spectra of gases, in: *Raman Spectroscopy of Gases and Liquids*, pp. 71–121, Springer, 1979.
- Weitkamp, C.: *Lidar: range-resolved optical remote sensing of the atmosphere*, vol. 102, Springer Science & Business, 2006.
- Whiteman, D., Melfi, S., and Ferrare, R.: Raman lidar system for the measurement of water vapor and aerosols in the Earths atmosphere, *Applied Optics*, 31, 3068–3082, 1992.
- Zuev, V. V., Gerasimov, V. V., Pravdin, V. L., Pavlinskiy, A. V., and Nakhtigalova, D. P.: Tropospheric temperature measurements with the pure rotational Raman lidar technique using nonlinear calibration functions, *Atmospheric Measurement Techniques*, 10, 315–332, 2017.

## Chapter 2

# Application of the OEM to retrieve temperature from the pure Rotational Raman lidar measurements<sup>1</sup>

### 2.1 Overview

High time and space resolution measurements of atmospheric temperature are necessary to improve our understanding of many atmospheric processes, both dynamical and chemical. Radiosounding is the most widely used method for temperature profiling in the troposphere and lower stratosphere, and has the advantage of operation in most weather conditions, but is typically limited to 2 flights per day. Pure rotational Raman (PRR) lidars have excellent vertical and temporal resolution, and can be combined with vibrational Raman channels to determine relative humidity. Lidar temperature measurements can be assimilated with atmospheric models to improve weather forecasts, as recently demonstrated by Adam et al. (2016).

The traditional Raman lidar temperature retrieval method, introduced by Cooney (1972), uses the ratio of two PRR signals from the Stokes branch which have been corrected for saturation, background and other instrumental effects as required. The PRR spectrum contains

---

<sup>1</sup>A version of this chapter has been published as a discussion paper: Mahagammulla Gamage, S., Sica, R. J., Martucci, G., and Haefele, A.: Retrieval of Temperature From a Multiple Channel Pure Rotational Raman-Scatter Lidar Using an Optimal Estimation Method, *Atmos. Meas. Tech. Discuss.*, <https://doi.org/10.5194/amt-2019-107>, 2019, Creative Commons 4.0 copyright.



two branches: Stokes and anti-Stokes. Both branches have approximately the same intensity and they are positioned symmetrically in wavelength on either side of the excitation line. The traditional Raman lidar temperature retrieval algorithm requires the assumption of an analytic form of a lidar calibration function whose coefficients are usually determined with external measurements, such as radiosondes (Behrendt, 2005). The calibration function is an approximation of the relationship of the signal ratio and temperature and depends on two or more coefficients. Calibration errors exceeding 0.5 K can arise if the calibration data do not cover a sufficient temperature range (Behrendt, 2005).

Of primary importance is in the calibration of the lidar returns to allow absolute temperature measurements. In the traditional method, the ratio of the corrected photo-counts is fit to a set of corresponding temperature data points usually obtained from radiosondes. The application of the Optimal Estimation Method (OEM) for temperature retrievals using Pure Rotational Raman (PRR) lidar measurements has several advantages over the traditional method, including the ability to find temperature without assuming an analytic form of the temperature/count ratio relation. Our OEM retrieval does not use the ratio of the counts. Rather we use a forward model which includes complete physics to describe the raw count profiles. For calibration the ratio of the lidar constants, here from referred to as coupling constant, needs to be determined. The coupling constant can in principle be estimated at a single point, such as a nearby flux tower or surface measurement. Our OEM retrieval has other important benefits over the traditional method, as it can directly retrieve ancillary parameters to temperature, such as geometrical overlap, particle extinction, and the lidar constant. Our OEM is also capable of providing a full uncertainty budget, including both random and systematic uncertainties on a profile-by-profile basis, including the systematic uncertainty introduced in the retrieved temperature by the estimation of the coupling constant. The OEM is an inverse method, and is a standard tool in the retrieval of geophysical parameters from passive atmospheric remote sensing instruments. Recent studies including (Povey et al., 2014; Sica and Haeferle, 2015, 2016; Farhani et al., 2019) have shown that OEM can be used to retrieve atmospheric aerosol, water vapor mixing ratio, middle and upper atmospheric temperature and ozone using lidar measurements.

In Section 2.2 a brief description of the instrument and the measurements used in this study is presented. Section 2.3 presents the development of the PRR lidar equation. The develop-

ment of a forward model for application of the OEM to PRR temperature retrieval is given in Section 2.4. The OEM-retrieved temperature results from the PRR measurements for different atmospheric conditions are shown in Section 2.5. A discussion of these results is presented in Section 2.6 followed by conclusions.

## 2.2 The Raman Lidar for Meteorological Observations

PRR measurements from the Raman Lidar for Meteorological Observations (RALMO), located in Payerne, Switzerland (46°48' N, 6°56' E) are used for the OEM temperature profiling. RALMO is a fully automated lidar built at the École Polytechnique Fédérale de Lausanne and operated by MeteoSwiss (Dinoev et al., 2013). It is dedicated to operational meteorology, validating models and satellite measurements, and climate studies. RALMO has been operating nearly continuously since 2008, with an average data availability of 50%. Data gaps are due to rain and low clouds (approximately 30% of the time), maintenance (1 - 2 days per month) and other occasional technical problems. RALMO consists of a frequency tripled, Q-switched Nd:YAG laser of 354.7 nm producing up to 400 mJ emission energy at 30 Hz repetition rate. The pulse duration is 8 ns. The laser is operated at 300 mJ energy per pulse to extend the lifetime of the flash lamps from 20 to approximately 60 million shots. The RALMO receiver uses four parabolic mirrors each with 1 m focal length and 30 cm diameter, and it is fiber coupled to a two stage grating polychromator. The data acquisition system consists of photo-multipliers and analog/digital transient recorders from Licel. The system obtains a measurement by adding together 1800 laser shots (every minute) at a vertical resolution of 3.75 m. For a detailed description of the lidar and a detailed validation of the temperature measurements the reader is referred to Dinoev et al. (2013).

The returns of the Raman-shifted backscatter arising from rotational energy state transitions of nitrogen and oxygen molecules due to the excitation at the laser wavelength at 354.7 nm are detected in analog and photon counting mode. The high quantum number channel (JH) of RALMO is assigned to the backscattered signals from the energy exchange that occurs in the high quantum states for both the Stokes (355.77-356.37 nm) and anti-Stokes (353.07-353.67 nm) branches. The low quantum number channel (JL) is assigned to the signals from

the energy exchange occurring in the low quantum states in the Stokes (355.17-355.76 nm) and anti-Stokes (353.67-354.25 nm) branches.

## 2.3 The PRR lidar equation

The backscattered PRR signal is given by the Raman lidar equation:

$$N_{RR,t}(z) = \frac{C_{RR}}{z^2} O(z) n(z) \Gamma_{atm}^2(z) \left( \sum_{i=O_2, N_2} \sum_{J_i} \eta_i \left( \frac{d\sigma}{d\Omega} \right)_\pi^i (J_i) \right) + B_{RR}(z) \quad (2.1)$$

where the true backscattered PRR signal  $N_{RR,t}$ , is a function of height  $z$ ,  $C_{RR}$  is the lidar constant,  $n(z)$  is the number density of the air molecules,  $O(z)$  is the geometrical overlap,  $\eta_i$  is the volume mixing ratio of nitrogen and oxygen,  $\Gamma_{atm}(z)$  is the atmospheric transmission,  $\left( \frac{d\sigma}{d\Omega} \right)_\pi^i (J_i)$  is the attenuated differential backscatter cross-section for each RR line  $J_i$  and  $B_{RR}(z)$  is the background of the measured signal. For different lidar systems the background can either be a constant or vary with height. Since air below 80 km is a constant mixture of oxygen and nitrogen,  $\eta_i$  is a constant. The lidar constant  $C_{RR}$  depends on the number of transmitted photons, detector efficiency, and the area of the telescope.

The attenuated differential backscatter cross section for Stokes and anti-Stokes line pairs of equal quantum number of the PRR spectrum is expressed as (Penney et al., 1974):

$$\begin{aligned} \left( \frac{d\sigma}{d\Omega} \right)_\pi^i (J) &= \frac{112}{15} \cdot \frac{\pi^4 g_i(J) h c B_{0,i} (v_o + \Delta v_i(J))^4 \zeta_i^2}{(2J_i + 1)^2 k T} \\ &\times (X^+(J) \tau^+(J_i) + X^-(J) \tau^-(J_i)) \exp \left( \frac{-E_{rot,i}(J)}{k T} \right) \end{aligned} \quad (2.2)$$

where for the Stokes branch

$$X^+(J) = \frac{(J+1)(J+2)}{2J+3} \quad \text{for } J = 0, 1, 2, \dots \quad (2.3)$$

and for the anti-Stokes branch

$$X^-(J) = \frac{J(J-1)}{2J-1} \quad \text{for } J = 2, 3, 4, \dots \quad \text{and} \quad X^-(J) = 0 \quad \text{for } J = 0, 1. \quad (2.4)$$

$\tau^+(J_i)$  and  $\tau^-(J_i)$  are the transmissions of the receiver for the Stokes and anti-Stokes lines  $J_i$ , respectively.  $g_i(J)$  is the statistical weighting factor, which depends on the nuclear spin  $I_i$  for each atmospheric constituent,  $h$  is Planck's constant,  $c$  is the velocity of light,  $k$  is Boltzmann's constant,  $B_{0,i}$  is the ground-state rotational constant,  $\nu_0$  is the frequency of the incident light, and  $\zeta_i$  is the anisotropy of the molecular polarizability. The rotational energy  $E_{rot,i}(J)$  for each Stokes and anti-Stokes branch is estimated based on the assumption of a homonuclear diatomic molecule in the quantum state  $J$  for nitrogen and oxygen molecules with no electronic momentum coupled to the scattering (Behrendt, 2005).

The response of photomultiplier tubes used as detectors in lidar systems and operated in the digital mode can become nonlinear at high count rates. In the case of RALMO, the true and observed counts are related by the equation for non-paralyzed systems where  $\gamma$  is the counting system dead time:

$$N_{observed} = \frac{N_{true}}{1 + N_{true}\gamma}. \quad (2.5)$$

The dead time of the counting system is the time taken after measuring a photon, before the detector is able to accurately measure another incident photon.

## 2.4 Application of the OEM for PRR temperature retrieval

### 2.4.1 Brief review of the optimal estimation method

The OEM is an inverse method that uses the measurements  $\mathbf{y}$  to estimate the state (retrieval) variables  $\mathbf{x}$  of a system via a forward model. The forward model  $\mathbf{F}$  contains all the atmospheric and instrumental physics that describe the measurements. The forward model can include model parameters  $\mathbf{b}$ , which are assumed and not retrieved, and their effect on the retrieved quantity uncertainties can be calculated.

The measurements are related to the forward model by:

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \boldsymbol{\epsilon} \quad (2.6)$$

where  $\boldsymbol{\epsilon}$  represents measurement noise. Under the assumption that all parameters have Gaus-

sian probability density functions Bayes theorem can be applied to determine the cost function,

$$\text{cost} = [\mathbf{y} - \mathbf{F}(\mathbf{x}, \mathbf{b})]^T \mathbf{S}_y^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x}, \mathbf{b})] + [\mathbf{x} - \mathbf{x}_a]^T \mathbf{S}_a^{-1} [\mathbf{x} - \mathbf{x}_a], \quad (2.7)$$

where  $\mathbf{x}_a$  is the *a priori*,  $\mathbf{S}_y$  is the measurement covariance, which describes the random measurement uncertainty and  $\mathbf{S}_a$  is the *a priori* covariance. The cost function measures the goodness of fit for a solution, and for good models the cost is on the order of unity. For non-linear forward models, the Marquardt-Levenberg method can be used iteratively to minimize the cost of the retrieval (see Section 5.7 in Rodgers (2000) for details).

### 2.4.2 The forward model for a PRR lidar

The forward model describes the measurement as a function of both the state of the atmosphere and instrumental parameters. The core of our forward model is the lidar equation as presented in Section 2.3. It is called 4 times to generate the measurements corresponding to high and low quantum numbers, i.e. JH and JL, with digital and analog detection. Analog detection is assumed to be linear and hence the saturation equation (Equation 2.5) is not applied.

We use Eq. 2.1 as the forward model after replacing the density  $n(z)$  in terms of pressure and temperature.

$$n(z) = \frac{P(z)}{kT(z)} \quad (2.8)$$

where  $k$  is the Boltzmann constant. The pressure required in the forward model obtained from an ancillary measurements, from an atmospheric model, or can be replaced as a function of temperature by assuming hydro-static equilibrium (Eq. 1.36).

Moreover, there are three different ways density can be included in the forward model.

1. Density is replaced by ideal gas law given in Eq. 2.1 (FM1). The forward model will include the temperature dependency of both the differential cross-section and density.
2. Forward model as given in Eq. 2.1 in terms of density. Thus, the temperature dependency of the density is not included (FM2). The only temperature dependent term in the forward model is the differential cross-section.

3. Replacing the density in the forward model given in Eq.2.1 using the ideal gas law and replacing the differential cross-section term as a model parameter (FM3). Thus, we need to calculate the differential cross-section using ancillary measurements. This third type of the forward model will be similar to the forward model used by Sica and Haeferle (2015), and should be less sensitive to temperature since the rotational temperature information is not used. Cooney (1972) used this model to retrieve temperature from Rayleigh lidar measurements.

To find out the forward model with maximum sensitivity to the measurements we compared the temperature Jacobians for each forward model. It was evident from the Jacobians (Appendix A figures) that FM1 has the most sensitivity to the temperature. FM2 only include the temperature dependency of the differential cross-sections terms similar to the traditional rotational temperature method (as when the ratio of PRR measurements is considered the temperature dependency of the density cancels). In FM3 the only temperature dependent term is density. For FM3 we assume the differential cross sections are model parameters, discarding the temperature dependency of the Raman measurements. Therefore, neither of FM2 or FM3 are complete or do not have the maximum sensitivity to retrieve temperature from the Raman back-scatter measurements. We recommend to use a forward model with the complete physics as given in Eq. 2.1 with the density replaced with Eq. 2.8 and pressure replaced with hydro-static equilibrium.

As stated earlier, in our work we use the most sensitive forward model, FM1; Eq. 2.1 where the density is replaced by ideal gas law. The pressure,  $P(z)$ , and temperature,  $T(z)$ , required can be taken from either a radiosonde measurement or an atmospheric model. The background noise,  $B_{RR}$ , is in general a function of height,  $z$ , but is constant with height for RALMO. Unlike all the other existing forward models for lidar except Povey et al. (2012) (which was designed specifically to determine overlap) we retrieve  $O(z)$  the geometrical overlap function in addition to temperature.

The atmospheric transmission,  $\Gamma_{atm}(z)$  in Eq. 2.1, includes both molecular and particle scattering.

$$\Gamma_{atm}(z) = \exp \left[ - \int_0^z (\alpha_{mol} + \alpha_{par}) dz \right] \quad (2.9)$$

where  $\alpha_{mol}$  is the molecular extinction coefficient and  $\alpha_{par}$  is the particle extinction coefficient. The molecular extinction can be expressed using the Rayleigh cross section  $\sigma_{Ray}$  and air density  $n_{air}$  as:

$$\alpha_{mol} = \sigma_{Ray} \cdot n_{air} \quad (2.10)$$

where  $\sigma_{Ray}$  is calculated using the expressions given by Nicolet (1984).

For each channel the subscript **RR** is replaced by JL and JH, the high and low quantum number PRR channels. Then  $C_{RR}$ ,  $B_{RR}$  and  $J_i$  then have different values.

RALMO detects multiple Stokes and anti-Stokes lines from both nitrogen and oxygen PRR spectrum. Therefore, to determine the attenuated backscatter cross-section in the forward model we require knowledge of the exact number of quantum states detected by each the RALMO PRR channel. From the JH and JL channel characteristics we can calculate the range of frequency shifts for each channel relative to the elastic wavelength 354.7 nm. Then using the equations given by Herzberg (2013) we can determine the quantum numbers  $J_i$  for both nitrogen and oxygen molecules. This calculation process is repeated to determine the  $J_i$  numbers for the JL channel of the RALMO. The summary of the findings are given in Table 2.1.

Table 2.1: Return PRR wavelengths detected by the RALMO and the respective quantum lines from nitrogen and oxygen PRR spectrums.

Channel		Nitrogen	Oxygen
JL	Detected wavelengths	355.17-355.76 nm	353.67-354.25 nm
	Quantum lines (Stokes and anti-Stokes)	3,4,5,7,8,9	5,7,9,11,13
JH	Detected wavelengths	355.77-356.37 nm	353.07-353.67 nm
	Quantum lines (Stokes and anti-Stokes)	10,11,12,13,14,15	15,17,19,21

In order to establish absolute calibration, we define the coupling constant  $R$  as the ratio of the lidar constants  $C_{JL}$  and  $C_{JH}$

$$R = \frac{C_{JH}}{C_{JL}} \quad (2.11)$$

and use the substitution  $C_{JH} = RC_{JL}$ . The coupling constant is height-independent and can be determined with no assumptions at, if desired, a single altitude using the following equation derived from Eq. 2.1.

$$R = \left( \frac{N_{t,JH} - B_{JH}}{N_{t,JL} - B_{JL}} \right) \bigg/ \left( \frac{\sigma_{JH}(T, z)}{\sigma_{JL}(T, z)} \right). \quad (2.12)$$

The differential cross section terms are determined by applying temperature from a coinci-

dent reference temperature, typically from a radiosonde. For a well designed lidar system the coupling constant should be stable over weeks. Unlike the fitting of an analytic calibration function to the data as in the traditional method,  $\mathbf{R}$  can be estimated at a specific height or range of heights, which can be over a narrow range of temperature without introducing extrapolation errors. We extensively tested this assertion using both synthetic and real measurements. The results show that the estimation of  $\mathbf{R}$  is indeed height independent. The value of  $\mathbf{R}$  is only affected by the measurement noise. Hence, we recommend using a range of heights or a specific height where the photocounts have a high signal-to-noise ratio.

Using  $\mathbf{R}$  in the forward model allows us to retrieve only one lidar constant, while constraining the two channels to vary so as to satisfy Eq. 2.11. We will see in the next section that any variations or uncertainty in the determination of  $\mathbf{R}$  introduces an uncertainty on the order of 0.2 K to the retrieved temperature profile.

The retrieval parameters (Table 2.2) in our OEM algorithm are temperature, background signals (including photo multiplier shot noise, sky background, and offset for analog channels), the lidar constants, dead times of the digital photon counting systems, geometrical overlap, and particle extinction as a function of height. In OEM we can retrieve parameters on a height grid where the resolution can be same or different than the vertical resolution of the height grid that the measurements obtained. If the retrieval grid is coarser than the measurement grip we use linear interpolation on retrieved quantities when they are required in the forward model.

Table 2.2: Values and associated uncertainties for the OEM retrieval and forward model parameters.

Parameter	Value	Standard Deviation
<b>Measurements</b>		
Digital	Measured	Poisson Statistics
Analog	Measured	Auto Covariance Method
<b>Retrieval Parameters ( <i>a priori</i> )</b>		
Temperature	US Standard Model	35 K
geometrical overlap Function	Estimated using the forward model and measurements	50% below and at transition height 10 <sup>-3</sup> above transition height
Particle Extinction	Estimated using measurements	10 <sup>-6</sup> km <sup>-1</sup> below and at transition height 50% above transition height
Lidar Constants (analog/digital)	Estimated using the forward model	100%
Digital Background Noise	Mean above 50 km	Standard Deviation above 50 km
Analog Background Noise	Mean above 50 km	nighttime- Standard Deviation above 50 km daytime- normalized standard deviation above 50 km
Dead Time	Empirical fitting	10%
<b>Forward Model Parameters</b>		
Pressure	Radiosonde	30 Pa
Coupling Constants (analog/digital)	Estimated measurements and sonde temperature	Standard deviation of the coupling constants over a height range
Air density	Radiosonde	1%



### 2.4.3 Implementation of the RR temperature retrieval

The OEM solver in the Qpack software package is used for our retrieval (Eriksson et al., 2005). The solver requires the following as inputs: the measurements from each lidar channel and their covariances, *a priori* values for the retrieval parameters and their covariances, model (*b*) parameters and their covariances, and the Jacobians of the forward model.

The lidar measurements from the digital channels follow Poisson counting statistics in the range where the counts are linear. Thus, the measurement variance  $S_y$  is equal to the number of photons in each height bin, assuming no correlation between the height bins (that is, the off-diagonal terms in the  $S_y$  matrix are zero). The auto-correlation function method of Lenschow et al. (2000) is used to estimate the measurement covariance of the analog and digital measurements in the non-linear region. For both analog and digital channels, the *a priori* backgrounds and their variances are taken as the mean and the variance of the measurements above 50 km height.

The U. S. Standard Atmosphere (Atmosphere, 1976) model temperature profile is normalized to the surface temperature from the coincident sonde temperature, and then used as the *a priori* temperature profile in our retrievals. Due to the high variability of the temperature, a standard deviation of 35 K for all heights is used to specify the covariance matrix for *a priori*. Other choices of *a priori* temperature profile are possible, but as an operational, fully automated lidar system RALMO retrievals must be performed automatically every 30 min, so the choice of the US Standard Model with this covariance simplifies this procedure. As discussed in Eriksson et al. (2005), the elements of the retrieval and model parameters are often correlated, and some of the covariance matrices should have off-diagonal elements. Off-diagonal elements are parametrized with the correlation length and an appropriate analytical function describing the decay of the correlation. For this study, we used a tent function with a 1 km correlation length for temperature retrievals (Eriksson et al., 2005).

Molecular and particle extinction terms occur in the atmospheric transmission term of Eq. 2.9. An *a priori* particle extinction profile is estimated based on the following expression:

$$\alpha_{par} = LR \cdot \beta_{par} = LR \cdot \beta_{mol} \cdot (\mathfrak{R}_\beta - 1) \quad (2.13)$$

where  $LR$  is the lidar ratio,  $\beta_{par}$  is the particle backscatter coefficient.  $\beta_{par}$  can be related to the backscatter ratio  $\mathfrak{R}_\beta$  as (Whiteman, 2003):

$$\mathfrak{R}_\beta = \frac{(\beta_{mol} + \beta_{par})}{\beta_{mol}} \quad (2.14)$$

The backscatter ratio  $\mathfrak{R}_\beta$  is estimated using the RALMO PRR and elastic measurements. Bucholtz (1995) gives a method for calculating  $\beta_{mol}$  using pressure, temperature and Rayleigh cross sections. The Rayleigh extinction cross sections required for  $\beta_{mol}$  estimation are computed using the formula of Nicolet (1984). Calculated Rayleigh extinction cross sections are also used to estimate the air density profile used as a  $b$  parameter in the forward model, assuming an uncertainty of 1% for the standard deviation.

The lidar ratio  $LR$  is chosen based on the  $\mathfrak{R}_\beta$  values for the given measurements. Typically  $\mathfrak{R}_\beta$  values inside clouds are greater than 2. Thus, for this study the height at which  $\mathfrak{R}_\beta$  is first equal to 2 is considered as the height of the cloud base or the height of an aerosol layer (cloud/aerosol layer base height). The cloud/aerosol layer base height is later used to determine the transition height that constrains the range of the geometrical overlap and the particle extinction retrievals. In clear sky conditions (that is if  $\mathfrak{R}_\beta$  does not exceed 2),  $LR$  is assumed to be 80 sr inside the boundary layer and 50 sr elsewhere. In cloudy conditions,  $LR$  is assumed to be 20 sr within the clouds present below 6 km. If the cloud is above 6 km,  $LR$  is assumed to be 15 sr within the cloud. These choices for lidar ratios are taken from Ansmann et al. (1992a) and Pappalardo et al. (2004). Accurate  $LR$  is not crucial, as it is used to estimate an *a priori* particle extinction profile. However, we can calculate a  $LR$  profile using the OEM-retrieved  $\alpha_{par}$  and compare it with the initially chosen  $LR$  values to evaluate how good a choice of the initial value is.

The effect of geometrical overlap and particle extinction on the signals are strongly linearly dependent and hence retrieving both parameters simultaneously with the given data channels is not possible unless at least one of the effects is highly constrained. In this work we assume that particle extinction is well known from the backscatter ratio outside clouds and that overlap is well known above the height of full overlap, i.e. above 6 km. Hence we define a transition height at 6 km or cloud base height, whatever is lower, below which particle extinction is

retrieved and above which overlap is retrieved. The *a priori* overlap function is estimated from the measurements in clear sky conditions with little effect due to particles. A 50% standard deviation is used for geometrical overlap below the transition height and a constant value of  $10^{-3}$  is used above this height, constraining the geometrical overlap to the *a priori* above the transition height. For particle extinction, a small standard deviation of  $10^{-6}\text{km}^{-1}$  is used below the transition height, but a 50% standard deviation is used above this height, allowing the OEM to retrieve exclusively the particle extinction. The *a priori* covariance matrices for both particle extinction and geometrical overlap are determined using a tent function with a 100 m correlation length.

*A priori* lidar constants for the JL analog and JL digital channels are estimated by fitting the measurements generated using the sonde temperature and pressure used in the forward model to the PRR measurements. For analog measurements, the fitting range is between 1.5 to 2 km height. For digital measurements with clear conditions or cloud/aerosol presence above 8 km, 6 to 8 km is used as the fitting range, to insure the photocounts are linear. If the digital measurements contain a cloud or aerosols in the geometrical overlap region, a fitting range below this is used, typically 3.5 - 4 km height. The fitting uncertainty for each analog and digital lidar constants is used as the variance of the *a priori* lidar constant.

The *a priori* dead times for the two digital photon counting systems are considered to be 3.8 ns, consistent with estimations from previous studies for RALMO and with values specified by the manufacturer (Sica and Haefele, 2016, 2015; Dinov et al., 2010). The uncertainty in the dead time is taken as 10%. Coincident radiosonde pressure profile are used assuming a 10% standard deviation. The coupling constants for analog ( $R_a$ ) and digital ( $R$ ) channels are estimated by fitting the ratio of PRR measurements with the ratio of the differential cross section (Eq. 2.12). The coupling constants are estimated using the same fitting range as the lidar constants. Table 2.2 gives a summary of the parameters and associated uncertainties of the retrieval and  $\mathbf{b}$  parameters used in the forward model.

## 2.5 Results from the temperature retrieval

We present four different measurement situations which demonstrate the robust nature of our OEM temperature retrieval. Details of each case study are given in Table 2.3. The RALMO measurements used in the retrievals are added in time over 30 min and to 15 m in height. Analog measurements are used from the surface to 6 km height, while digital measurements are used from 4 to 28 km. The retrieval grid has a vertical resolution of 60 m at all heights. For all the cases given in Table 2.3 we used radiosonde measurements that coincide within 1 hour of the lidar measurements for comparison purposes, to estimate the required *a priori* information, and to determine the forward model *b* parameters (Table 2.2).

Table 2.3: Details of the 4 cases in different sky conditions we present to demonstrate the flexibility of our OEM temperature retrieval.

Case	Date	Time (UT)	Sonde Launch (UT)	Sky Condition
1	20110909	2300-2330	2200	Clear-nighttime
2	20110910	1100-1130	1100	Clear-daytime
3	20110705	2300-2330	2300	Cirrus Cloud ( $\sim 6$ km)-nighttime
4	20110621	2300-2330	2300	Lower Cloud (4 km)-nighttime

The traditional temperature profiles shown are calculated using count profiles consisting of glued analog and digital measurements which are corrected for non-linearity and background before processing. The vertical resolution of the traditional temperature profiles is 30 m at the lowest heights, increasing to 400 m by the upper heights to decrease the magnitude of the statistical uncertainty. A calibration function linear in  $1/T$  is used and the two coefficients have been determined with radiosonde data using the altitude range from surface to 10 km. Hence, for this comparison the temperature profile is smoother than the OEM retrieved temperature profile. The change in vertical resolution is based on the random uncertainty of the temperature profile at each height. The vertical resolution is decreased until the temperature uncertainty becomes less than a threshold value, set here as 1 K.

### 2.5.1 Case 1: Nighttime with clear conditions

Fig. 2.1 shows the RALMO 30 min coadded count measurements in the four PRR channels for case 1 given in Table 2.3. Analog signals suffer day and night from a high electrical offset which

becomes dominant above 5 km. On the other hand, the analog signals stay linear over the entire encountered range. The digital signals on the other hand have a very low system background but become saturation above 2 MHz. Fig. 2.2 shows the residuals, which are defined as the difference between the forward model and the measurements. For a good retrieval with cost on the order of unity, the residuals (blue curve) should be on the order of the standard deviation of the measurement noise (red curve), and indeed this is the case, hence the forward model is not over-fitting the measurements (e.g. cost much less than unity).

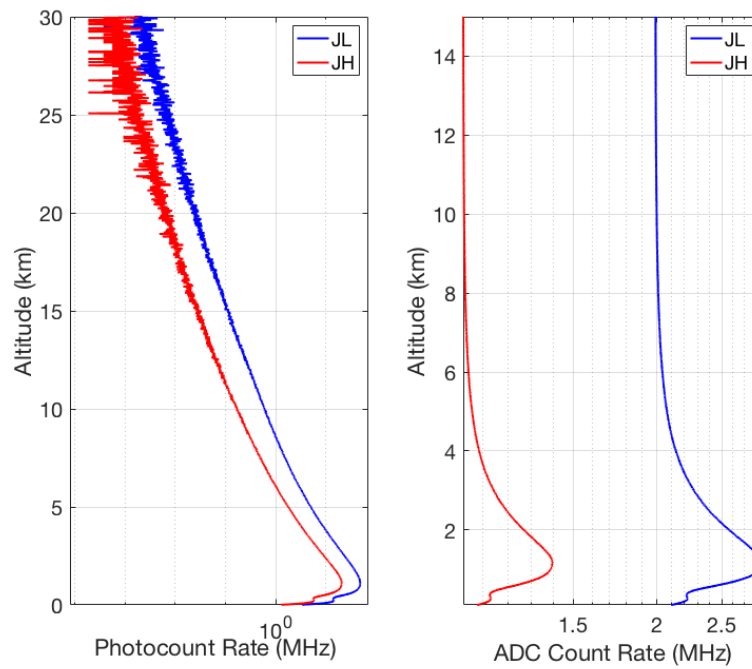


Figure 2.1: Count rate for 30 min of coadded RALMO measurements from 2300 UT on 09 September 2011, a clear night. Left panel: digital channels (blue curve, JL; red curve, JH). Right panel: analog channels.

The averaging kernels of temperature and the vertical resolution of the retrievals are shown in Fig. 2.3. The area of the averaging kernels is defined as the response function of the retrievals and is close to 1.0 below 24 km, meaning that contribution of the *a priori* temperature profile to the retrieved temperature is negligible (Rodgers, 2000). With decreasing signal-to-noise ratio the measurement response quickly decreases above about 27 km (Fig. 2.3). The full-width at half-maximum of the averaging kernels is the retrieval's vertical resolution (Fig. 2.3, right

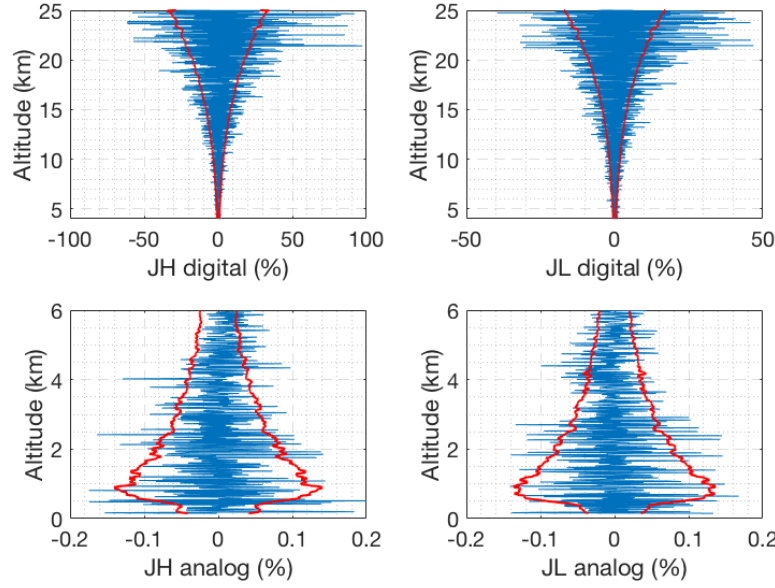


Figure 2.2: Difference between the forward model and clear nighttime RALMO measurements on 09 September 2011 for the four channels (blue). The red curves show the standard deviation of the measurements noise.

panel). The vertical resolution of the retrieval is smaller than the resolution of the retrieval grid above about 10 km. We consider the height at which response function decreases to 0.9 as the cutoff height for the OEM retrievals as at this level the contribution of the *a priori* temperature profile is about 10%.

Figure 2.4 shows a comparison of the OEM-retrieved temperature profile with coincident sonde temperature and temperature obtained using the traditional method. The traditional method profile from the RALMO glued (digital and analog) measurements provided by MeteoSwiss has a vertical resolution of 30 m, below 12.5 km and 400 m above this height, and is interpolated to the same retrieval grid as the OEM and shown in black. The change in vertical resolution and the cutoff height of the traditional temperature retrieval are based on temperature uncertainty thresholds.

As shown in the right panel of Fig. 2.4, the temperature difference between OEM-retrieved and sonde temperature (blue curve) fits inside the statistical uncertainty of the OEM-retrieved temperature. Temperature from the traditional method follow the same trend as the sonde and the OEM-retrieved temperature.

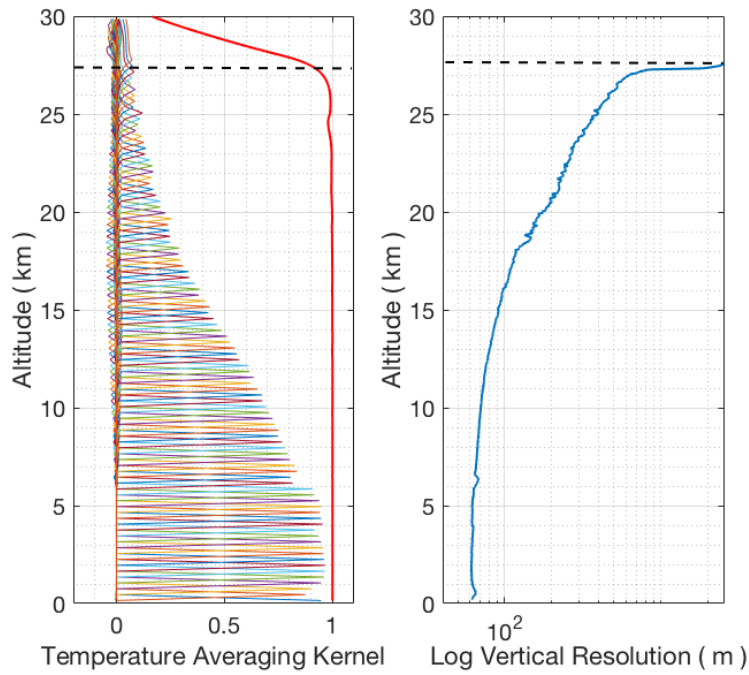


Figure 2.3: Averaging kernels (left) and vertical resolution (right) for temperature retrievals from the clear nighttime RALMO measurements on 09 September 2011. The area of the averaging kernels at each height, the response function, is the red solid line. The horizontal dashed line shows the height below which the retrieval is due primarily to the measurement and not the *a priori* temperature profile. For clarity averaging kernels for every fifth height bin of the retrieval grid are shown.

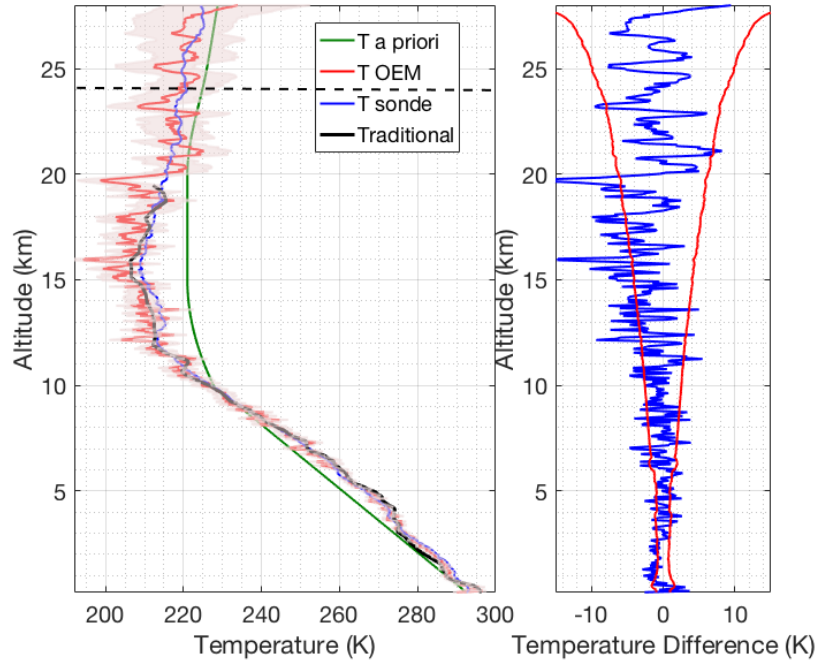


Figure 2.4: Left panel: Retrieved temperature profile and the statistical uncertainty (red curve and shaded area) using the OEM from the clear nighttime RALMO measurements on 09 September 2011. The blue curve is the radiosonde measurement. The sonde was launched at 2200 UT. The green curve is the *a priori* temperature profile used by the OEM. The black curve shows the temperature retrieved using the traditional analysis method from glued analog and digital signals. The horizontal dashed line shows the height below which the retrieval is due primarily to the measurement and not the *a priori* temperature profile. Right panel: The blue curve shows the temperature difference between OEM-retrieved temperature and the sonde temperature. The red curves in the figure show the statistical uncertainty of the OEM temperature.



The OEM provides a complete uncertainty budget of both random and systematic uncertainties (Figure 2.5). Uncertainties due to the Rayleigh cross section is in the order of 0.01 mK. Pressure accounts for up to 0.1 K below 10 km and up to 0.7 K up to 25 km and is a non-negligible source of uncertainty in the stratospheric part of the retrieval. Note that this error could be reduced by choosing a better pressure profile. The uncertainty due to the analog coupling constant  $R_a$  is in the order of 0.07 K up to 4 km and the uncertainty due to digital coupling constant  $R$  is 0.15 K in 4-7 km height range and less than 0.1 K everywhere else. The largest contribution to the temperature uncertainty is from measurement noise, which increases with height.

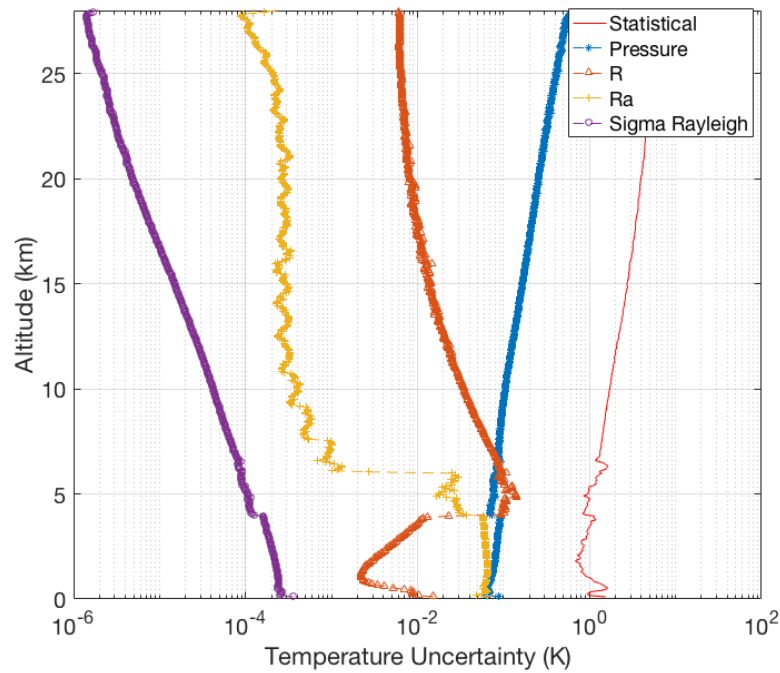


Figure 2.5: Random uncertainties (red curve) and systematic uncertainties due to the forward model parameters for the temperature retrievals from the clear nighttime RALMO measurements on 09 September 2011. The systematic uncertainties included are Rayleigh-scatter cross section (purple dot-dash curve),  $R$  digital coupling constant (orange triangle-dash curve),  $R_a$  analog coupling constant (yellow cross-dash curve), and pressure (blue dot-dash curve).

Figure 2.6 shows the OEM-retrieved geometrical overlap function for the RALMO PRR channels. It illustrates that the overlap retrieval is constrained to be equal to 1 above the transition height (6 km), above which the extinction coefficient is retrieved (see Sec. 2.4.3).

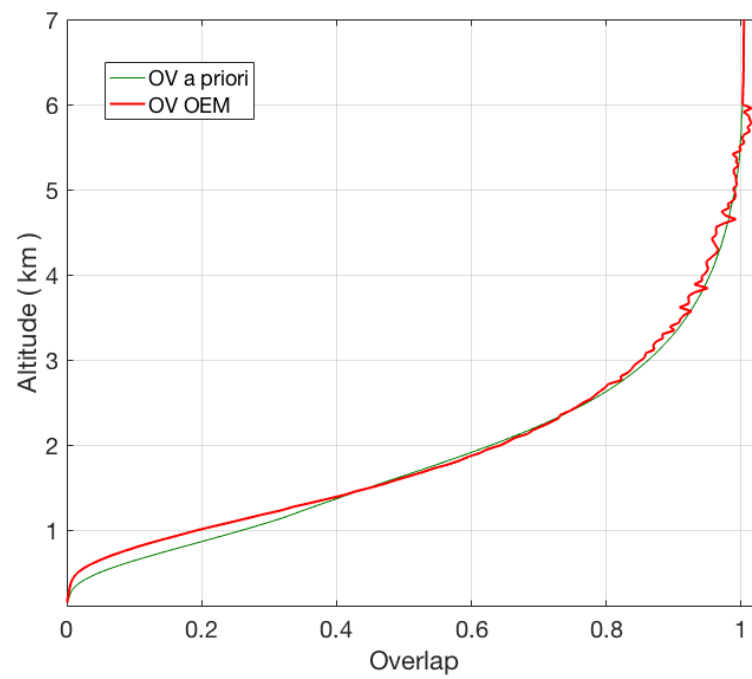


Figure 2.6: Retrieved geometrical overlap function (red curve) from the clear nighttime RALMO measurements on 09 September 2011 compared to the *a priori* geometrical overlap function (green curve).

### 2.5.2 Case 2: Daytime with clear conditions

The retrieval setup for second case study, which is a daytime measurement (Table 2.3), is identical to the one used for nighttime. The major difference between daytime and nighttime retrievals is the large solar background in the measurements, which is evident in the measurements of the digital PRR channels (Fig. 2.7).

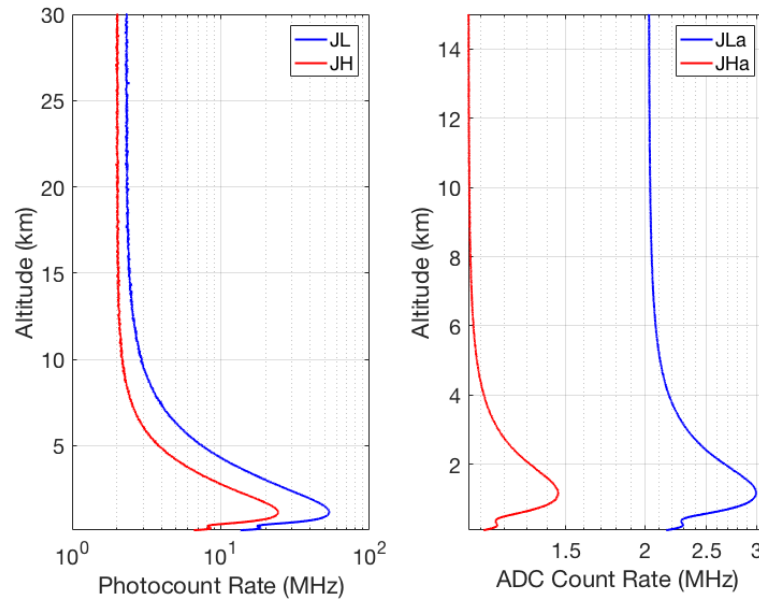


Figure 2.7: Count rate for 30 min of clear coadded RALMO measurements from 1100 UT on 10 September 2011. Left panel: digital channels (blue curve, JL; red curve, JH). Right panel: analog channels.

The residuals are unbiased and fall within the limits of the measurement standard deviation (Fig. 2.8). This result confirms the capability of our forward model in daytime conditions. As shown in Fig. 2.9, the vertical resolution (right panel) is the same as the retrieval grid below 13 km where response function (left panel) is equal to 0.9. The vertical resolution starts to deviate from the retrieval grid as the response function decreases and the background starts to dominate the digital channels.

Similar to the clear nighttime case, the OEM-retrieved temperature agree with the sonde temperature within the statistical uncertainty (Fig. 2.10). It is also evident for this specific case study that the temperature from the traditional method deviate from the sonde temperature more

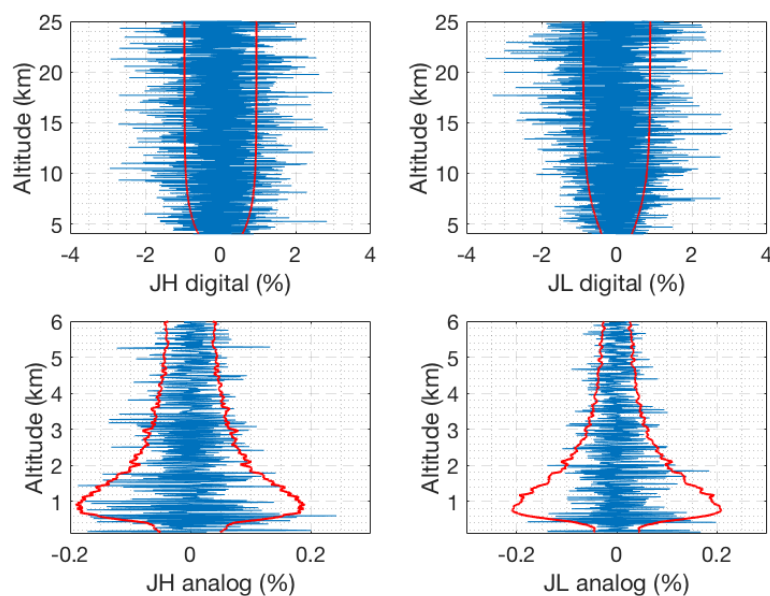


Figure 2.8: Difference between the forward model and the clear daytime RALMO measurements on 10 September 2011 for the four channels (blue). The red curves show the standard deviation of the measurements.

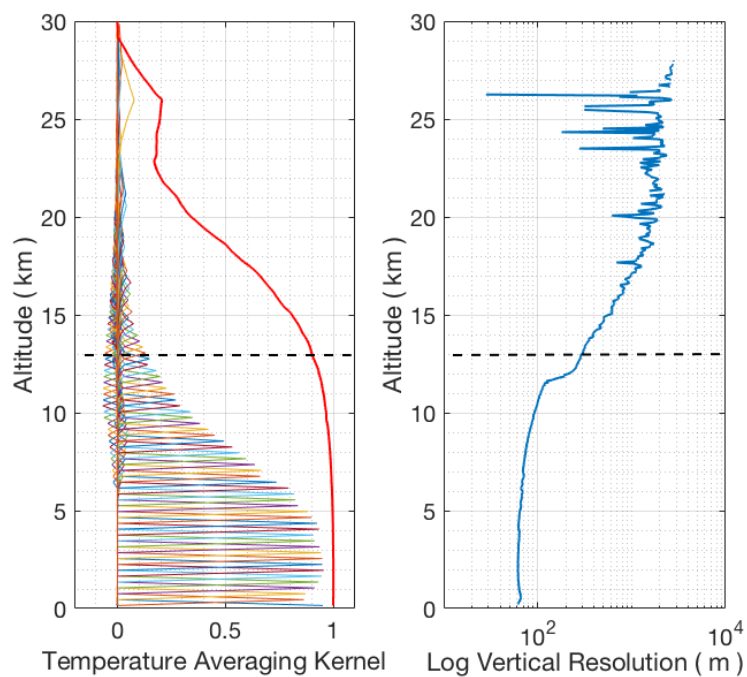


Figure 2.9: Same as Fig. 2.3 but for 10 September 2011.

than the OEM retrieved temperature. This difference compared to the traditional method is due to the fact that the OEM adapts the vertical resolution in an optimal way as a function of height while the traditional method is, constrained by the filter employed to a specific signal-to-noise ratio of the measurements from both channels.

The analog coupling constant  $R_a$  uncertainty of the temperature from the daytime measurements are also in the order of 0.07 K and the uncertainty due to digital coupling constant  $R$  is less than 0.2 K for all heights.

The retrieved geometrical overlap function for the clear daytime case (not shown) agrees with the geometrical overlap retrieved for the nighttime case within 10% statistical uncertainty.

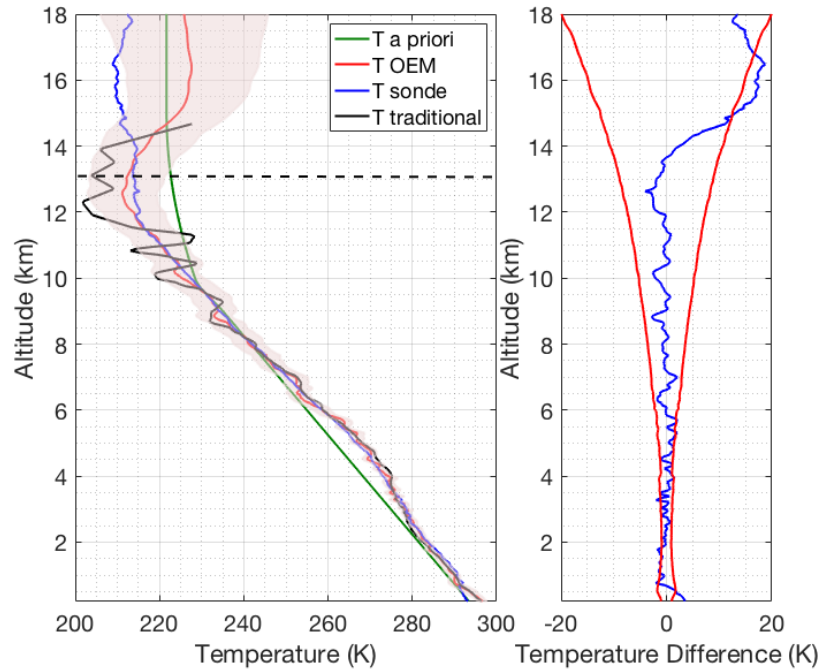


Figure 2.10: Same as Fig. 2.4 but for 10 September 2011.

### 2.5.3 Case 3: Nighttime with cirrus cloud

The third case (details are given in Table 2.3) features a cirrus cloud at 6 km height (Fig. 2.12). The retrieval setup is identical to the previous cases, as the cloud base is above the height of

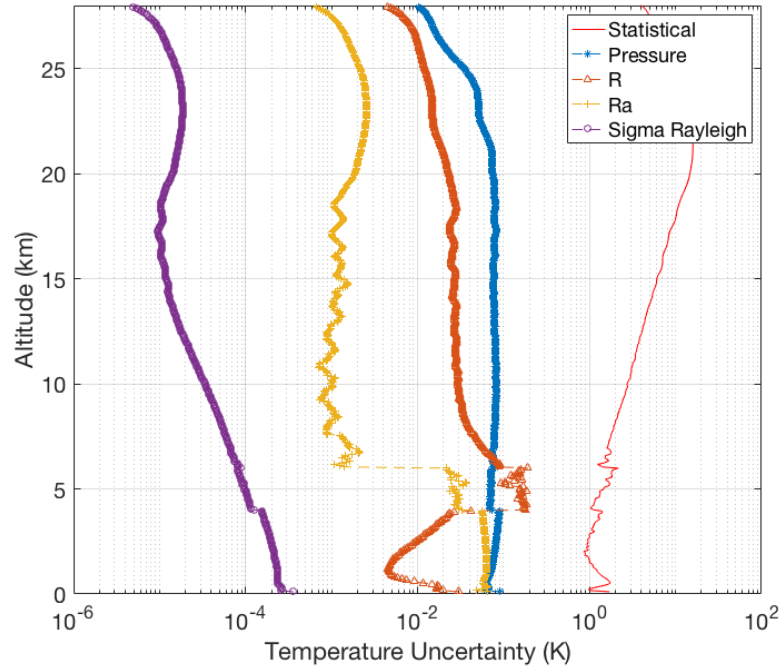


Figure 2.11: Same as Fig. 2.5 but for 10 September 2011.

full geometrical overlap of the transmitter and receiver. The *a priori* profile of the particle extinction coefficient is derived from the back-scatter ratio assuming a lidar ratio of 15 sr.

The residuals (Fig. 2.13) are unbiased and fall within the square root of the measurement variance. This is also true for the altitude range of the cirrus cloud demonstrating that the particle extinction coefficient was determined correctly. The response function (left panel, Fig. 2.14) decreases to the 0.9 cutoff value at about 23.5 km, clearly lower than the clear-sky nighttime case because of the attenuation of the cirrus cloud. Similar to the two previous cases, the OEM-retrieved temperature agree with the sonde temperature within the statistical uncertainty of the OEM retrieved temperature (Fig. 2.15).

The retrieved geometrical overlap function from the measurement with the cirrus cloud (not shown) agrees within 10% uncertainty with the geometrical overlap functions retrieved from the measurement with clear sky conditions, as the cloud is above the region of complete geometrical overlap. The red curve in the first plot in Fig. 2.16 shows the OEM-retrieved particle extinction and the green curve is the *a priori* particle extinction estimated using the RALMO PRR and elastic measurements, assuming a lidar ratio for cirrus clouds in order to

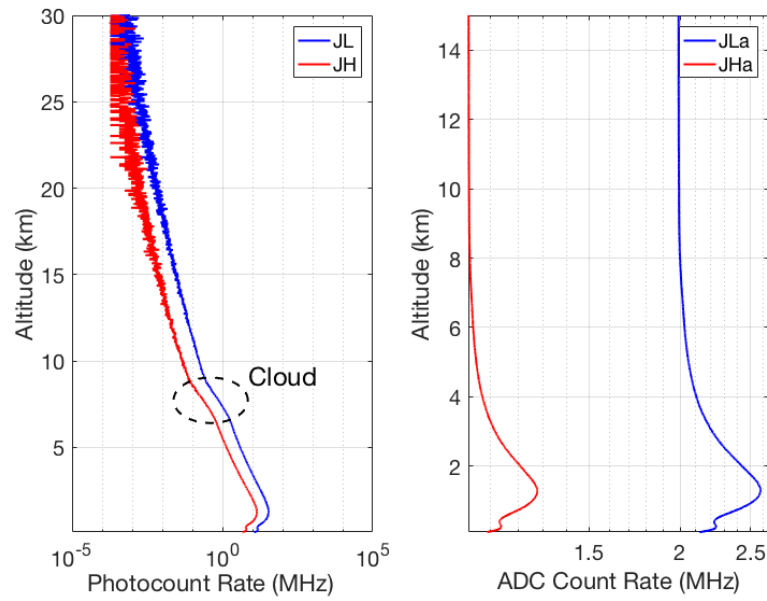


Figure 2.12: Count rate for 30 min of coadded RALMO measurements from 2300 UT on 05 July 2011. Left panel: digital channels (blue curve, JL; red curve, JH). Right panel: analog channels.

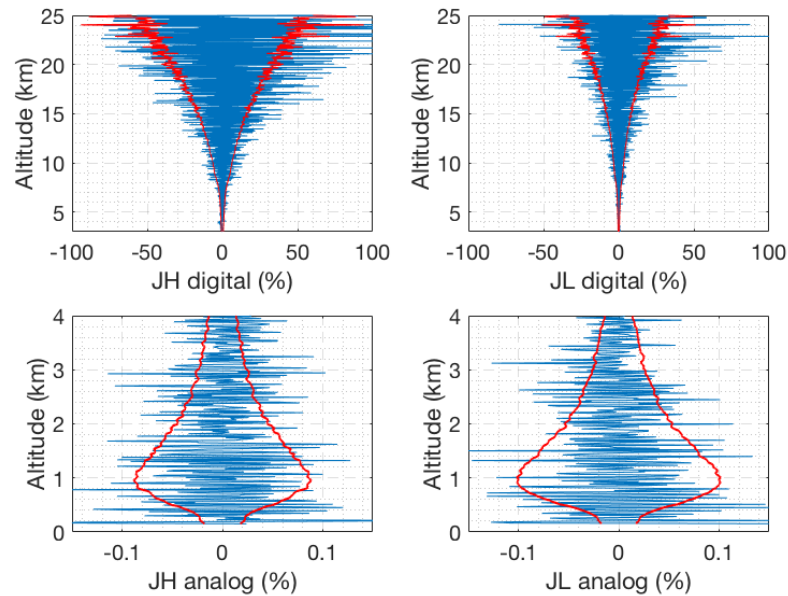


Figure 2.13: Difference between the forward model and the nighttime RALMO measurements on 05 July 2011 with the presence of a cirrus cloud for the four channels (blue). The red curves show the standard deviation of the measurements.

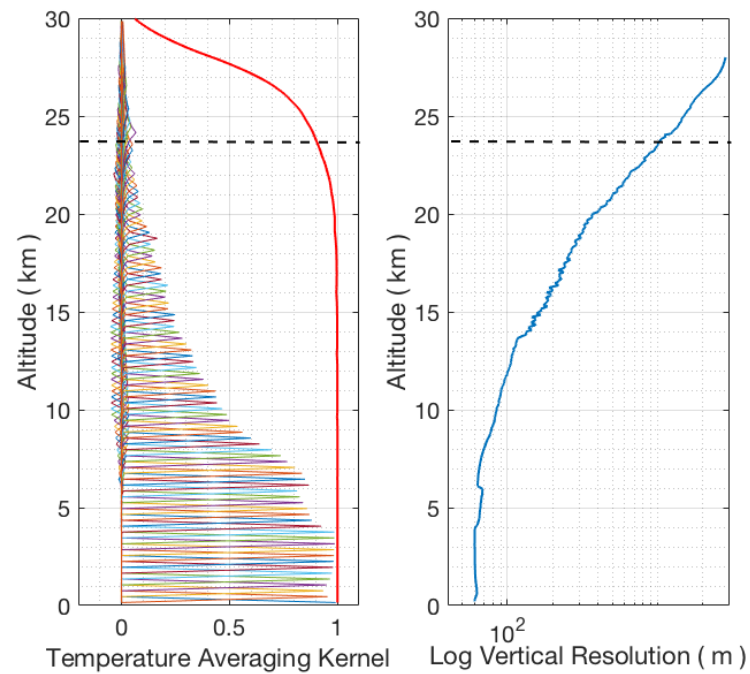


Figure 2.14: Same as Fig. 2.3 but for 05 July 2011 with a cirrus cloud at 6 km height.

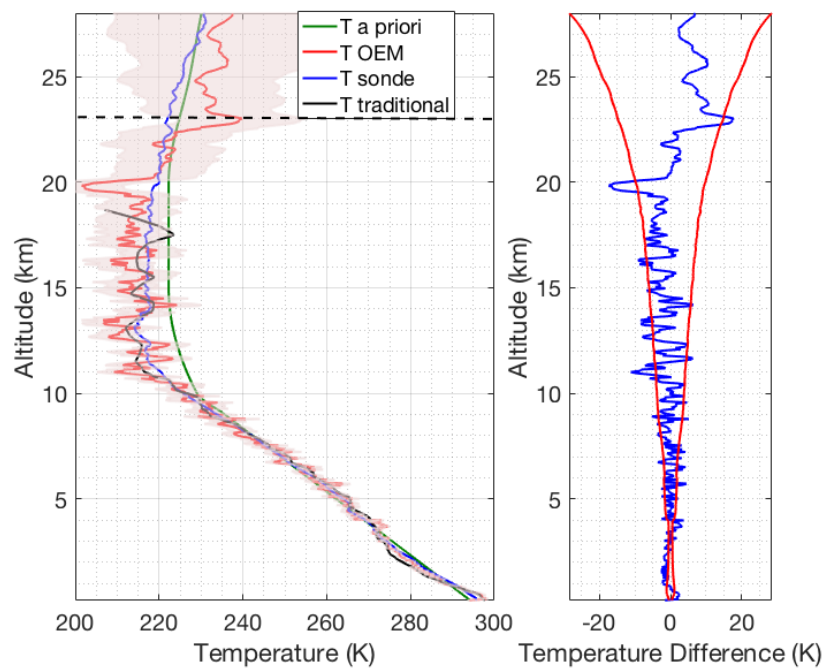


Figure 2.15: Same as Fig. 2.4 but for 05 July 2011 with a cirrus cloud at 6 km height, using the OEM.



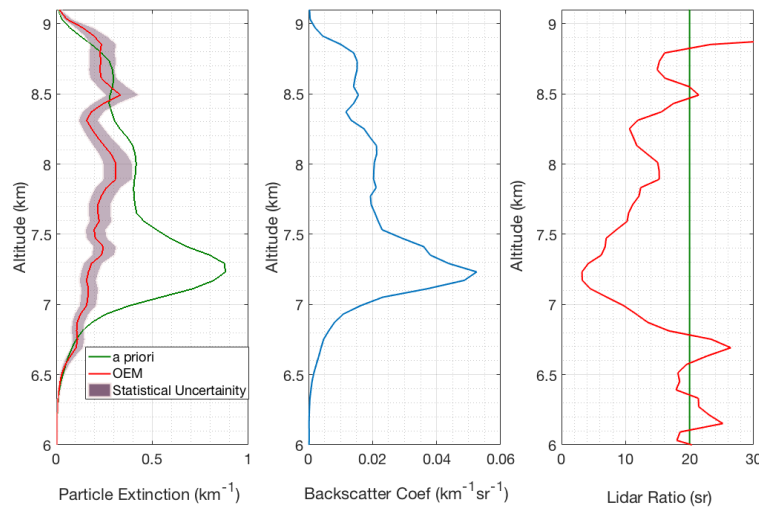


Figure 2.16: Left: Retrieved particle extinction (red) and the *a priori* particle extinction used in the OEM (green). Center: Backscatter coefficient calculated from the nighttime RALMO measurements on 05 July 2011 with of a cirrus cloud present at 6 km height. Right: Lidar ratio used to determine *a priori* particle extinction (green) and the estimated lidar ratio using the OEM-retrieved particle extinction (red).

estimate an *a priori* extinction. Above 6.75 km, the OEM-retrieved particle extinction is around  $0.25\text{km}^{-1}$  and approximately two times smaller than the *a priori* yielding a lidar ratio of 5-15 sr while the initial guess was 20 sr (Fig. 2.16. Ansmann et al. (1992b), using independent measurements of particle extinction and backscatter profiles in cirrus clouds, show similar extinction values ( $0 - 0.5\text{km}^{-1}$ ) and also similar values for the lidar ratio inside the cloud 0-10 sr. Thus, the OEM-retrieved extinctions for this cirrus cloud appears to be reasonable. Below 6.75 km, the lidar ratio is around 20 sr which could be an indication that this part of the cloud is super-cooled liquid.

The uncertainty budget for this case (not shown) is similar to the previous 2 cases shown; the cloud has little effect on the uncertainty values. As before, the statistical uncertainty makes the largest contribution to the full uncertainty.

#### 2.5.4 Case 4: Nighttime with lower level cloud

A cloud at about 4 km is present in measurements used for the last case study (Table 2.3). In this situation we use our OEM-retrieved geometrical overlap during clear conditions as our *a*

*priori* geometrical overlap profile. We then retrieve geometrical overlap to the cloud base (4 km height) and particle extinction above 4 km. In this case the retrieved geometrical overlap up to 4 km agrees within 10% uncertainty with the OEM-retrieved geometrical overlap for clear conditions.

Figure 2.17 shows the measurements in the four PRR channels and the elastic channel measurement (left panel, green curve). It can be seen in the elastic signal that a cloud base is at around 4 km height. The Raman measurements drop above 4 km and are fully attenuated at 7 km.

We use these measurements obtained at a cloudy condition as input to our OEM, and obtain unbiased residuals which fall within the standard deviation of the measurements, meaning the forward model accurately retrieve temperature in the presence of the cloud.

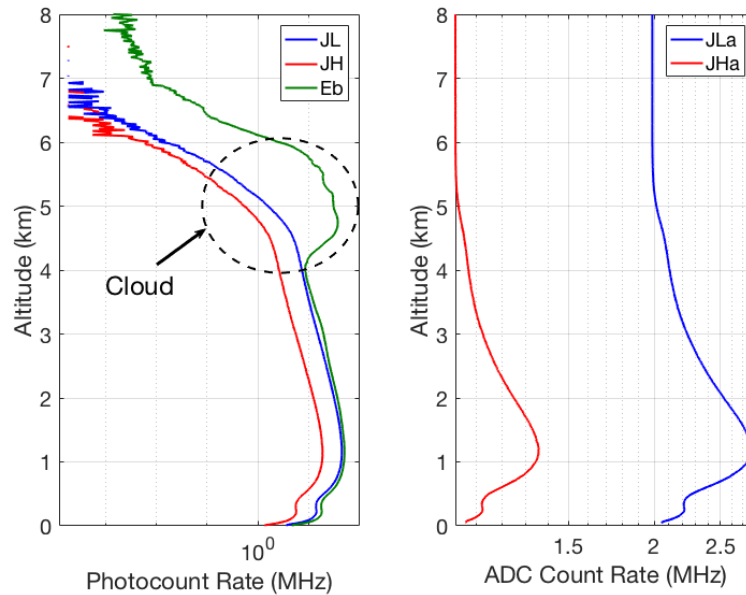


Figure 2.17: Count rate for 30 min of coadded RALMO measurements from 2300 UT on 21 June 2011, which has a cloud base at an height about 4 km. Left panel: digital channels (blue curve, JL; red curve, JH; green, Elastic). Right panel: analog channels.

The response function (left panel, Fig. 2.19) is 0.9 at 6 km, which is considered the maximum height where the OEM-retrieved temperature are valid. At this height the vertical resolution (right panel, Fig. 2.19) rapidly increases as the cloud thickens. As shown in Fig. 2.20 (right panel), up to 6 km, the temperature from the sonde launched at 2300 UT from Payerne

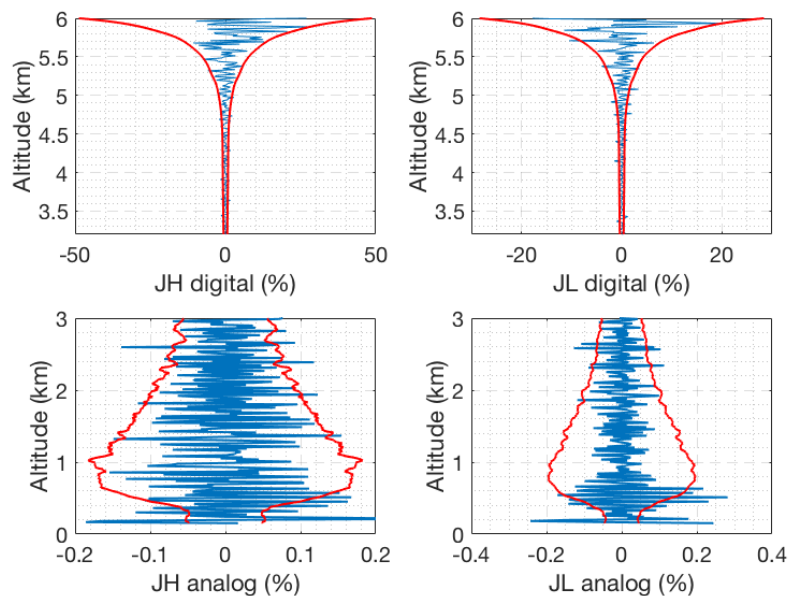


Figure 2.18: Difference between the forward model and the nighttime RALMO measurements on 21 June 2011 with the presence of a lower level cloud for the four channels (blue). The red curves show the standard deviation of the measurements.

and OEM temperature agree with each other within the statistical uncertainty of the OEM temperature. Temperature retrieved using the traditional method are similar to the OEM and sonde measurements up to 3.5 km, while inside the cloud the traditional temperature starts to deviate.

Below 5.25 km the retrieved particle extinction coefficient agrees well with the a priori values and the corresponding lidar ratio is between 15 and 20 sr indicating a liquid cloud. Above 5.25 km the retrieved particle extinction is smaller than the first guess yielding again a lidar ratio around 5 sr (left panel in Fig. 2.21). This could be an indication that the cloud became an ice cloud above 5.25 km.

## 2.6 Discussion

The four retrievals discussed in the previous section demonstrate that the OEM provides robust and accurate retrievals of temperature, geometrical overlap, and particle extinction coefficients, during both clear and cloudy day and night conditions. Unlike the traditional Raman lidar temperature analysis method (Cooney, 1972; Arshinov et al., 1983; Di Girolamo et al.,

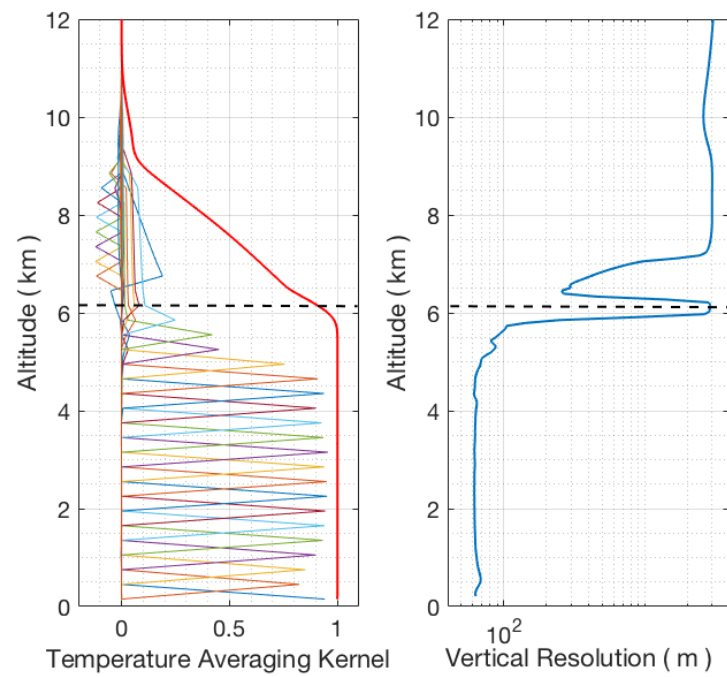


Figure 2.19: Same as Fig. 2.3 but for 21 June 2011 with the presence of lower level cloud.

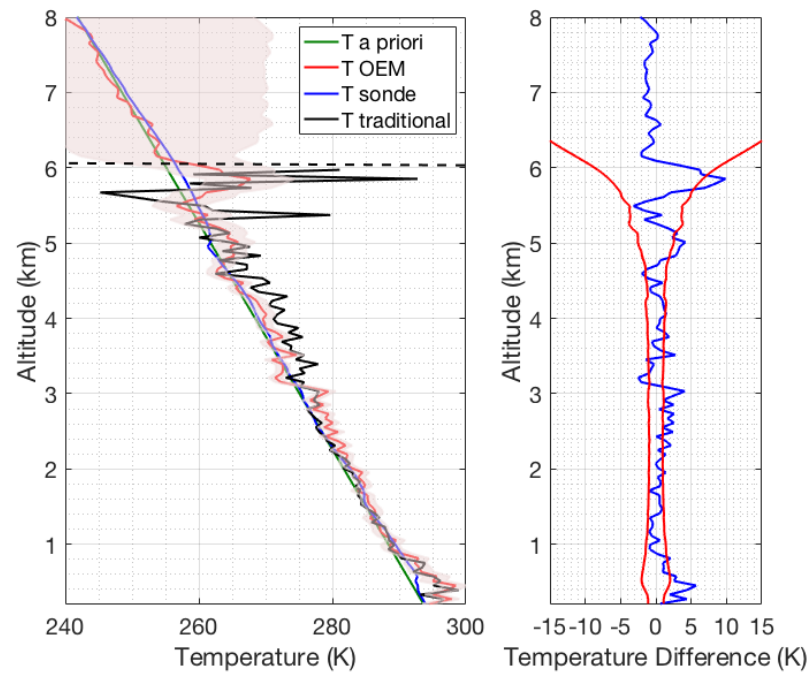


Figure 2.20: Same as Fig. 2.3 but for 21 June 2011 with the presence of a lower level cloud, using the OEM.

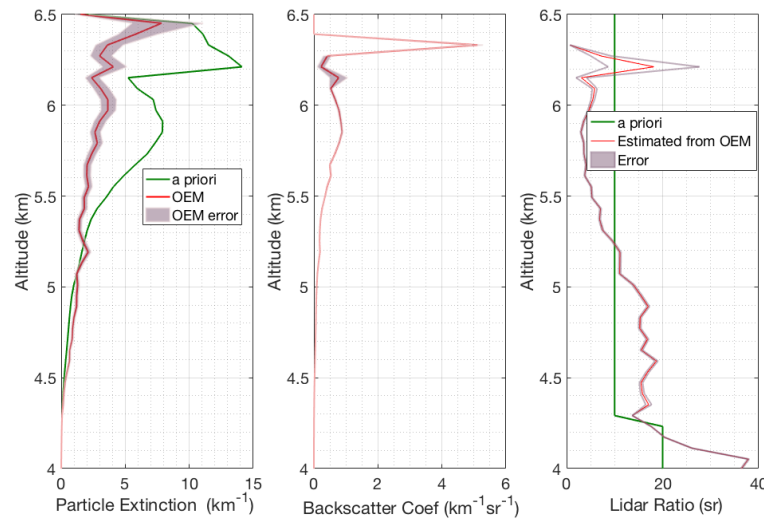


Figure 2.21: Same as Fig. 2.16 but for 21 June 2011 with the presence of a lower level cloud.

2004; Behrendt, 2005; Zuev et al., 2017), the OEM does not require an analytic form of a calibration function; rather a single calibration coefficient has to be estimated using a reference temperature profile and were shown to have a small effect on the retrieved temperature.

The calibration function plays a key role in the traditional temperature retrieval algorithm from the PRR backscattered signals, in particular if calibration is not done over the entire observed temperature range. Typically, a calibration function linear in  $1/T$  is used for systems that detect only one or multiple RR lines (Behrendt, 2005), although other forms of calibration functions have been employed. Recently Zuev et al. (2017) showed closer agreement at times with temperature calculation used a higher order polynomial for the calibration function. All calibration coefficient estimation methods require multiple reference data points which span ideally the entire range of temperatures to avoid extrapolation errors.

The only calibration required in our OEM scheme is the determination of the two coupling constants,  $R$  and  $R_a$ . The coupling constants can be estimated at a specific height (that is over a narrow range of temperature) without introducing extrapolation errors. Using the OEM we can show that the contribution of the coupling constant to the temperature uncertainty is in the order of 0.07 K in the height below 4 km and about 0.2 K or less above 4 km for a wide variety of sky conditions.

The OEM temperature retrievals of 4 very different sky conditions have been compared against coincident radiosonde temperature measurements. cases presented is the US Standard model normalized to the surface temperature from the coincident sonde temperature. We successfully used other *a priori* temperature profiles, such as the smoothed sonde temperature measurements and temperature from the Mass Spectrometer and Incoherent Scatter radar (MSIS) model to retrieve temperature using our OEM algorithm. All the retrieved temperature profiles using each *a priori* profile for heights where the response function is 0.9 or greater are identical within the statistical uncertainty.

In our study, we have successfully retrieved a geometrical overlap function for the RALMO system using the PRR measurements simultaneously with the temperature retrieval. Ray-tracking studies have concluded that the RALMO system reaches its full geometrical overlap by 5.0-5.5 km in height. These calculations are consistent with our geometrical overlap retrievals in both clear daytime and nighttime conditions. Measuring the geometrical overlap function and its uncertainty allows a more accurate estimation of the particle extinction coefficient when clouds or aerosols are present. The particle extinction profiles we retrieved in the two cloudy condition cases are consistent with measurements collected by Ansmann and Müller (2005) for cirrus clouds and with O'Connor et al. (2004) for liquid clouds.

The particle extinction coefficient is retrieved in the full geometrical overlap region, i.e. above 6 km or above the cloud base. The extinction values and lidar ratios we obtained for high and mid-level clouds are in agreement with other publications. The two case studies featuring clouds suggest, that both clouds consisted of a liquid and a ice part with lidar ratios at 18 and 5 sr.

For all the case studies we presented, the lidar constants for the lower quantum channels (analog and digital), the dead times for each digital channel (JL and JH), and background for all four channels are also retrieved. The retrieved lidar constants for each channel agreed within 20% uncertainty for all four cases. The retrieved dead times are about 3.8 ns and consistent with the dead times specified by the manufacturer and other independent estimates.

The uncertainty budget provided by the OEM contains both random and systematic uncertainties. Estimation of the uncertainty budget requires assignment of appropriate covariances to the model parameters. Using the standard deviations given in Table 2.2, the uncertainty

budgets for all the case studies are estimated. The largest contribution towards the temperature uncertainty originates from the statistical uncertainty due to the measurement noise. Overall contribution from the coupling constants to the temperature uncertainty are less than 0.2 K for all heights. Given the fact that the measurement noise can be reduced with longer integration times this result suggests that the OEM method very precise temperature measurements are possible even if calibration is only possible over a small temperature range. The systematic uncertainties of pressure and air density are on the order of 0.1 K to 0.1 mK respectively. Understanding the full uncertainty budget of temperature is of particular importance for trend analysis and process studies. The observational basis for super-saturation studies in the upper troposphere is still unsatisfactory and the OEM framework allows to combine different data sources to provide a high quality data set including profile-by-profile uncertainty budgets. .

### 2.6.1 Conclusion

We have demonstrated the ability of the OEM to retrieve multiple geophysical and instrumental parameters from PRR lidar measurements. The first-principle forward model adequately represents the raw PRR measurement and allows us to retrieve temperature, geometrical overlap, particle extinction, lidar constants, background counts, and dead time using multiple analog and digital channels. The retrievals discussed for four different cases that represent different (and typical) sky conditions. We found the following results from our OEM temperature retrievals from PRR measurements:

- The forward model presented, based on the lidar equation, contains the essential physics to reproduce the analog and digital measurements, leading to unbiased residuals and robust estimates of temperature.
- Our OEM retrieval does not require a calibration function as used in the traditional temperature retrieval method. It only requires determination of the two coupling constants,  $R$  and  $R_a$ , using a reference temperature profile that can be estimated at a specific height bin (or over a range). Retrieved temperature profiles from both day and night uncorrected PRR measurements in clear and cloudy conditions agree well with coincident radiosonde measurements.

- The OEM provides a cutoff height for the temperature retrievals that specify up to which height the retrieved profile is primarily due to the measurements and not the *a priori* temperature profile.
- Vertical resolution is determined at each height, and is automatically adapted in the retrieval in response to increasing measurement noise with height.
- The OEM provides a complete uncertainty budget, including random and systematic uncertainties due to model parameters, including the assumed pressure, air density and the coupling constants.
- Simultaneous analog, which are linear, and digital measurements allow the dead time to be retrieved.
- The OEM-retrieved geometrical overlap function for the RALMO using the measurements in clear conditions is determined and shown to be consistent with, but not the same, as that calculated by Dinoev et al. (2013). Hence, retrievals of the particle extinction coefficient are possible using the OEM from the measurements in cloudy conditions or when aerosol layers are present.
- The OEM is a computationally fast and practical for routine temperature retrievals from lidar measurements as required for operational lidar systems.

We have demonstrated that the OEM allows retrieval of temperature from Pure Rotational Raman lidar measurements that are consistent with the coincident sonde temperature. We discussed the advantages of the OEM over the traditional temperature retrieval algorithm. We can use the OEM-retrieved temperature to study temperature trends with the benefit of a full uncertainty budget provided by our OEM. Our OEM temperature retrieval can also be used for routine measurements in a wide variety of observing conditions, and is applicable to any similar PRR lidar system.

We are in the process of implementing the OEM for routine temperature measurements from the RALMO system. Currently we are also combining the OEM PRR temperature retrieval with the OEM water vapor mixing ratio retrieval of Sica and Haeferle (2016) to directly



retrieve relative humidity from the RALMO measurements, both for its importance to operational forecasting and to allow the study of ice super-saturation events.

## Bibliography

- Adam, S., Behrendt, A., Schwitalla, T., Hammann, E., and Wulfmeyer, V.: First assimilation of temperature lidar data into an NWP model: impact on the simulation of the temperature field, inversion strength and PBL depth, *Quarterly Journal of the Royal Meteorological Society*, 142, 2882–2896, 2016.
- Ansmann, A. and Müller, D.: Lidar and atmospheric aerosol particles, in: *Lidar*, pp. 105–141, Springer, 2005.
- Ansmann, A., Riebesell, M., Wandinger, U., Weitkamp, C., Voss, E., Lahmann, W., and Michaelis, W.: Combined Raman elastic-backscatter lidar for vertical profiling of moisture, aerosol extinction, backscatter, and lidar ratio, *Applied Physics B*, 55, 18–28, 1992a.
- Ansmann, A., Wandinger, U., Riebesell, M., Weitkamp, C., and Michaelis, W.: Independent measurement of extinction and backscatter profiles in cirrus clouds by using a combined Raman elastic-backscatter lidar, *Applied optics*, 31, 7113–7131, 1992b.
- Arshinov, Y. F., Bobrovnikov, S., Zuev, V. E., and Mitev, V.: Atmospheric temperature measurements using a pure rotational Raman lidar, *Applied Optics*, 22, 2984–2990, 1983.
- Atmosphere, U. S.: National oceanic and atmospheric administration, National Aeronautics and Space Administration, United States Air Force, Washington, DC, 1976.
- Behrendt, A.: Temperature measurements with lidar, in: *Lidar*, pp. 273–305, Springer, 2005.
- Bucholtz, A.: Rayleigh-scattering calculations for the terrestrial atmosphere, *Applied Optics*, 34, 2765–2773, 1995.
- Cooney, J.: Measurement of atmospheric temperature profiles by Raman backscatter, *Journal of applied meteorology*, 11, 108–112, 1972.

- Di Girolamo, P., Marchese, R., Whiteman, D., and Demoz, B.: Rotational Raman Lidar measurements of atmospheric temperature in the UV, *Geophysical Research Letters*, 31, 2004.
- Dinoev, T., Simeonov, V., Calpini, B., and Parlange, M.: Monitoring of Eyjafjallajökull ash layer evolution over payerne Switzerland with a Raman lidar, *Proceedings of the TECO*, 2010.
- Dinoev, T., Simeonov, V., Arshinov, Y., Bobrovnikov, S., Ristori, P., Calpini, B., Parlange, M., and Van den Bergh, H.: Raman Lidar for Meteorological Observations, RALMO-Part 1: Instrument description, *Atmospheric Measurement Techniques*, 6, 1329–1346, 2013.
- Eriksson, P., Jiménez, C., and Buehler, S. A.: Qpack, a general tool for instrument simulation and retrieval work, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 91, 47–64, 2005.
- Farhani, G., Sica, R. J., Godin-Beekmann, S., Ancellet, G., and Haefele, A.: Improved ozone DIAL retrievals in the upper troposphere and lower stratosphere using an optimal estimation method, *Appl. Opt.*, 58, 1374–1385, <https://doi.org/10.1364/AO.58.001374>, 2019.
- Herzberg, G.: *Molecular spectra and molecular structure*, vol. 1, Read Books Ltd, 2013.
- Lenschow, D. H., Wulfmeyer, V., and Senff, C.: Measuring second-through fourth-order moments in noisy data, *Journal of Atmospheric and Oceanic Technology*, 17, 1330–1347, 2000.
- Nicolet, M.: On the molecular scattering in the terrestrial atmosphere: An empirical formula for its calculation in the homosphere, *Planetary and Space Science*, 32, 1467–1468, 1984.
- O'Connor, E. J., Illingworth, A. J., and Hogan, R. J.: A technique for autocalibration of cloud lidar, *Journal of Atmospheric and Oceanic Technology*, 21, 777–786, 2004.
- Pappalardo, G., Amodeo, A., Pandolfi, M., Wandinger, U., Ansmann, A., Bösenberg, J., Matthias, V., Amiridis, V., De Tomasi, F., Frioud, M., et al.: Aerosol lidar intercomparison in the framework of the EARLINET project. 3. Raman lidar algorithm for aerosol extinction, backscatter, and lidar ratio, *Applied optics*, 43, 5370–5385, 2004.

- Penney, C., Peters, R. S., and Lapp, M.: Absolute rotational Raman cross sections for  $N_2$ ,  $O_2$ , and  $CO_2$ , *JOSA*, 64, 712–716, 1974.
- Povey, A., Grainger, R., Peters, D., and Agnew, J.: Retrieval of aerosol backscatter, extinction, and lidar ratio from Raman lidar with optimal estimation, *Atmospheric Measurement Techniques*, 7, 757–776, 2014.
- Povey, A. C., Grainger, R. G., Peters, D. M., Agnew, J. L., and Rees, D.: Estimation of a lidars overlap function and its calibration by nonlinear regression, *Applied optics*, 51, 5130–5143, 2012.
- Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, vol. 2, World scientific, 2000.
- Sica, R. and Haeferle, A.: Retrieval of temperature from a multiple-channel Rayleigh-scatter lidar using an optimal estimation method, *Applied optics*, 54, 1872–1889, 2015.
- Sica, R. and Haeferle, A.: Retrieval of water vapor mixing ratio from a multiple channel Raman-scatter lidar using an optimal estimation method, *Applied optics*, 55, 763–777, 2016.
- Whiteman, D. N.: Examination of the traditional Raman lidar technique. I. Evaluating the temperature-dependent lidar equations, *Applied Optics*, 42, 2571–2592, 2003.
- Zuev, V. V., Gerasimov, V. V., Pravdin, V. L., Pavlinskiy, A. V., and Nakhtigalova, D. P.: Tropospheric temperature measurements with the pure rotational Raman lidar technique using nonlinear calibration functions, *Atmospheric Measurement Techniques*, 10, 315–332, 2017.

# Chapter 3

## 1DVar reanalysis of 10 years of ERA5 assimilating a Raman lidar for temperature and humidity<sup>1</sup>

### 3.1 Overview

Water vapor is a key factor for the atmosphere's radiative budget, chemistry and dynamics. Upper tropospheric humidity exhibits a large radiative forcing (Comstock et al., 2004; Sinha and Harries, 1995) and is a driving factor of the formation of cirrus clouds, which in turn modify the radiative budget. However, there are still large uncertainties associated with cirrus cloud formation since observations of high levels of ice super saturation up to 200%RH (we refer the unit of relative humidity as %RH and standard percentages as %) have been reported by Gensch et al. (2008) and references therein. These extreme observations may reveal a gap in our knowledge of the physics and chemistry of cirrus cloud formation (Peter et al., 2006). These observations, though, still show large uncertainties and do not bring the necessary observational evidence to confirm or prove incomplete our understanding of cirrus cloud formation (Gierens et al., 2012; Comstock et al., 2004). To improve the situation good quality profiles with uncertainty budgets are needed.

---

<sup>1</sup>A version of this chapter is in its final preparation stage for submission in the near future.

The global radiosonde network provides most of the relative humidity information for forecast models. Even though the radiosonde network is widely spread over the world, the temporal resolution of the routine sonde measurements is typically two radiosondes per day. Also, it is well known that the radiosonde relative humidity measurements are often not reliable in the upper troposphere (Miloshevich et al., 2001; Noh et al., 2016; Ferreira et al., 2019).

Among several other techniques available for improved water vapor measurements, such as satellites and microwave radiometers, Raman lidar has become one of the best tools that provides water vapor measurements throughout the troposphere with high spatial and temporal resolutions (Whiteman et al., 1992; V  r  mes et al., 2016). For Raman lidars which possess temperature profiling capability exploiting pure rotational Raman scattering from nitrogen and oxygen, the water vapor information can be combined with temperature to yield relative humidity (Mattis et al., 2002). These authors report a relative uncertainty of the calculated relative humidity between 5-25%RH, the uncertainty of the temperatures being the dominating source of uncertainty. A 1-2 K temperature accuracy is required for reducing the relative uncertainty in the relative humidity to, on average, less than 10%RH. Brief description of the conventional Raman lidar relative humidity estimation method is given in Appendix B.

Here we apply a 1D Var data assimilation scheme to reanalyze the ERA5 relative humidity profiles above Payerne assimilating data from a Raman water vapor and temperature lidar. The 1D Var scheme is based on the work in (Mahagammulla Gamage et al., 2019) and (Sica and Haefele, 2016) and uses raw data from the lidar, i.e. Raman backscatter profiles, as opposed to retrievals, i.e. water vapor or temperature profiles. The reanalyzed ERA5 profiles, hereafter referred to as ERA5-reRH, come along with a complete profile-by-profile characterization in terms of uncertainty, sensitivity and vertical resolutions. We have chosen to reanalyze relative humidity profiles in units of relative humidity with respect to water,  $RH_w$ , since this allows an easy validation with radiosondes which generally report  $RH_w$  (Dirksen et al., 2014; Miloshevich et al., 2009).  $RH_w$  is derived from temperature and water vapor mixing ratio using the Hyland and Wexler formulation (see Section 3.3.1). Note that for this study this is simply a choice of the unit and the data can be converted to relative humidity with respect to ice if needed.

The main purpose of the ERA5-reRH data set is to study ice super saturation events. The

paper is organized as follows: In Section 3.2 we provide a description of the Raman lidar and ERA5 data we used in the study. The forward model and the implementation of the 1D Var algorithm together with a characterization of the ERA5-reRH profiles are given in Section 3.3. Section 3.5 contains a validation of the ERA5-reRH data set versus radiosondes. A detailed discussion of the results, conclusions and an outlook to future work are given in Sections 3.6, 3.7.

## **3.2 Data used in 1DVar reanalysis**

### **3.2.1 Raman lidar for Meteorological Observations (RALMO)**

For this study we use Raman lidar measurements from the Raman lidar for Meteorological Observations (RALMO), operated by MeteoSwiss and located in Payerne, Switzerland. (46°48' N, 6°56' E). RALMO is a fully automated lidar, operating since 2008, with an average up-time of 50%. The transmitting system of RALMO consists of a frequency tripled, Q-switched Nd:YAG laser at 354.7 nm generating up to 400 mJ per shot at a 30 Hz repetition rate. The laser pulses are 8 ns pulse in duration. The lidar telescope receiver consists of four 30 cm diameter telescopes that are tightly arranged around a 15x beam expander. The telescope is fiber coupled to the spectral unit. A near range optical fiber, located off of the optical axis of one of the four telescopes improves the signal-to-noise ratio in the partial overlap region and allows water vapor and temperature measurements down to roughly 100 m above ground. The RALMO detection system consists of two polychromators isolating the water vapor and nitrogen Raman return at 407 and 387 nm, respectively, as well as four portions of the pure rotational Raman spectrum including high and low quantum number lines in the Stokes and anti-Stokes branches. All signals are acquired in analog and digital mode which yields a total of 8 channels. A detailed description of RALMO is given by Dinoev et al. (2013), and the instrument's validation is given in Brocard et al. (2013).

### **3.2.2 ERA5 reanalysis data**

ERA5 is the fifth generation of the European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric reanalyses of the global climate. ERA5 was produced using 4D-Var

data assimilation. It provides hourly temperature, relative humidity (over water above freezing and over ice below freezing temperature), specific humidity, geopotential and many other atmospheric parameters with an uncertainty estimate at different pressure levels between the surface and the mesopause from the year 1979 on wards. ERA5 is combined with observations from several satellites and in-situ measurements worldwide to provide a complete and consistent data set. The data assimilation is done twice per day using 12 hour windows from 09 UTC to 21 UTC and 21 UTC to 09 UTC (the following day). The two types of ERA5 data sets contain one (31 km) high resolution realisation (HRES) and a reduced resolution ten member ensemble (EDA) (Hennermann and Berrisford, 2017).

Humidity is reported both as specific and as relative humidity. For temperatures above  $0^{\circ}\text{C}$  relative humidity is given with respect to water, for temperatures below  $-23^{\circ}\text{C}$  it is given with respect to ice, and a mix of the two for temperatures between  $0^{\circ}\text{C}$  to  $-23^{\circ}\text{C}$  (ECMWF, 2018). Since here we systematically use relative humidity with respect to water,  $RH_w$ , as unit, we convert ERA5 specific humidity, temperature and pressure data to relative humidity with respect to water for all temperatures, hereafter referred to as  $RH_{w,ERA5}$ . For this calculation we use the Hyland and Wexler formulation as given in Eq. 3.7.

The ERA5 reanalysis assimilate both satellites and in-situ measurements. A list of all the observations is presented in Hennermann and Berrisford (2017) and the data usage in ERA5 for the segment from 1979 is presented in Hersbach et al. (2019). The number of temperature observation sources in the model is comparatively larger than the number of humidity observation sources. ERA5 also provides uncertainty (ensemble spread) estimates from a 10-member ensemble data assimilation. In general lower ensemble spreads in ERA5 indicate the high confidence in the data due to the bulk data coverage. As shown in (Hersbach et al., 2019) the accuracy of the ERA5 temperature data is improved over the years due to the increase of the number of temperature observations that are assimilated into the model. The temperature ensemble spreads for ERA5 reanalysis in year 2000's are about 0.5 K and hence close to the uncertainty of the Raman lidar temperature. Thus, Raman lidar temperatures will have only a limited impact on the ERA5-reRH data quality. As there is no available comparison of the accuracy of the ERA5 specific humidity data over the years, we studied ERA5 specific humidity ensemble spreads for the same date in every decade starting from 1980-2010. Our results

(not shown in the paper) showed no significant improvements in the ERA5 specific humidity data over the time even with the higher observational coverage. Thus, the accuracy of the relative humidity data derived from the ERA5 data is not validated. However, the ERA5 relative humidity ensemble spreads for the same date in every decade starting from 1980-2010 from surface to about 15 km are 10%RH . Where the relative humidity standard uncertainty of the lidar is better than 5%RH. Thus, the we expect significant impact on the ERA5-reRH relative humidity data from the lidar measurements.

Further details of the ERA5 reanalysis can be obtained from Hennermann and Berrisford (2017) and the ERA5 data can be accessed either from the Meteorological Archival and Retrieval System (MARS) archive or from Climate Data Store (CDS) cloud server that has comparatively fast access (Hersbach et al., 2019).

### 3.3 1D Var retrieval of relative humidity from Raman lidar measurements and ERA5 (ERA5-reRH)

For the 1D Var data assimilation we use the optimal estimation method (OEM) solver in the Qpack software package, developed by Eriksson et al. (2005) to retrieve relative humidity, temperature and other parameters. The OEM solver requires the following inputs: the measurements and their error covariances, *a priori* values and their error covariances (more commonly referred to as background and background error covariance in data assimilation), forward model (**b**) parameters and their error covariances, and the forward model and its Jacobians.

#### 3.3.1 Forward model

The forward model for the lidar measurements is based on the Raman lidar equation and relates the observed backscattered photon-counts of the  $i$ th channel,  $N_{obs,i}$ , to the atmospheric parameters of ERA5:

$$N_{obs,i}(z) = \frac{C_i}{z^2} O_i(z) \left( n_i(z) \frac{d\sigma_i(\pi)}{d\Omega} + \beta_{i,aer} \right) \exp \left( - \int_0^z [\alpha_{\lambda_0}(z) + \alpha_{\lambda_i}(z)] dz \right) + B_i(z), \quad (3.1)$$



where  $C_i$  is the lidar constant for channel  $i$  that depends on the number of transmitted photons, detector efficiency and area of the telescope.  $O_i(z)$  is the geometrical overlap that describes the overlap between the transmitted laser beam and the field of view of the telescopes.  $n_i(z)$  is the number density of the scattering molecule,  $\beta_{i,aer}$  is the aerosol backscatter coefficient and  $B_i(z)$  is the background of the observed signal.  $\frac{d\sigma_i(\pi)}{d\Omega}$  is the differential Raman backscatter cross-section.  $\alpha_{\lambda_0}(z)$  and  $\alpha_{\lambda_i}(z)$  are the extinction coefficients at the transmitted wavelength and at the wavelength of receiver channel  $i$ .

The lidar equation given in Eq. 3.1 presumes that the count rates linearly depend on the number of received photons that are valid for analog channels and for digital channels when the count rates are below 1 MHz. For the digital channels at higher count rates, the true and observed counts are related by the saturation equation. For non-paralyzable counting system observed counts  $N_{obs}$  are given by Kovalev and Eichinger (2004):

$$N_{obs} = \frac{N_{tr}}{1 + N_{tr}\gamma}. \quad (3.2)$$

The dead time,  $\gamma$ , characterizes the response speed of the digital acquisition system. The forward model (Eq. 3.1 and 3.2 for digital channels only) is evaluated 8 times to produce the 4 digital and 4 analog signals corresponding to rotational-vibrational Raman scattering of water vapor ( $i = Wd, Wa$ ) and nitrogen ( $i = Nd, Na$ ) and pure rotational Raman (PRR) scattering of high ( $i = JHd, JHa$ ) and low ( $i = JLD, JLa$ ) quantum numbers.

For the PRR channels the number densities are equal to the air number density,  $n_{air} = n_{JHd} = n_{JHa} = n_{JLD} = n_{JLa}$ , which is replaced by pressure and temperature assuming the hydro-static equilibrium and the ideal gas law (Behrendt, 2005).

$$n_{air}(z) = \frac{P_{HSEQ}(z)}{kT(z)}, \quad (3.3)$$

where  $P_{HSEQ}(z)$  is:

$$P_{HSEQ}(z) = P_0 \exp \left[ \int_{z_0}^{z_{top}} \frac{M(z)g(z)}{R^*T(z)} dz \right]. \quad (3.4)$$

$M(z)$  is the height-dependent mean molecular mass,  $g(z)$  is the acceleration due to gravity,  $P_0$  is the pressure at the maximum altitude, and  $R^*$  is the universal gas constant.

The relative humidity over water is defined as:

$$RH_w(z) = \frac{e(z)}{e_w(z)}, \quad (3.5)$$

where  $e(z)$  is the water vapor pressure and  $e_w(z)$  is the saturation vapor pressure. The water vapor pressure can be given as a function of the water vapor mixing ratio (Yau and Rogers, 1996),

$$e(z) = \frac{P(z)m_{wv}(z)}{0.622 + m_{wv}(z)}. \quad (3.6)$$

The saturation vapor pressure can be expressed as a function of temperature (Hyland, 1983):

$$\begin{aligned} \text{Log}e_w(z) = & \frac{-0.5800220610^4}{T} + 0.1391499310^1 - 0.4864023910^{-1}T \\ & + 0.4176476810^{-4}T^2 - 0.1445209310^{-7}T^3 + 0.6545967310^1 \text{Log}(T). \end{aligned} \quad (3.7)$$

The saturation vapor pressure is obtained by the log inverse of Eq. 3.7.

The number density of water vapor  $n_{wv} = n_{wd} = n_{wa}$  can be replaced using the definition of the relative humidity given in Eq. 3.5.

$$n_{wv}(z) = \frac{e(z)}{kT(z)} = \frac{RH_w(z)e_w(z)}{kT(z)}. \quad (3.8)$$

The nitrogen number density,  $n_{n2} = n_{Nd} = n_{Na}$ , can be related to relative humidity as:

$$n_{n2}(z) = \frac{P_{HSEQ}(z) - RH_w(z)e_w(z)}{0.643kT(z)}. \quad (3.9)$$

It can be seen from the lidar equations that temperature information comes predominantly from the PRR channels humidity information from the nitrogen and water vapor Raman channels. Substituting water vapor mixing ratio with relative humidity couples the two sets of channels and hence relative humidity information comes from all channels.

The extinction coefficients  $\alpha_{\lambda_0}(z)$  and  $\alpha_{\lambda_i}(z)$  in Eq. 3.1 contains the effects of both molecules and particles.

$$\alpha_{\lambda_0}(z) = \alpha_{\lambda_0,mol}(z) + \alpha_{\lambda_0,par}(z) \quad (3.10)$$

where  $\alpha_{\lambda_0,mol}(z)$  is the molecular extinction coefficient at the emitted wavelength and  $\alpha_{\lambda_0,par}(z)$  is the particle extinction coefficient at the emitted wavelength. Similarly,  $\alpha_{\lambda_i}(z)$  can be represented as a summation of molecular and particle extinction. For PRR channels ( $i = JHd, JHa, JLa$ ),  $\alpha_{\lambda_0}(z)$  is assumed to be equal to  $\alpha_{\lambda_i}(z)$ , as the emitted and the received wavelengths are very close to each other.  $\alpha_{\lambda_0,mol}(z)$  given in Eq. 3.10 can be expressed using the Rayleigh cross-section,  $\sigma_{\lambda_0,Ray}$  and air density,  $n_{air}(z)$ .

$$\alpha_{\lambda_0,mol}(z) = \sigma_{\lambda_0,Ray} n_{air}(z) \quad (3.11)$$

$\sigma_{\lambda_0,Ray}$  is calculated for each wavelength using the expressions given by Nicolet (1984). In our work the air density term in Eq. 1.39 is considered as a model parameter; thus estimated using the pressure and temperature either from sonde measurements or reanalysis data.

All the particle extinction coefficients are expressed at their respective return wavelength using the following expression:

$$\alpha_{\lambda_i,par}(z) = \alpha_{\lambda_0,par}(z) \left( \frac{\lambda_i}{\lambda_0} \right)^{-\ddot{a}(z)} \quad (3.12)$$

where  $\ddot{a}(z)$  is the Ångstrom exponent as a function of height.

For a stable calibration of the lidar we further introduce the calibration factors,  $R$ , to eliminate  $C_{Wd}$ ,  $C_{Wa}$ ,  $C_{JHd}$  and  $C_{JHa}$  from the forward model:

$$R_{WVd} = \frac{C_{Wd}}{C_{Nd}} \quad (3.13a)$$

$$R_{WVa} = \frac{C_{Wa}}{C_{Na}} \quad (3.13b)$$

$$R_{PRRd} = \frac{C_{JHd}}{C_{JLd}} \quad (3.13c)$$

$$R_{PRRa} = \frac{C_{JHa}}{C_{HLa}} \quad (3.13d)$$

$$(3.13e)$$

These calibration factors are determined outside the 1D Var process using collocated radiosonde profiles. For details concerning the calibration see Sica and Haefele (2016) and Mahagam-

mulla Gamage et al. (2019).

### 3.3.2 Error covariance matrices

In this section we explain how the various error covariance matrices are constructed prior to run the 1D Var retrieval. An overview is given in Table 3.1.

#### Measurement noise

The covariance matrices of the 2 sets of Raman lidar measurements, analog and digital, are diagonal assuming no correlation of the noise between channels. For the analog channels and the digital measurements that are not in the linear range (count rate is  $>10$  MHz) the variances are estimated using the auto-correlation function method introduced by Lenschow et al. (2000). The measurements from the digital channels that are in the linear range follow Poisson statistics and the variance is equal to the signal itself.

#### *A priori* (background) relative humidity and temperature

The *a priori* or background error covariance matrices of temperature and relative humidity is a key parameter in the 1D Var process since it controls directly the weight that is given to the background, i.e. ERA5. As it was mentioned earlier, reanalysis data is produced assimilating an extensive set of global observations. Thus, the accuracy the reanalysis data is higher than typical forecasts. To construct numerically suited *a priori* error covariance matrices for temperature and relative humidity we use a tent function to parametrize the off-diagonal elements which decay linearly from the variance on the diagonal  $1/e$  over a correlation length. Negative values are set to zero.

We first calculated the mean and the standard deviation of the temperature and relative humidity differences between sonde measurements (observations) and ERA5 data from 2004-2015. Since the routine sonde measurements made at 1100 and 2300 UT from Payerne are assimilated in the ERA5 data, we only consider a set of special sonde measurements that were made at times between 0600-0900 UT, 1300-1500 UT, and 1800-2100 UT. Fig. 3.1 shows dates and times of the total of 56 special soundings that were considered in the calculation. We

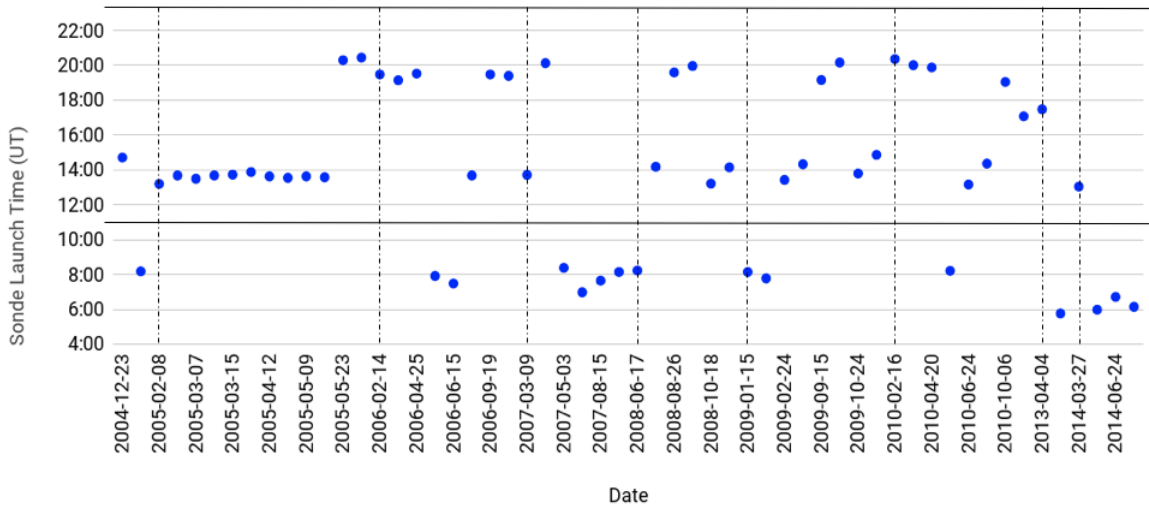


Figure 3.1: Dates and times of the sondes were launched from Payerne, Switzerland that coincide with the ERA5 reanalysis data used to estimate the correlations of the temperature and relative humidity. The two black lines indicate the usual sonde launching times 1100 and 2300 UT from Payerne.

observed a systematic bias in the ERA5 temperature and relative humidity with respect to the sonde measurements (shown in Fig. 3.2), as has been previously noted by Hersbach and Dee (2016); Yegui et al. (2018); Lompar et al. (2019) (Fig. 3.2). Up to about 12 km from the ground ERA5 shows a warm bias with maximum temperature biases of about 4.5 K and 3.6 K at the surface and at about 8 km. About 0.5 K cold bias is observed in the ERA5 temperatures above 12 km. From the ground to about 1.5 km ERA5 relative humidity is too low by up to 20%RH while above 1.5 km ERA5 shows a dry bias of 0 to 15%RH. To obtain bias corrected ERA5 data we subtract the temperature and relative humidity biases from the ERA5 reanalysis temperature and relative humidity data. For the rest of the paper ERA5 refers to the bias corrected profiles above Payerne and constitutes the a priori information used in the 1D Var process.

To determine the correlation lengths of the temperature and relative humidity errors we first computed the correlation matrices of the differences between ERA5 and the coincident sonde measurements (see Fig. 3.3). By eye the correlation lengths for temperature and relative humidity was then estimated to be 1 km and 750 m, respectively, throughout the entire troposphere.

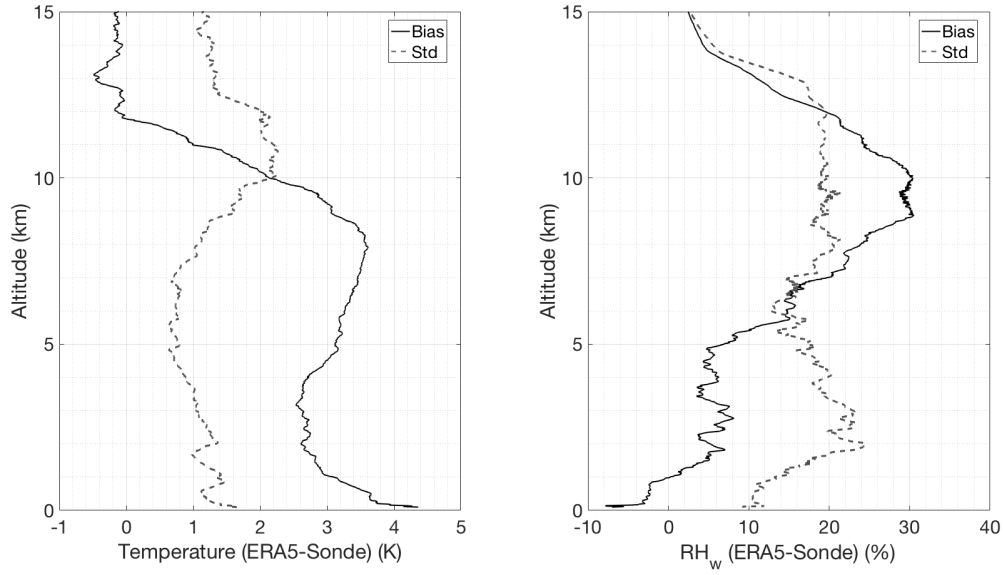


Figure 3.2: Left Panel: Profiles of ERA5-sonde mean temperature bias (solid black curve) and standard bias (dash ash curve). Right Panel: Profiles of ERA5-sonde mean relative humidity bias (solid black curve) and standard bias (dash ash curve). The data are obtained from 56 available sonde measurements launched between year 2004-2015 from Payerne, Switzerland. Sonde launch date and time information are given in Fig. 3.1.

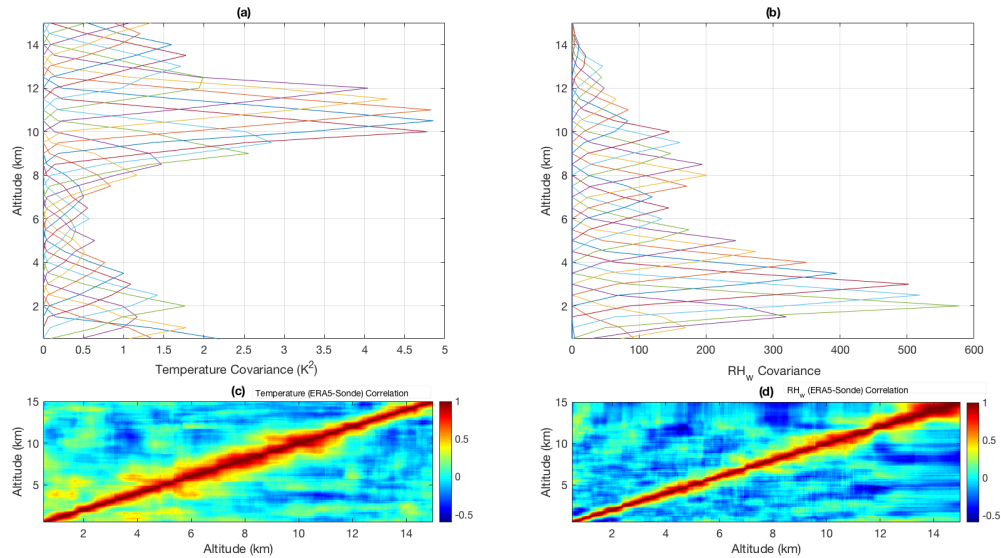


Figure 3.3: (a)Temperature and (b) relative humidity *a priori* covariance matrices. For clarity covariance for every tenth altitude bin the retrieval grid are shown. The correlation matrices of temperature and relative humidity differences used to estimate the covariances matrices that are calculated based on sonde and ERA5 data are shown in the subplots (c) and (d) respectively.

### Particle extinction and overlap

For particle extinction we derive an *a priori* or background profile from the backscatter ratio measured by the lidar rather than to use data from ERA5. To convert the backscatter ratio to particle extinction we assume a lidar ratio and use the same molecular extinction profile as in the OEM retrieval. As in previous studies we assume the lidar ratio for clear sky conditions (backscatter ratio smaller than 2) to be 80 sr inside the boundary layer and 50 sr elsewhere. Inside clouds we assume the lidar ratio to be 20 sr below 6 km (liquid cloud), and 15 sr above (cirrus cloud; Ansmann et al. (1992); Pappalardo et al. (2004)). We identify the presence of cloud based on the backscatter ratio using a threshold of 2. The cloud layer base height is used to define the altitude where the retrieval of overlap hands over to the retrieval of particle extinction since retrieving both simultaneously at the same altitude is mathematically not possible due to the high degree of linear dependence. In cloud free conditions, this handover takes place at 6 km where full overlap has been reached. For the particle extinction (overlap) error covariance matrix we assume a standard deviation of 50%RH above (below) the handover height and  $10^{-6} \text{ km}^{-1}$  ( $10^{-3}$ ) below (above). The off diagonal elements are parametrized using a tent function with a correlation length of 100 m. Same as for temperature and humidity, negative values are set to zero. For further details the reader is referred to Mahagammulla Gamage et al. (2019).

### Background, lidar constants, and dead times

The *a priori* backgrounds and their variances for both analog and digital channels are determined by the mean and the variance of the measurements above 50 km altitude. The four *a priori* lidar constants ( $C_{JLd}$ ,  $C_{JLa}$ ,  $C_{Nd}$ , and  $C_{Na}$ ) required for the 1D Var process are estimated by fitting the forward model to the respective Raman lidar measurements. We consider 3.8 ns as the *a priori* dead times for the digital photon counting systems, values which were found by previous studies of using RALMO, and also consistent with the values specified by the manufacturer (Sica and Haefele, 2016, 2015; Dinoev et al., 2010; Mahagammulla Gamage et al., 2019).

### Forward model parameters

The model (**b**) parameters used in the forward model are the calibration factors (for analog channels:  $R_{PRR,a}, R_{wv,a}$  and for digital channels:  $R_{PRR,d}, R_{wv,d}$ ), Ångstrom exponent, seed pressure, and density of air for the atmospheric transmission. The values and uncertainties of the other **b** parameters and a summary of all the parameter uncertainties are given in Table 3.1.

Table 3.1: Values and associated uncertainties for the retrieval and forward model parameters.

Parameter	Value	Standard Deviation
<b>Measurements</b>		
Digital (JL, JH, Water vapor, Nitrogen)	Measured	Poisson Statistics
Analog (JL, JH, Water vapor, Nitrogen)	Measured	Auto Covariance Method
<b>Retrieval Parameters (<i>a priori</i>)</b>		
Temperature	ERA5 reanalysis	Standard deviation of ERA5 and coincident sonde temperature measurements
Relative humidity	Recalculated using ERA5 reanalysis temperature, pressure, and specific humidity	Standard deviation of ERA5 and coincident sonde relative humidity measurements
Overlap Functions	Estimated using the forward model and measurements	50% below and at transition height $10^{-3}$ above transition height
Particle Extinction	Estimated using measurements	$10^{-6}\text{km}^{-1}$ below and at transition height 50% above transition height
Lidar Constants (analog/digital)	Estimated using the forward model	100%
Digital Background Noise	Mean above 50 km	Standard Deviation above 50 km
Analog Background Noise	Mean above 50 km	nighttime- Standard Deviation above 50 km daytime- normalized standard deviation above 50 km
Dead Time	Empirical fitting	10%
<b>Forward Model Parameters</b>		
Coupling constants (analog/digital)	Estimated measurements and sonde temperature	Standard deviation of the linear fitting
Air density	Radiosonde	1%
Seed Pressure	HSEQ: pressure at the heights point of the retrieval grid	10%
Ångstrom exponent	1 for all heights	10%



### 3.3.3 Other 1D Var retrieval specifications

Prior to the assimilation the lidar raw data are co-added to 30 m bins in height and 30 min in time ( $\pm 15$  min around ERA5 analysis time). The retrieval grid spans from 600 m asl to 20 km asl with a grid spacing of 90 m.

### 3.3.4 Characterization of ERA5-reRH

In this section we present two representative instances from the new ERA5-reRH data set corresponding to day and nighttime. For comparison, ERA5-reRH is compared to coincident sonde measurements. While the focus of this section is the characterization in terms of the uncertainty budget and lidar data impact, a statistical validation of ERA5-reRH is given in 3.5.1.

#### Case 1: Nighttime, cloudy sky, 28 August 2012 2241-2311 UT

We consider 30 min of measurements started from the launch time of the coincident sonde from Payerne. The calibration coupling constants required in the OEM, were estimated using the 30 min raw lidar measurements and coincident sonde measurements. During the time of the lidar measurement period a cirrus cloud was present around 8 to 10 km.

Figure. 3.4 shows the ERA5-reRH results in comparison with coincident sonde measurements. Figure 3.4(a), shows the ERA5-reRH temperature (red curve) with the statistical uncertainty (shaded area), coincident sonde temperature (blue curve) and ERA5 temperature (black curve). In the 6 to 10 km region the ERA5-reRH temperatures are about 1.5 K colder than the sonde and ERA5 temperatures. However, as shown in Fig. 3.4(d), the ERA5-reRH relative humidity (red curve) is in good agreement with the coincident sonde measurements (blue curve) in the same region.

The measurement response function (impact of lidar data) is the sum of the averaging kernels and indicates the contribution of the measurement to the retrieval. For a retrieval fully dependent on measurements the response function is equal to one; when a retrieval fully dependent on the *a priori* profile the response function is equal to zero. The measurement response function (red curve) for temperature (Fig. 3.4(b)), reveals between 50 and 60% lidar

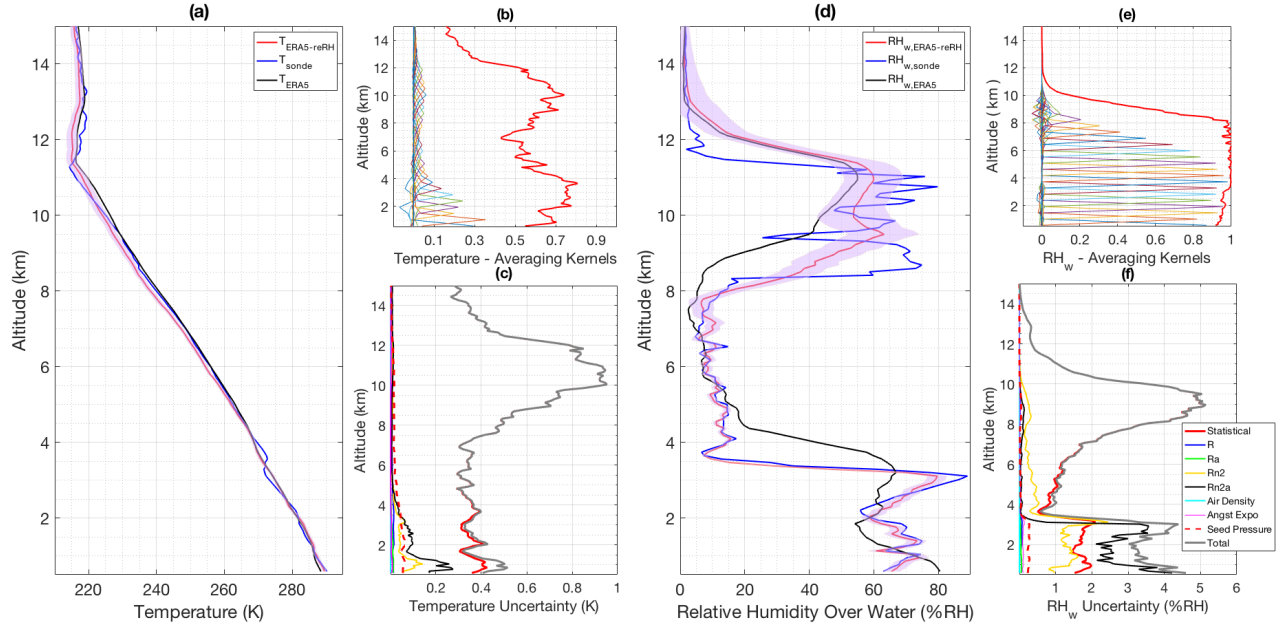


Figure 3.4: (a) The OEM-retrieved temperature profile and the statistical uncertainty (red curve and shaded area) from RALMO measurements on 28 August 2012 with 30 min temporal and 90 m vertical resolutions. The blue curve is the radiosonde measurement. The sonde was launched at 2241 UT. The black curve is the ERA5 bias corrected *a priori* temperature profile used by the OEM. (b) Averaging kernels for temperature retrievals. The red curves show the response functions. For clarity averaging kernels for every fifth altitude bin the retrieval grid are shown. (c) Random and systematic uncertainties due to the forward model parameters for the temperature retrievals. Total uncertainty (ash curve), statistical uncertainty (red curve),  $R_{PRR,d}$  digital coupling constant for PRR (blue curve),  $R_{PRR,a}$  analog coupling constant for PRR (green curve),  $R_{wv,d}$  digital coupling constant for WV/N2 (yellow curve),  $R_{wv,a}$  analog coupling constant for WV/N2 (black curve), air density (cyan curve), Ångström exponent (purple curve), and seed pressure (red-dash curve). (d) The OEM-retrieved Relative humidity profile and the statistical uncertainty (red curve and shaded area). (e) Averaging kernels for relative humidity retrievals. (f) The total uncertainty budget for the relative humidity retrievals.

measurements contribution up to 12 km. Above 12 km the lidar impact drops quickly and the ERA5-reRH becomes identical to ERA5. Previous work by the authors (Sica and Haeefe, 2015, 2016; Farhani et al., 2018; Mahagammulla Gamage et al., 2019) used a cutoff height of 0.9 in the response function for each retrieval. However, here we are not interested in a fair representation of the measurement but in the best combination of measurement and ERA5 reanalysis and hence such a cut-off does not make sense. Unlike for temperature, the measurement response for relative humidity (red curve in Fig. 3.4(e)), is greater than 90% on the measurements up to about 8 km. The difference in lidar impact on temperature and humidity is related to the fact that the ERA5 reanalysis assimilates many temperature data sets while there are fewer humidity data sets available. Hence the background error is much smaller for temperature than for relative humidity which reduces the impact of the lidar data on temperature.

One of the main features of ERA5-reRH is the full uncertainty budget on a profile-by-profile basis that contains both random and systematic uncertainties. Fig. 3.4(c) shows the uncertainty budget for the temperature retrievals. Measurement noise is the dominant source of uncertainty where the total temperature uncertainty is less than 0.5 K up to 8 km in height and maximizes around 8 to 12 km to a value of 0.9 K. The second most important contribution is the uncertainty of the analog coupling constant for WV/N<sub>2</sub> channels. Below 2 km the uncertainty due to analog coupling constant for WV/N<sub>2</sub> channels is in the order of 0.3 K and it drops to about 0.02 K above. Contribution from each of the other forward model parameters such as seed pressure, digital and analog coupling constants from PRR channels, Ångström exponent, and etc. to the temperature uncertainty is less than 0.1 K.

The full relative humidity uncertainty budget is shown in Fig. 3.4(f). For all height the total uncertainty is in the order of 5%RH and maximum of 5.15%RH is observed at about 9 km. Below 3 km contribution from the analog coupling constant for WV/N<sub>2</sub> channels is about 4%RH. Above 3 km the statistical uncertainty dominates most to the total uncertainty and it is as same order as the total uncertainty.

### **Case 2: Daytime, clear sky, 10 September 2011 1010 - 1040 UT**

For the second case study measurements from the coincident radiosonde from Payerne that was launched at 1010 UT is used to determine the model parameters and for comparisons. During

the time of the measurements sky conditions remained clear but the signal to noise ratio (SNR) of the RALMO daytime water vapor measurements drop below 1 at 5 to 6 km.

Day time ERA5-reRH retrieved temperature (red curve) shown in Fig. 3.5(a) is in good agreement with both sonde temperature (blue curve) and bias corrected ERA5 temperature (black curve) in most heights. The ERA5-reRH retrieved temperature is about 1.5 K colder than both sonde and bias corrected ERA5 temperatures in between 2 to 3 km region. Similar to the nighttime case study, the measurement response function for temperature (red curve) in Fig. 3.5(b) drops below 0.5 at 7 km. Up to 6 km ERA5-reRH temperature retrieval depends about 70% on measurements.

Daytime ERA5-reRH retrieved relative humidity (red curve) shown in Fig. 3.5(d) is in good agreement with the sonde measured relative humidity (blue curve) below 2 km. Above 2 km ERA5-reRH retrieved relative humidity closely follow the same trend as sonde relative humidity measurements. Below 8 km in height ERA5 (black curve) is significantly different than the sonde measurements. From the measurement response function for relative humidity (red curve) shown in Fig. 3.5(e), it is evident that the ERA5-reRH retrieved relative humidity depends to over 90% of lidar measurements up to about 5 km. As the lidar water vapor signals gets weaker the relative humidity retrievals starts to rely more on the *a priori* relative humidity profile. Above 5 km we can clearly see that the ERA5-reRH retrieved relative humidity becomes identical to ERA5.

The temperature and the relative humidity uncertainty budgets are shown in Fig. 3.5(c) and (f) respectively. The total temperature uncertainty is in the order of  $\sim 0.5$  K for most heights and it maximize to  $\sim 0.9$  K from 10 to 12 km. Uncertainty due to the water vapor calibration factors ( $R_{PRR,a}$  and  $R_{PRR,d}$ ) dominates the total temperature uncertainty below 1 km ( $\sim 0.5$  K). Everywhere else the statistical uncertainty dominates the total uncertainty. Uncertainty from other model parameters lies in the order of  $\sim 0.1$  K each.

The total relative humidity uncertainty is less than 7%RH for all heights. The maximum value of  $\sim 7\%$ RH is observed below 1 km and uncertainty again peaks to about  $\sim 6$  to  $6.5\%$ RH between 5.5 to 6 km in height. Uncertainty due to digital WV/N<sub>2</sub> coupling constant ( $R_{WV,d}$ ) is dominant below 3 km ( $< \sim 5\%$ RH) while the statistical uncertainty dominates above. Uncertainty due to other model parameters are in the order of  $< 1\%$ RH each for all heights.

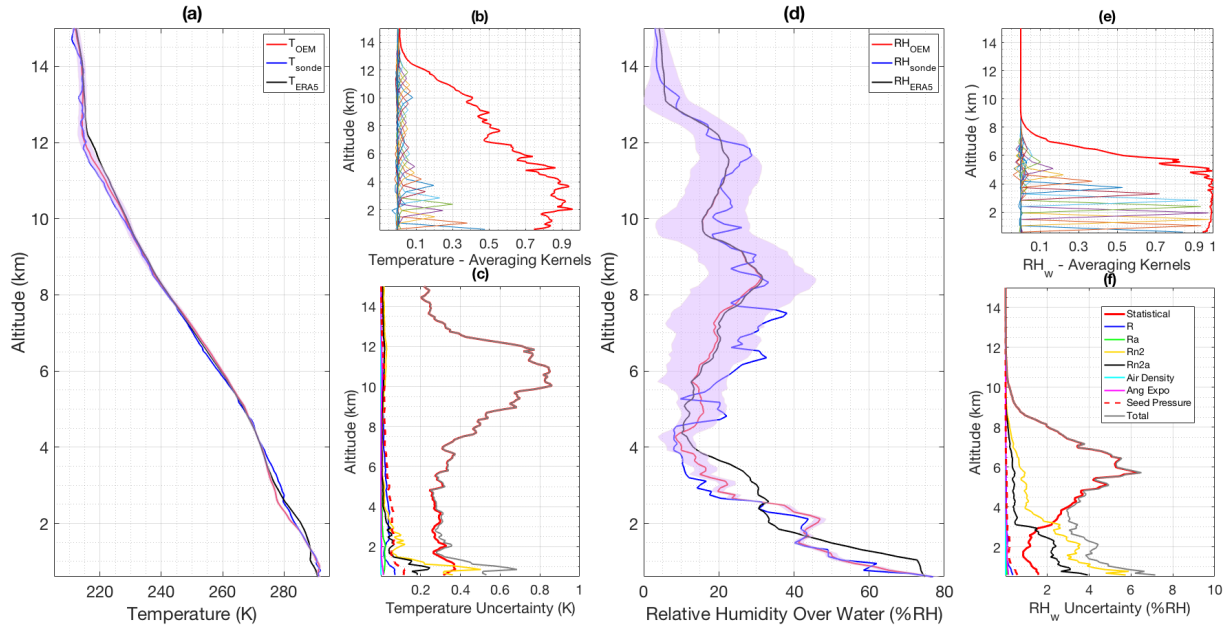


Figure 3.5: (a) The OEM retrieved temperature profile and the statistical uncertainty (red curve and shaded area) from RALMO measurements on 10 September 2011 with 30 min temporal and 90 m vertical resolutions. The blue curve is the radiosonde measurement. The sonde was launched at 1010 UT. The black curve is the ERA5 bias corrected *a priori* temperature profile used by the OEM. (b) Averaging kernels for temperature retrievals. The red curves show the response functions. For clarity averaging kernels for every fifth altitude bin the retrieval grid are shown. (c) Random and systematic uncertainties due to the forward model parameters for the temperature retrievals. Total uncertainty (ash curve), statistical uncertainty (red curve),  $R_{PRR,d}$  digital coupling constant for PRR (blue curve),  $R_{PRR,a}$  analog coupling constant for PRR (green curve),  $R_{wv,d}$  digital coupling constant for WV/N<sub>2</sub> (yellow curve),  $R_{wv,a}$  analog coupling constant for WV/N<sub>2</sub> (black curve), air density (cyan curve), Ångström exponent (purple curve), and seed pressure (red-dash curve). (d) The OEM retrieved Relative humidity profile and the statistical uncertainty (red curve and shaded area). (e) Averaging kernels for relative humidity retrievals. (f) The total uncertainty budget for the relative humidity retrievals.

### 3.4 1D Var retrieval of relative humidity from Raman lidar and U.S. standard atmospheric model

In order to be able to demonstrate the benefit of combining Raman lidar with ERA5, we repeated the same processing using the U.S. standard atmospheric climatology instead of ERA5. Apart from the *a priori* and *a priori* error covariance matrix, all parameters are kept the same and this data set is referred to as RALMO, indicating that this is essentially a pure lidar measurement, since the choice of a large error covariance greatly reduces the impact of the *a priori*.

We use an *a priori* temperature error covariance matrix with a standard deviation of 35 K and off-diagonal elements parametrized using a tent function with 1 km correlation length. The U.S. Standard Atmosphere model temperature serves as a priori profile. The *a priori* relative humidity profile is constant in height with a value of 50%RH. The error covariance matrix is constructed in the same way as for temperature with a standard deviation of 100%RH and a correlation length of 1 km.

Two cases studies and a time series of relative humidity retrievals using RALMO OEM scheme are presented in Appendix C.

## 3.5 Results

### 3.5.1 Validation of the reanalysis against radiosonde measurements

In this section we provide a comparison of ERA5, RALMO, and ERA5-reRH temperature and relative humidity profiles with coincident sonde measurements to evaluate the improvements in the ERA5-reRH retrievals. The comparison includes 14 nighttime and 6 daytime cases from 2011 to 2015. The dates that are used in the comparison were not effected by precipitation, thick cloudy conditions or missing data.

Calibration of the lidar is performed with respect to coincident sonde measurements for all 20 day and nighttime profiles to estimate coupling constants for temperature (Mahagama-mulla Gamage et al., 2019).

### Nighttime

Figure 3.6 shows the differences (black curves) between ERA5, RALMO and ERA5-reRH with respect to coincident sonde temperatures in terms of temperature and relative humidity for nighttime. The red curve is the mean of the 14 differences and the green shaded area shows the standard deviation. To improve readability we call the mean of the differences between ERA5/RALMO/ERA5-reRH and sonde the bias and the standard deviation of the differences of ERA5/RALMO/ERA5-reRH and sonde the spread.

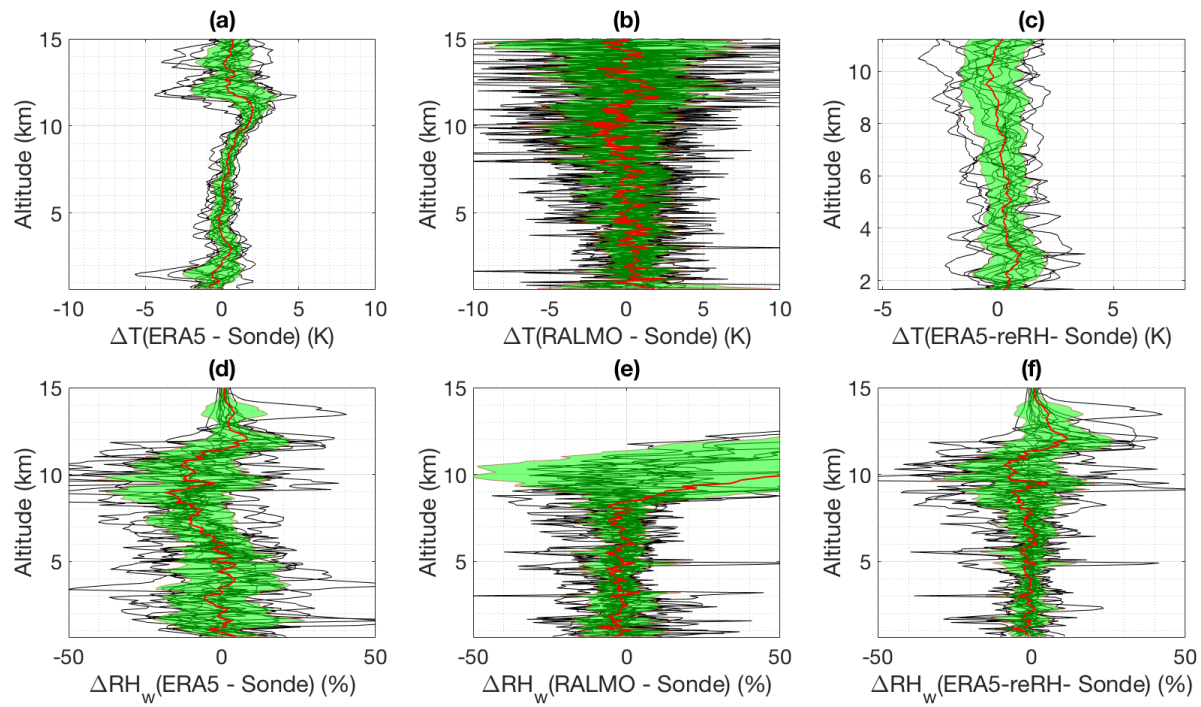


Figure 3.6: (a) Nighttime temperature differences between bias corrected ERA5 and sonde measurements in black curves for 14 nights. (b) Temperature difference between nighttime RALMO retrievals and sonde measurements. (c) Temperature difference between nighttime ERA5-reRH retrievals and sonde measurements. (d) Relative humidity differences between nighttime bias corrected ERA5 and sonde measurements in ash curves for 14 nights. (e) Relative humidity difference between nighttime RALMO retrievals and sonde measurements. (f) Relative humidity difference between nighttime ERA5-reRH retrievals and sonde measurements. Red curves and green shaded areas show the mean and the standard deviations of the temperature and relative humidity differences of each subplot.

In comparison with ERA5 and ERA5-reRH (Fig. 3.6 subplots (a) and (c)), RALMO temperatures show more scatter. There is a significant warm bias in ERA5 between 10 to 12 km

that is not detected in the RALMO and ERA5-reRH data sets. For a quantitative comparison we have shown bias and spread of the three data sets (red curves and green shaded areas shown in each subplot (a), (b), and (c) in Fig. 3.6) in Fig. 3.7.

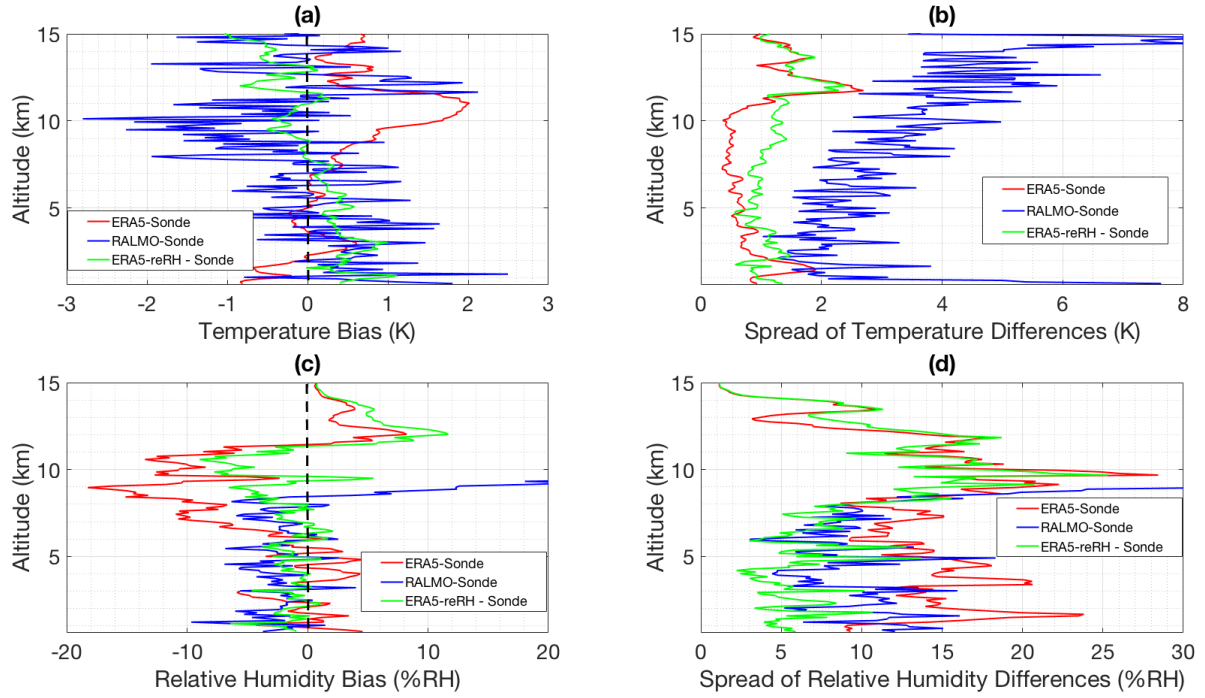


Figure 3.7: (a) Nighttime temperature biases of ERA5-Sonde (red curve), RALMO-Sonde (blue curve), and ERA5-reRH -Sonde (green curve) (b) Spread of the temperature differences. (c) Nighttime relative humidity biases. (d) Spread of the relative humidity differences.

Figure 3.7 subplots (a) and (b) show the nighttime temperature bias and spread of ERA5 (red curve), RALMO (blue curve), and ERA5-reRH (green curve). Below 4 km the temperature bias of RALMO and ERA5-reRH follow the same trend. This indicates that the ERA5-reRH temperature retrievals rely more on the lidar measurements in the heights below 4 km. Even though both RALMO and ERA5-reRH temperature retrievals rely on lidar measurements in heights below 4 km the spreads are significantly different. The spread of RALMO is in the range of 1 to 7 K and the spread of the ERA5-reRH is in the range 0.5 to 1.5 K for heights below 4 km. This large difference in spread is due to the use of different *a priori* temperature profiles and error covariances.

ERA5 shows a significant warm bias between 0.5 to 2 K in the heights above 8 km. In the same height range RALMO's bias varies from -3 to 2 K and the ERA5-reRH bias is in the range



of -1 to 0.5 K. Overall, above 8 km ERA5-reRH temperature has the smallest bias compared to ERA5 and RALMO. Thus, ERA5-reRH temperatures agrees best with the coincident sonde measurements.

As shown in Fig. 3.7 subplot(b), the spread of RALMO temperatures increases with height. The spread of ERA5 is smaller than the one of ERA5-reRH and RALMO except at 1 to 2 km and 11.5 to 12.5 km height ranges. In those two height regions the spread of ERA5-reRH becomes the smallest. Thus, ERA5-reRH overcome the variations in the ERA5 data and lidar measurements and retrieve an optimal temperature profile. The overall performance of ERA5-reRH temperatures retrievals are high as they agree best with the sonde measurements and also ERA5-reRH could eliminate the biases in ERA5 and lidar measurements (as seen in RALMO retrievals).

The subplots (d), (e), and (f) in Fig. 3.6 show the nighttime relative humidity differences (black curves) between ERA5, RALMO and ERA5-reRH with respect to sonde. The spread of RALMO relative humidity shown in subplot (e) in Fig. 3.6 is comparatively smaller than the spread of ERA5. However, the RALMO relative humidity retrievals are restricted to an average height of about 11 km where the RALMO retrievals response function reaches 0.9. Above the cutoff height the RALMO retrievals depend on the *a priori* and the spread (green shaded area) increases significantly. The spread of ERA5-reRH shown in subplot(f) in Fig. 3.6 is smaller compared to ERA5 and RALMO. Thus, by comparing the three subplots (d), (e), and (f) we can conclude that by assimilating the lidar into ERA5 we have improved the relative humidity retrievals in a way that they agree best with the coincident sonde measurements. Also, it is evident that above 11 km (average cutoff height of the RALMO retrievals) bias and spread of ERA5-reRH are identical to ERA5 and hence the lidar impact is negligible.

Figure 3.7 subplots (c) and (d) shows bias and spread ERA5 (red curve), RALMO (blue curve), and ERA5-reRH (green curve) in relative humidity. Below 6 km the bias of ERA5 varies between  $\pm 6\%RH$ . The bias of RALMO is between -10 to 4%RH and the bias of ERA5-reRH is in the range of -6 to 2%RH. Also, below 6 km ERA5-reRH has the smallest spread while ERA5 shows the largest spread. Therefore, below 6 km ERA5-reRH relative humidity agrees best with the coincident sonde measurements and ERA5 relative humidity is in the least agreement with the sonde measurements.

As shown in Fig. 3.7 subplots (c), from 6 to 11.5 km ERA5 has a wet bias with a maximum of 18%RH at 9 km. In the same height range ERA5-reRH shows a smaller bias than ERA5 and at 9 km the ERA5-reRH bias is about 8%RH. RALMO and ERA5-reRH biases between 6 to 9 km are about -6 to 2%RH. However, above 9 km RALMO's bias increases significantly. In terms of spread, ERA5 shows the largest values up to about 9 km followed by RALMO. Above 9 km ERA5 and ERA5-reRH both have almost the same bias and spread indicating that the lidar impact reduces quickly above this level.

### Daytime

Daytime temperature differences between ERA5-sonde, RALMO-sonde, and ERA5-reRH-sonde are given in Fig. 3.8 subplots (a), (b), and (c). The red curves in each subplot indicates the temperature bias of the differences and the green shaded area shows the spread of the temperature differences. For comparison purposes we have shown the temperature biases and the spreads of the temperature differences of ERA5-sonde (red curve), RALMO-sonde (blue curve), and ERA5-reRH-sonde (green curve) in Fig. 3.9 subplots (a) and (b).

The RALMO-sonde daytime temperature differences (black curves) shown in subplot (b) Fig. 3.8 are more spread out than ERA5-sonde and ERA5-reRH-sonde differences. As seen in subplot (a) Fig. 3.9 at heights below 4 km, RALMO and ERA5-reRH temperature biases are in the same order of  $\pm 1.5$  K and the ERA5 bias is about  $\pm 0.5$  K. The spreads of the temperature differences in subplot (b) Fig. 3.9 shows below 4 km the spreads of ERA5-sonde, RALMO-sonde, and ERA5-reRH temperature differences are about 1.6 K,  $\sim 1$  K, and 1 K respectively. Below 4 km, ERA5 temperatures agree best with the coincident sonde measurements.

From 4 to 13 km in height, bias of the temperature differences of ERA5-reRH and sonde measurements is in between the biases of ERA5 and RALMO temperature differences with sonde. This indicates both lidar measurements and ERA5 have an impact on the ERA5-reRH temperatures in 4 to 13 km region. Above 13 km, bias and the spread of the temperature differences of ERA5-reRH and sonde are same as the ERA5 and sonde indicating *a priori* has more impact on the the ERA5-reRH temperature retrievals at heights above 13 km. For all heights below 13 km, both RALMO and ERA5-reRH temperatures show a significant bias compared to the sonde measurements.

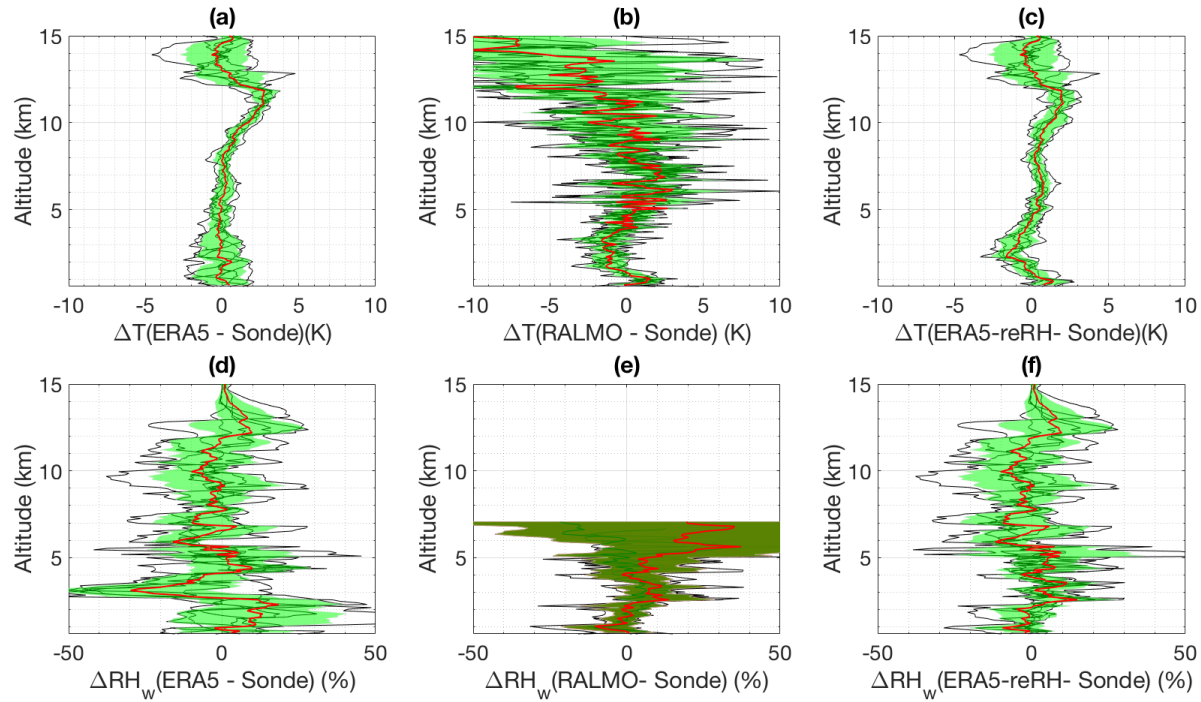


Figure 3.8: Daytime temperature differences between bias corrected ERA5 and sonde measurements in black curves for 6 days. (b) Temperature difference between nighttime RALMO retrievals and sonde measurements. (c) Temperature difference between nighttime ERA5-reRH retrievals and sonde measurements. (d) Relative humidity differences between nighttime bias corrected ERA5 and sonde measurements in ash curves for 14 nights. (e) Relative humidity difference between nighttime RALMO retrievals and sonde measurements. (f) Relative humidity difference between nighttime ERA5-reRH retrievals and sonde measurements. Red curves and green shaded areas show the mean and the spreads of the temperature and relative humidity differences of each subplot.

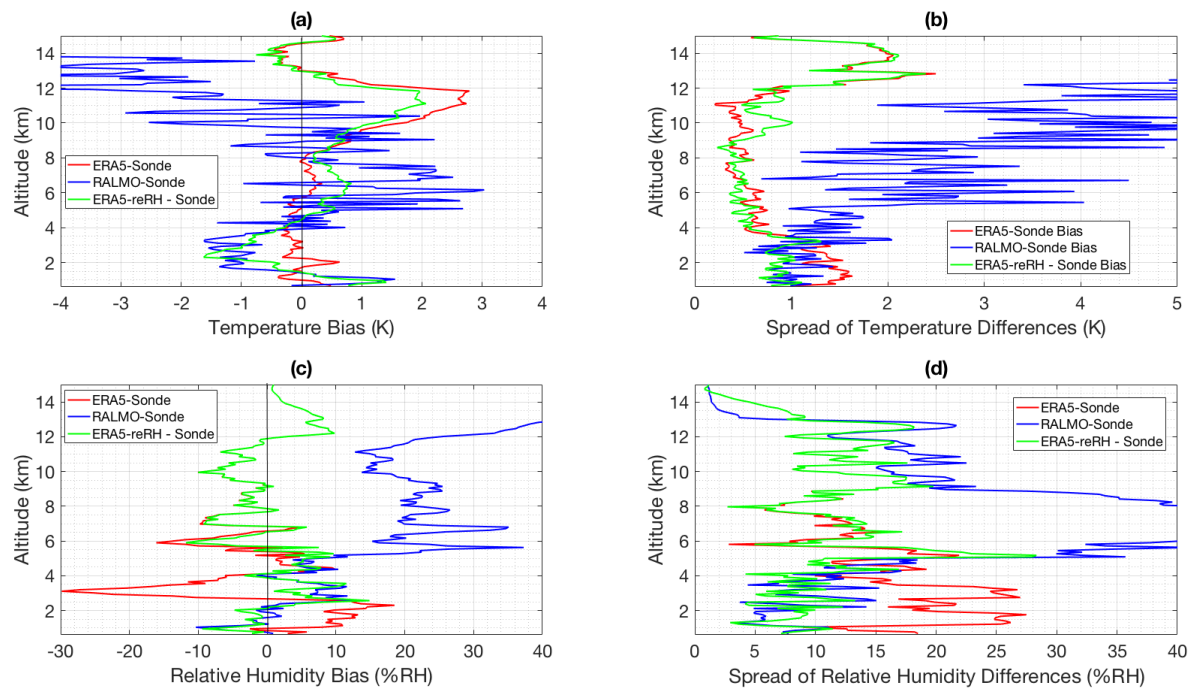


Figure 3.9: (a) Daytime temperature biases of ERA5-Sonde (red curve), RALMO-Sonde (blue curve), and ERA5-reRH -Sonde (green curve) (b) Spread of the relative temperature differences. (c) Daytime relative humidity biases.(d) Spread of the relative humidity differences.

Figure 3.8 subplots (d), (e), and (f) show the daytime relative humidity differences between ERA5-sonde (red curve), RALMO-sonde (blue curve), and ERA5-reRH-sonde (green curve). The corresponding relative humidity biases and spread of the differences are shown in Fig. 3.9 subplots (c) and (d). Unlike the ERA5 temperatures, the daytime relative humidity shows significantly large bias to sonde measurements in lower heights (see Fig. 3.9 subplot(c)). Below 2.5 km ERA5 has a dry bias with a maximum of 18%RH and from 2.5 to 4 km a large wet bias with a maximum of 30%RH are found.

Up to about 5 km in height the relative humidity biases of RALMO and ERA5-reRH with the sonde are in the same order indicating, that both relative humidity retrievals are in good agreement with the sonde measurements. Moreover, it also shows that the impact of the Raman lidar measurements on the ERA5-reRH relative humidity retrievals are high in heights below 5 km. The spreads of the relative humidity differences in Fig. 3.9 subplot (d) shows that below 5 km, the spread of the ERA5 and sonde relative humidity difference is the greatest. Thus, ERA5 relative humidity has the worst agreement with the sonde measurements in the heights below 5 km.

Above 5 km the biases and the spreads of the relative humidity differences of ERA5-sonde and ERA5-reRH-sonde start to overlap with each other, indicating the ERA5-reRH relative humidity retrievals essentially depend on the *a priori* relative humidity. Thus, above 5 km both ERA5 and ERA5-reRH relative humidity profiles are in good agreement with the sonde measurements.

### 3.6 Discussion

In this paper, we have combined Raman lidar measurements with ERA5 reanalysis data using a 1D Var data assimilation approach (optimal estimation method) to generate the ERA5-reRH temperature and humidity data set. Data assimilation of the lidar measurements into ERA5 is done using raw lidar measurements, i.e. backscatter profiles from rotational and vibrational Raman scattering, without any data treatments such as filtering or smoothing. Output of the 1D Var data assimilation process are profiles of relative humidity with respect to water and temperature above Payerne, Switzerland. The data set comes along with a full characterization

on a profile per profile basis in terms of uncertainty and vertical resolution. Prior to assimilation ERA5 temperature and relative humidity have been bias corrected using a set of special radiosoundings which have not been assimilated into ERA5. The same set of sonde has been used to determine the ERA5 background (*a priori*) error covariance matrix.

The comparison of ERA5, ERA5-reRH and RALMO temperature and relative humidity profiles with coincident sonde measurements given in Section 3.5.1 reveals that ERA5-reRH relative humidity profiles could be significantly improved compared to both ERA5 and RALMO. The lidar impact is most visible where the raw signals have a sufficient signal to noise ratio, i.e. below 5 km during daytime and below 11 km during nighttime. This improvement could be expected since ERA5 assimilates relatively few humidity data sets and since water vapor is very variable and therefore difficult to model accurately. For temperature the improvement is much less significant and most visible in the lowest layers below 2 km and at tropopause level. This result could be expected since many temperature data sets are assimilated in ERA5 and hence ERA5 quality is already excellent, apart from a bias. The ERA5 reanalysis temperature data has being improved over the last few decades (Hersbach et al., 2019) and the temperature uncertainty is in the same order as the standard uncertainty of radiosonde. Where as the ERA5 relative humidity data for the last few decades does not show any significant improvement.

In general the lidar has a positive impact on the reanalysis up to 6 km during day and up to about 11 km during night. ERA5-reRH retrieved temperature and relative humidity both agree with the coincident sonde measurements. Full uncertainty budgets showed the total ERA5-reRH temperature uncertainties for both day and night are less than 1 K for all heights. The ERA5-reRH relative humidity uncertainty is less than 4%RH for nighttime and for daytime it goes up as high as 7%RH (refer Fig. 3.4 and Fig. 3.5 subplot (f)). Statistical and WV/N<sub>2</sub> coupling uncertainties dominates the most to the total relative humidity uncertainty. For comparison, ERA5 relative humidity uncertainty is greater than 10%RH up to 10 km reaching a maximum of 25%RH at 2 km (Fig. 3.2). The relative humidity uncertainty of less than 10%RH up to the height of 2 km is reported by Wang et al. (2011), where they calculated relative humidity using water vapor mixing ratio and temperature measurements from two different Raman lidars. Also, relative humidity study by Mattis et al. (2002) report a relative uncertainty of

5 to 25%RH for relative humidity and state a 1 to 2 K temperature accuracy is required for reducing the relative uncertainty in the relative humidity to, on average, less than 10%RH. Our Era5-reRH relative humidity uncertainties for both day and nighttime retrievals are less than 10%RH with a temperature accuracy of 1 K. Thus, our findings prove the speculation by Mattis et al. (2002) is correct. Moreover, ERA5-reRH relative humidity uncertainties can be compared to the uncertainties recorded by the studies mentioned above and studies by Brocard et al. (2013), Sakai et al. (2000); thus, ERA5-reRH OEM scheme provides accurate relative humidity retrievals.

Accurate relative humidity retrievals and uncertainties are essential to detect ice supersaturation layers (ISS) in the upper troposphere. The first known study of atmospheric ISS using Raman lidar measurements was made by Comstock et al. (2004). A year's worth of nighttime Raman lidar water vapor mixing ratio measurements calibrated against microwave radiometer water vapor measurements was used with radiosonde temperature measurements to estimate relative humidity over ice ( $RH_i$ ). The study focuses on the frequency of high ISS in cirrus clouds. The results indicated that  $RH_i > 120\%$ RH frequently occurs at temperatures above  $-70^\circ\text{C}$ . With a  $\pm 1^\circ\text{C}$  temperature accuracy, Comstock et al. (2004) report the  $RH_i$  variance to be  $\sim 8\%$ RH at  $-40^\circ\text{C}$  and  $\sim 10\%$ RH at  $-60^\circ\text{C}$ . The study by Immler et al. (2008) also used a combination of Raman lidar water vapor measurements with radiosonde temperature measurements to investigate cirrus, contrails, and ice supersaturated regions in high pressure systems at northern mid-latitudes. Raman lidar measurements made from August to September in 2000, in clear sky conditions (without low and mid-level clouds) were used to estimate  $RH_i$ . The results showed that the occurrence of cirrus and ISS are closely related. They observed frequent ice supersaturated regions in the uppermost troposphere (8 km to tropopause). Further investigations of optical depths, cirrus cloud classification, and contrails were also presented in the study by Immler et al. (2008). All these studies use partial amounts of Raman lidar measurements to determine  $RH_i$ ; thus to detect ISS layers. Also, there are no studies available that provide comprehensive error analysis with their  $RH_i$ . Therefore, our ERA5-reRH OEM scheme is the first method that provides quality relative humidity retrievals with well defined uncertainties directly retrieved from Raman lidar measurements to accurately study the ISS events occur in the upper troposphere.

## 3.7 Conclusion

We have successfully assimilated ERA5-reanalysis with the Raman lidar measurements to retrieve optimum temperature and relative humidity profiles. Our ERA5-reRH retrievals are influenced by both Raman lidar measurements and ERA5 data. Both day and nighttime ERA5-reRH retrieved temperature and relative humidity profiles are in excellent agreement with the coincident radio sonde measurements. Moreover, the quality of the ERA5-reRH relative humidity retrievals are improved, and the full uncertainty budget for each retrieved profile is well-defined; thus the proposed method can be used to detect ice supersaturation events occur in both day and night-times more precisely. As the future work we will use ERA5-reRH OEM scheme to study the ice super saturation events over Payerne, Switzerland.

## Bibliography

- Ansmann, A., Wandinger, U., Riebesell, M., Weitkamp, C., and Michaelis, W.: Independent measurement of extinction and backscatter profiles in cirrus clouds by using a combined Raman elastic-backscatter lidar, *Applied optics*, 31, 7113–7131, 1992.
- Behrendt, A.: Temperature measurements with lidar, in: *Lidar*, pp. 273–305, Springer, 2005.
- Brocard, E., Philipona, R., Haeferle, A., Romanens, G., Mueller, A., Ruffieux, D., Simeonov, V., and Calpini, B.: Raman Lidar for Meteorological Observations, RALMO-Part 2: Validation of water vapor measurements, *Atmospheric Measurement Techniques*, 6, 2013.
- Comstock, J. M., Ackerman, T. P., and Turner, D. D.: Evidence of high ice supersaturation in cirrus clouds using ARM Raman lidar measurements, *Geophysical research letters*, 31, 2004.
- Dinoev, T., Simeonov, V., Calpini, B., and Parlange, M.: Monitoring of Eyjafjallajökull ash layer evolution over payerne Switzerland with a Raman lidar, *Proceedings of the TECO*, 2010.
- Dinoev, T., Simeonov, V., Arshinov, Y., Bobrovnikov, S., Ristori, P., Calpini, B., Parlange, M.,



- and Van den Bergh, H.: Raman Lidar for Meteorological Observations, RALMO-Part 1: Instrument description, *Atmospheric Measurement Techniques*, 6, 1329–1346, 2013.
- Dirksen, R., Sommer, M., Immler, F., Hurst, D., Kivi, R., and Vömel, H.: Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde, *Atmospheric Measurement Techniques*, 7, 4463–4490, 2014.
- ECMWF: Part IV : Physical processes, no. 4 in IFS Documentation, ECMWF, URL <https://www.ecmwf.int/node/18714>, 2018.
- Eriksson, P., Jiménez, C., and Buehler, S. A.: Qpack, a general tool for instrument simulation and retrieval work, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 91, 47–64, 2005.
- Farhani, G., Sica, R. J., Godin-Beekmann, S., and Haeefe, A.: Optimal Estimation Method Retrievals of Stratospheric Ozone Profiles from a DIAL Lidar, *Atmospheric Measurement Techniques Discussions*, 2018, 1–22, <https://doi.org/10.5194/amt-2018-310>, 2018.
- Ferreira, A. P., Nieto, R., and Gimeno, L.: Completeness of radiosonde humidity observations based on the Integrated Global Radiosonde Archive, *Earth System Science Data*, 11, 603–627, 2019.
- Gensch, I., Bunz, H., Baumgardner, D., Christensen, L., Fahey, D., Herman, R., Popp, P., Smith, J., Troy, R., Webster, C., et al.: Supersaturations, microphysics and nitric acid partitioning in a cold cirrus cloud observed during CR-AVE 2006: an observation–modelling intercomparison study, *Environmental research letters*, 3, 035 003, 2008.
- Gierens, K., Spichtinger, P., and Schumann, U.: Ice supersaturation, in: *Atmospheric Physics*, pp. 135–150, Springer, 2012.
- Hennermann, K. and Berrisford, P.: ERA5 data documentation, Copernicus knowledge base, 2017.
- Hersbach, H. and Dee, D.: ERA5 reanalysis is in production, ECMWF newsletter, 147, 2016.

- Hersbach, H., Bell, W., Berrisford, P., Horányi, A., J., M.-S., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., and Dee, D.: Global reanalysis: goodbye ERA-Interim, hello ERA5, ECMWF Newsletter, pp. 17–24, <https://doi.org/10.21957/vf291hehd7>, URL <https://www.ecmwf.int/node/19027>, 2019.
- Hyland, R.: Formulations for Thermodynamic Properties of the Saturated Phases Of H<sub>2</sub>O from 173.15 K to 473.15 K, *Ashrae transactions*, A, 2, 500–513, 1983.
- Immler, F., Treffeisen, R., Engelbart, D., Krüger, K., and Schrems, O.: Cirrus, contrails, and ice supersaturated regions in high pressure systems at northern mid latitudes, *Atmospheric Chemistry and Physics*, 8, 1689–1699, 2008.
- Kovalev, V. A. and Eichinger, W. E.: Elastic lidar: theory, practice, and analysis methods, John Wiley & Sons, 2004.
- Lenschow, D. H., Wulfmeyer, V., and Senff, C.: Measuring second-through fourth-order moments in noisy data, *Journal of Atmospheric and Oceanic Technology*, 17, 1330–1347, 2000.
- Lompar, M., Lalić, B., Dekić, L., and Petrić, M.: Filling Gaps in Hourly Air Temperature Data Using Debiased ERA5 Data, *Atmosphere*, 10, 13, 2019.
- Mahagammulla Gamage, S., Sica, R. J., Martucci, G., and Haeferle, A.: Retrieval of Temperature From a Multiple Channel Pure Rotational Raman-Scatter Lidar Using an Optimal Estimation Method, *Atmospheric Measurement Techniques Discussions*, 2019, 1–34, <https://doi.org/10.5194/amt-2019-107>, URL <https://www.atmos-meas-tech-discuss.net/amt-2019-107/>, 2019.
- Mattis, I., Ansmann, A., Althausen, D., Jaenisch, V., Wandinger, U., Müller, D., Arshinov, Y. F., Bobrovnikov, S. M., and Serikov, I. B.: Relative-humidity profiling in the troposphere with a Raman lidar, *Applied optics*, 41, 6451–6462, 2002.
- Miloshevich, L. M., Vömel, H., Paukkunen, A., Heymsfield, A. J., and Oltmans, S. J.: Characterization and correction of relative humidity measurements from Vaisala RS80-A radiosondes at cold temperatures, *Journal of Atmospheric and Oceanic Technology*, 18, 135–156, 2001.

- Miloshevich, L. M., Vömel, H., Whiteman, D. N., and Leblanc, T.: Accuracy assessment and correction of Vaisala RS92 radiosonde water vapor measurements, *Journal of Geophysical Research: Atmospheres*, 114, 2009.
- Nicolet, M.: On the molecular scattering in the terrestrial atmosphere: An empirical formula for its calculation in the homosphere, *Planetary and Space Science*, 32, 1467–1468, 1984.
- Noh, Y.-C., Sohn, B.-J., Kim, Y., Joo, S., and Bell, W.: Evaluation of temperature and humidity profiles of Unified Model and ECMWF analyses using GRUAN radiosonde observations, *Atmosphere*, 7, 94, 2016.
- Pappalardo, G., Amodeo, A., Pandolfi, M., Wandinger, U., Ansmann, A., Bösenberg, J., Matthias, V., Amiridis, V., De Tomasi, F., Frioud, M., et al.: Aerosol lidar intercomparison in the framework of the EARLINET project. 3. Raman lidar algorithm for aerosol extinction, backscatter, and lidar ratio, *Applied optics*, 43, 5370–5385, 2004.
- Peter, T., Marcolli, C., Spichtinger, P., Corti, T., Baker, M. B., and Koop, T.: When dry air is too humid, *Science*, 314, 1399–1402, 2006.
- Sakai, T., Shibata, T., Kwon, S.-A., Kim, Y.-S., Tamura, K., and Iwasaka, Y.: Free tropospheric aerosol backscatter, depolarization ratio, and relative humidity measured with the Raman lidar at Nagoya in 1994–1997: contributions of aerosols from the Asian Continent and the Pacific Ocean, *Atmospheric Environment*, 34, 431–442, 2000.
- Sica, R. and Haeferle, A.: Retrieval of temperature from a multiple-channel Rayleigh-scatter lidar using an optimal estimation method, *Applied optics*, 54, 1872–1889, 2015.
- Sica, R. and Haeferle, A.: Retrieval of water vapor mixing ratio from a multiple channel Raman-scatter lidar using an optimal estimation method, *Applied optics*, 55, 763–777, 2016.
- Sinha, A. and Harries, J. E.: Water vapour and greenhouse trapping: The role of far infrared absorption, *Geophysical research letters*, 22, 2147–2150, 1995.
- Vérèmes, H., Payen, G., Keckhut, P., DufLOT, V., Baray, J.-L., Cammas, J.-P., de Bellevue, J. L., Evan, S., Posny, F., Gabarrot, F., et al.: A Raman lidar at Maïdo Observatory (Reunion

- Island) to measure water vapor in the troposphere and lower stratosphere: calibration and validation, *Atmospheric Measurement Techniques*, pp. Under-Review, 2016.
- Wang, Y., Hua, D., Mao, J., Wang, L., and Xue, Y.: A detection of atmospheric relative humidity profile by UV Raman lidar, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 112, 214–219, 2011.
- Whiteman, D., Melfi, S., and Ferrare, R.: Raman lidar system for the measurement of water vapor and aerosols in the Earths atmosphere, *Applied Optics*, 31, 3068–3082, 1992.
- Yau, M. K. and Rogers, R.: *A short course in cloud physics*, Elsevier, 1996.
- Yegui, W., Wenqi, J., Yuanyuan, C., Qifa, C., Weiren, L., Hanxian, F., Libin, W., and Jun, N.: Application of ERA5 Reanalysis to the Construction of Initial Conditions for WACCM Simulations, *ÅÅL*, 38, 460–468, 2018.

# Chapter 4

## Detecting ice supersaturation (ISS) layers over Payerne, Switzerland by assimilating Raman lidar measurements with ERA5 reanalysis <sup>1</sup>

### 4.1 Overview

In this chapter we present preliminary detection of ice supersaturation (ISS) layers over Payerne using an optimal estimation method which assimilated RALMO lidar measurements into ERA5 reanalysis. Supersaturation is necessary for cloud formation, for crystal and droplet growth and sedimentation, and for the formation of precipitation (Gierens et al., 2012). An ISS layer can be simply identified as a regions where relative humidity over ice ( $RH_i$ ) values are greater than 100%RH. ISS is an frequent phenomenon that is found in the atmospheric regions where the temperatures reach below 0°C. It has a small effect on the radiation flow through the atmosphere. However, Fusina et al. (2007) has shown that as soon as a thin cirrus cloud forms in ISS layers, the radiation effect grows by almost two orders of magnitude. Thus, ISS layers play an essential role in cloud formation and the evolution that makes supersaturation

---

<sup>1</sup>A version of this chapter will be submitted for publication in the near future.

important for weather and climate.

ISS layers can be found in both cloudy and cloud-free regions in the atmosphere. Various techniques such as radiosonde, satellites, aircraft, and Raman lidar are used to measure atmospheric  $RH_i$ . In our work we use Raman lidar measurements to detect ISS layers. Traditional Raman lidar techniques do not measure direct  $RH_i$ . However, study by Comstock et al. (2004) use Raman lidar water vapor mixing ratio measurements combined with radiosonde temperature measurements to estimate  $RH_i$ . The results show that  $RH_i > 120\%RH$  frequently occurs at temperatures above  $-70^\circ C$ . However, Comstock et al. (2004) does not provide uncertainties associated with the  $RH_i$  measurements. Immler et al. (2008) also used a combination of Raman lidar mixing ratios with radiosonde temperatures to investigate ISS layers. In their study they only used Raman lidar measurements are made in clear sky conditions (without low and mid-level clouds). They observed frequent ISS layers in the uppermost troposphere (8 km to tropopause). Overall, measuring ISS layers in clear and cloudy conditions is difficult due to instrumental limitations. Therefore, a clear distinction between ISS layers as cloud-free layers and ISS within clouds cannot be guaranteed in all databases (Gierens et al., 2012), especially with radiosonde ISS layer studies.

Accurate  $RH_i$  measurements are important in the upper tropospheric (UT) region to study ISS and formation. We have adapted the OEM algorithm to discussed in 3 to retrieve  $RH_i$ . We refer to this new process as 1D Var assimilation of lidar data into ERA5. The new OEM algorithm retrieves  $RH_i$ , temperature, particle extinction, dead times, overlap functions, backgrounds, and lidar ratios with a full uncertainty budget. The OEM-retrieved  $RH_i$  assimilated into ERA5 will allow us to identify the ISS layers with a better accuracy. Unlike other Raman lidar ISS studies, we do not require separate temperature measurements from sonde or other instruments. We retrieve  $RH_i$  directly from vibrational and pure rotational Raman lidar measurements (Mahagammulla Gamage et al., 2019b). Furthermore, we can use the OEM-retrieved particle extinction profiles to accurately estimate the lidar ratio profiles. Based on the lidar ratios we can distinguish between clear and cloudy conditions in the atmosphere.

We applied the OEM to RALMO 12 nighttime measurements made from June-November 2011 to calculate the frequency of ISS occurrence over Payerne. Further analysis of ISS layer thickness, temperatures and heights at which ISS layers occur, and ISS layers in clear and

cloudy conditions are also presented. In this study, we have only use a very small sample of the existing RALMO measurements to demonstrate the ability of our new method and to detect ISS layers. However, this study by far does not provide a comprehensive understanding of the ISS layer detection nor ISS layer related statistics. Therefore, in future extend our study by applying the proposed method to RALMO nighttime measurements from 2008-2019 to detect ISS layers over Payerne.

## 4.2 Method

In this study we use an optimal estimation method (OEM) to assimilate lidar data into ERA5 to retrieve relative humidity over ice ( $RH_i$ ). In a previous study we retrieved relative humidity over water from Raman lidar measurements and bias corrected ERA5 data using the OEM. The accuracy of the retrievals were greatly improved with the use of bias corrected ERA5 data as *a priori* and using *a priori* error covariances. Thus, we use the same OEM setup as presented in Mahagammulla Gamage et al. (2019b) with bias corrected ERA5 data used to retrieve relative humidity over water with a forward model that uses the saturation vapor pressure over ice. The only changes made to the forward model used in the Mahagammulla Gamage et al. (2019b) are:

- Replace the saturation vapor pressure over water (Equation 1.4) to saturation vapor pressure over ice (Equation 1.5).
- Replace the relative humidity over water terms to  $RH_i$  using Equation 1.6.

ERA5 reanalysis relative humidity is calculated over water (in temperatures above  $0^\circ\text{C}$ ), over ice (in temperatures below  $0^\circ\text{C}$ ), and over a mixed phase (which lies in between these temperatures). Hence, for studying ISS layers we calculated ERA5  $RH_i$  using ERA5 specific humidity and temperatures. We also estimated the ERA5 bias for the ERA5-derived  $RH_i$  and the error covariance for  $RH_i$  based on the differences between the lidar and RS92 radiosonde measurements as explained in Mahagammulla Gamage et al. (2019b) (see Appendix D).

### 4.3 Preliminary results

We applied our improved OEM scheme for a pilot study to detect ISS layers using bias corrected ERA5 reanalysis based on lidar measurements (henceforth ERA5-reice) to 12 nights of measurements (total of 94 hours) obtained from June-November 2011. Dates used in the study and the number of hours are shown in Table 4.1.

Table 4.1: Dates and hours of measurements used in the ISS study. The dates with the \* symbol indicates the dates that we performed the calibration to estimates coupling constants. We used 30 min co-added raw lidar signals with 90 m vertical resolution, starting from the sonde launch time for the calibration

Date	No	Number of Hours	Start and End times
20110621*	10		0000-0500, 1900-2359
20110705*	10		0000-0500, 1900-2359
20110706	10		0000-0500, 1900-2359
20110719	5		1900-2359
20110802*	10		0000-0500, 1900-2359
20110817	10		0000-0500, 1900-2359
20110901	5		1900-2359
20110909*	5		1900-2359
20110910	10		0000-0500, 1900-2359
20110913	9		0000-0500, 1900-2259
20111005*	5		1900-2359
20111116*	5		1900-2359

The grid spacing for the ERA5-reice retrieval is set to 90 m at all heights. For each night the raw lidar data have been co-added to 30 m bins in height and co-added over 30 min in time. Therefore, we use same hourly ERA5 reanalysis data as the *a priori* profiles for two consecutive iterations in the ERA5-reice algorithm. For example, if the lidar measurements were obtained from 18:00 to 18:30 UT and 18:30 to 19:00 UT, we picked 18:00 UT ERA5 reanalysis data in the ERA5-reice for both retrievals. We applied the ERA5-reice to a total of 188 time profiles and only 165 (82.5 hours) of those converged to give an optimal solution. The profiles that did not converge the measurements were affected by thick clouds. Thus, the results shown below are from 165 ERA5-reice retrieval profiles that were made using RALMO measurements from nights shown in Table 4.1. As the first step of our analysis, we selected the ERA5-reice  $RH_i$ , temperature and particle extinction retrievals from 4 to 15 km to identify ISS layers. Also,



our retrievals are constrained by the signal strength of RALMO water vapor signals. Typical 30 min coadded nighttime water vapor signal can reach up to about 12 to 15 km maximum before turning into background. Therefore, most nighttime ERA5-reice  $RH_i$  retrievals above 12 km in height have less impact from the water vapor lidar measurements and the retrievals errors are expected to be higher.

From all 165 profiles (82.5 hours) of measurements, 44 profiles contained  $RH_i$  values over 100%RH at temperature below 0°C in 4 to 15 km height region. Thus, 27% of the time an ISS region was present over Payerne in the months from June-November 2011. This finding from our study fits well with the findings from Jensen et al. (2001) based on *in situ* measurements of water vapor and temperature from aircraft campaigns that showed the occurrence frequency of ISS ranged from about 20% to 45%.

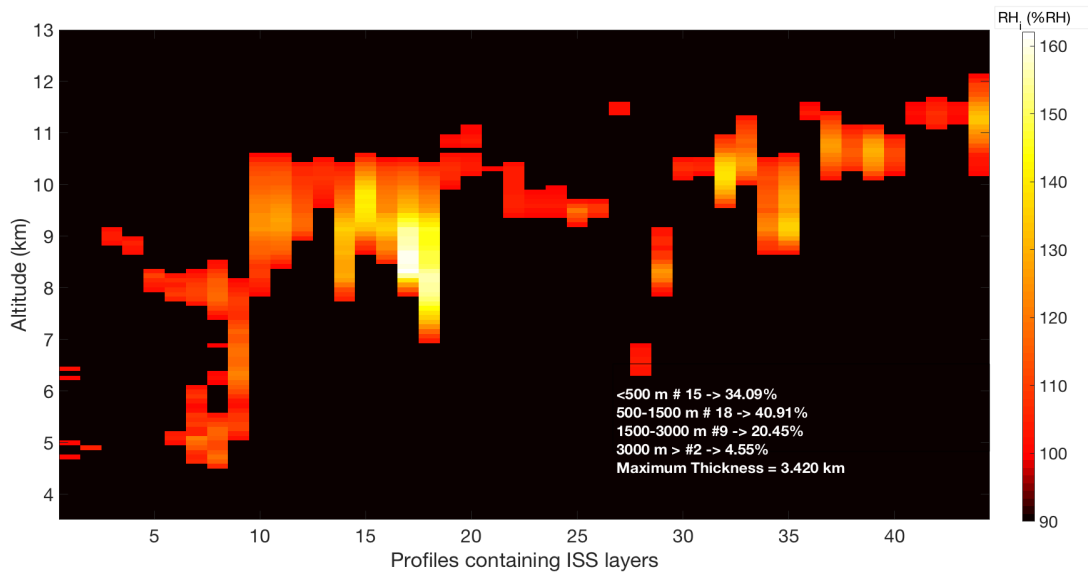


Figure 4.1: Thickness of the 44 individual ISS layers detected over Payerne, Switzerland with 90 m vertical resolution.

Thickness of the ISS layers found in the 44 individual profiles are shown in Fig. 4.1. We found 2 profiles contain ISS layers that are thicker than 3 km (5%). About 75% of the ISS layers were less than 1.5 km thick and about 20% had a thickness between 1.5 to 3 km. With our ERA5-reice retrievals we estimated an average thickness of 1070 m with 90 m for the ISS layers above Payerne. Previous studies by Spichtinger et al. (2003), Treffeisen et al. (2007), and

Rädel and Shine (2007) used radiosonde data to infer the thickness of ISS layers. The average layer thickness obtained by Spichtinger et al. (2003) is about 560 m with an vertical resolution of approximately 30 m. Treffeisen et al. (2007) and Rädel and Shine (2007) estimated a larger average vertical thickness of ISS layers at 600 to 1100 m and 1300 to 2400 m with lower vertical resolutions of 200 and 250 to 300 m respectively. A study by Dickson et al. (2009) with 10 m vertical resolution showed that more than 80% of the ISS layers are shallower than 1.5 km and roughly 5% of the ISS layers are over 3 km thick (these are the ones that satellite instruments preferentially detect). Results from our study agree with Dickson et al. (2009) and all studies agree that shallow ISS layers are more frequent than thick ones.

As seen in Fig. 4.1 most  $RH_i$  profiles contain multiple height bins with  $RH_i$  values over 100%RH. Therefore, we now refer to each height bin as an individual region; the total number of layers (bins) amounts to 23,760 (165 time profiles  $\times$  144 height bins, see Appendix D Fig. D.3). Figure 4.2 shows the histogram of all the ISS layers detected and the portion of ISS layers that occur in temperatures below  $-38^\circ\text{C}$ . Approximately 67% of the ISS layers were form at temperatures below  $-38^\circ\text{C}$ . Less than 1% of ISS layers recorded  $RH_i$  values greater than 160%RH and all those layers occur at temperatures between  $-34$  to  $-32^\circ\text{C}$ .

Further analysis of temperatures at which ISS layers are formed is presented in Fig. 4.3. About 40% of ISS layers occur at temperatures between  $-50$  to  $-40^\circ\text{C}$ . About 18%, 23%, and 16% of ISS layers were found in temperatures between  $-60$  to  $-50^\circ\text{C}$ ,  $-40$  to  $-30^\circ\text{C}$ , and below  $-30^\circ\text{C}$  respectively.

Although it is expected that most ISS layers occur in between the 4 to 15 km height region in the atmosphere where the sub zero temperatures exist, we further investigated the altitudes where ISS events are most prominent. Fig. 4.4 shows that 44% of the ISS events are in the range of 8 to 10 km height; similar to the findings of Immler et al. (2008), at this altitude, the typical atmospheric temperature is about  $-60$  to  $-40^\circ\text{C}$ . In our pilot study no ISS layers are found in heights above 14 km and only about 0.3% found in heights between 12 to 14 km. However, it is possible to have ISS layers in the heights above 14 km.

We used the ERA5-reice retrieved particle extinction profiles and aerosol scattering ratios calculated using RALMO elastic and PRR measurements to estimate lidar ratios profile-by-profile basis (Mahagammulla Gamage et al., 2019a). We use the calculated lidar ratios to

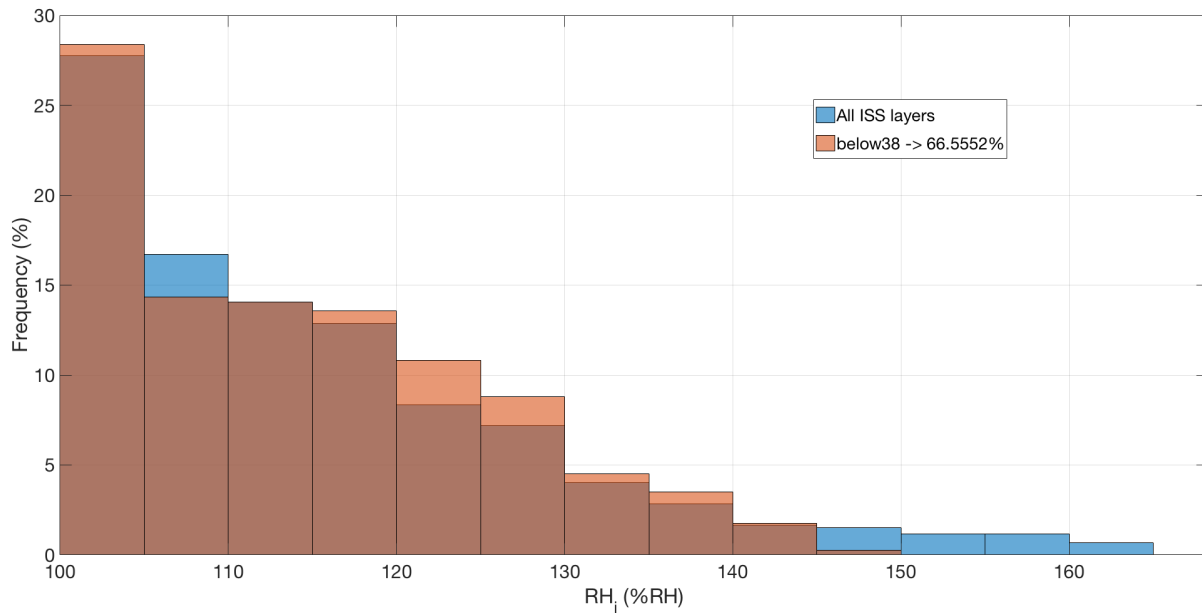


Figure 4.2: The frequency of ISS layers that occur at temperatures below freezing point are shown in blue bars while layers made from super-cooled water (temperatures below  $-38^{\circ}\text{C}$ ) are represented by the orange bars. The dark orange color indicates areas of the histogram where overlap between the red and orange bars occur.

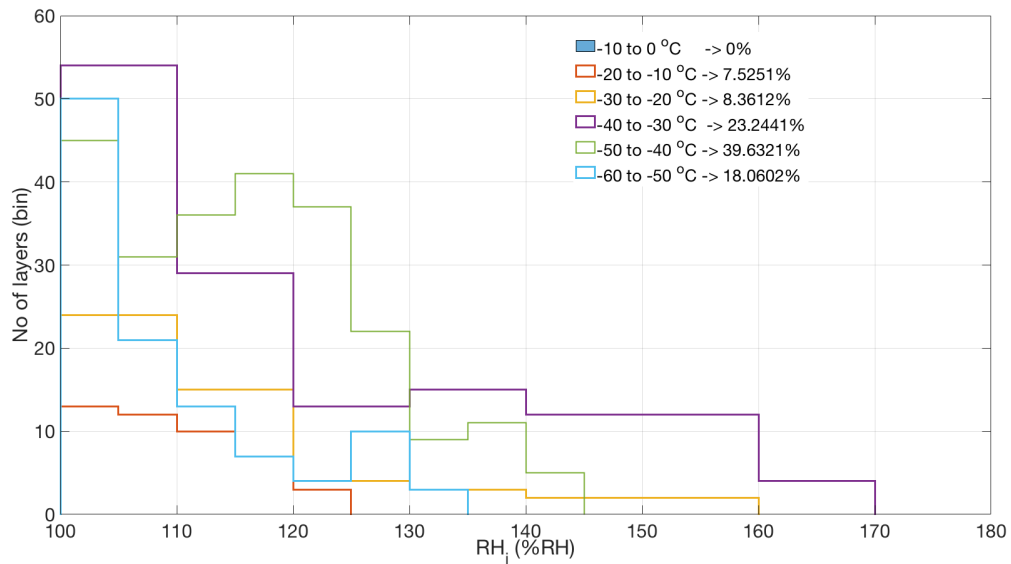


Figure 4.3: Number of ISS layers occur at every 10 degree temperature ranges below freezing temperature.

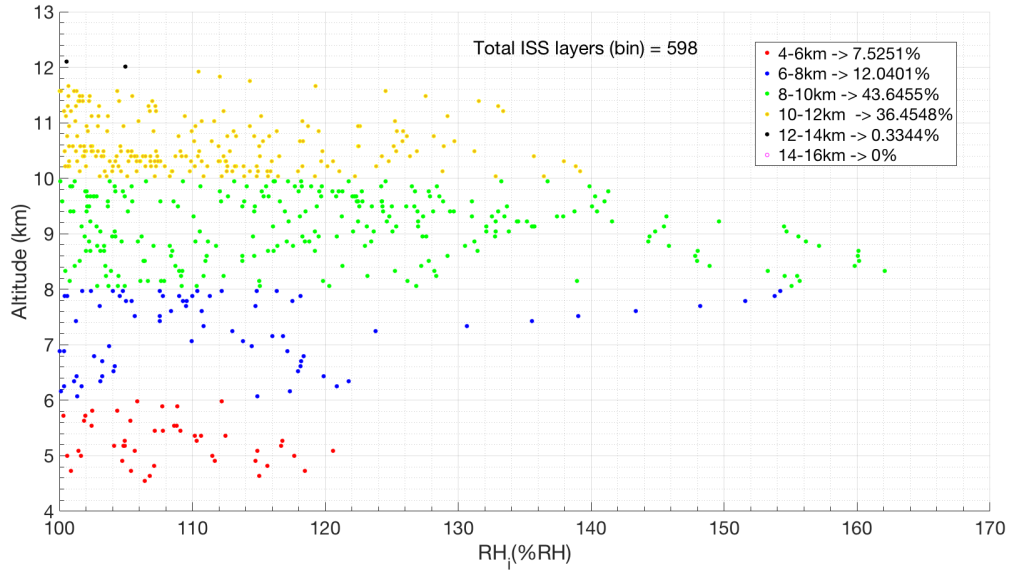


Figure 4.4: ISS layers presented at different height layers in the atmosphere. Red scatters shows the ISS layers occur between 4 to 6 km, blue scatters between 6 to 8 km, green scatters between 8 to 10 km, yellow scatters between 10 to 12 km, and black scatters between 12 to 14 km. No ISS layers are detected above 14 km.

determine the clear/cloudy conditions in the atmosphere when the  $RH_i$  measurements were made. A study by Haarig et al. (2016) showed the mean lidar ratio of  $31 \pm 5$  sr for cirrus clouds and Chen et al. (2002) has found an average lidar ratio of  $29 \pm 12$  sr for all clouds that they measured in 1999 and 2000. Moreover, Chen et al. (2002) show lidar ratios are in the 20 to 40 sr range for clouds at heights of 12.5 to 15 km. Studies by Ansmann et al. (1992), Young (1995), O'Connor et al. (2004), Sakai et al. (2003) also present lidar ratio values for optically thin clouds, cirrus clouds and as well as liquid clouds all in the range of 10 to 40 sr. Therefore, in our study, height bins with lidar ratios in the range of 10 to 40 sr are referred to as clouds and anything not in that range is clear. In our study what we mean by clear sky is a cloud-free atmosphere with sufficiently small aerosol particles. the lidar ratios of typical aerosol types such as urban/industrial, biomass burning, and dust are recorded as 71 sr, 60 sr, and 42 sr respectively (Catrall et al., 2005).

As stated before, ISS layers can occur in both clear and cloudy conditions. Figure 4.5 shows the ISS layers that formed in clear and cloudy sky conditions. The left panel shows all ISS (temperature below freezing) that is present during clear and cloudy conditions. From all

the ISS layers, 72% were found in clear conditions and about 30% of those regions were in  $RH_i$  100-105%RH. From the 28% of the cloudy ISS bin layers, about 42% had  $RH_i$  values of 100-110%RH. The ISS layers forming in temperatures below  $-38^\circ\text{C}$  are shown in the right panel of the Fig. 4.5. From all the bin layers 46% of them had ISS layers occur at temperature below  $-38^\circ\text{C}$  in clear sky conditions and 20% of all bin layers had ISS layers in cloudy conditions. Overall, the majority of ISS layers we detected have occurred in clear sky conditions.

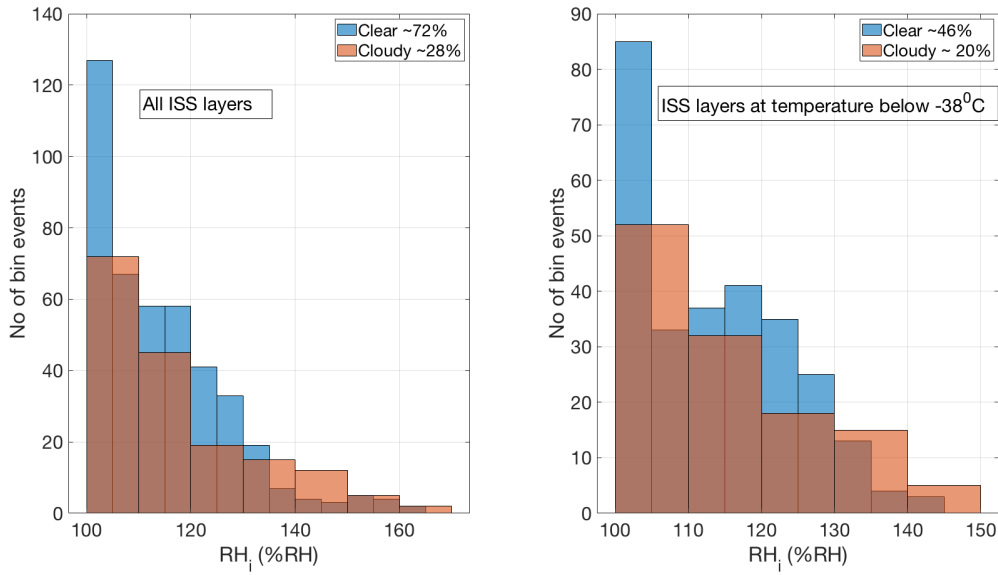


Figure 4.5: Left Panel: Number of ISS layers presence at clear and cloudy sky conditions. Right Panel: Number of ISS layers presence at temperature below  $-38^\circ\text{C}$  in clear and cloudy sky conditions.

The spatial extensions of ISS layers and their lifetimes are two other parameters that are in interest of the climate scientists. Unfortunately, we can not use lidar measurements to detect the spatial extensions. However, we can study the lifetime of ISS layers by analyzing a continuous data set. According to Gierens et al. (2012), the lifetime of an ISS region is an ill-defined notion for open dynamic systems that incessantly change their shape and size and where air is flowing in and out. These layers can perhaps stay in the atmosphere from minutes to days. Spichtinger et al. (2005) shows a case an where ISS region that stays for about 6 hours duration and states the possibility of ISS region lifetimes shorter than that. Due to the data gaps in our study we did not explicitly investigate the lifetime of ISS layers. However, in our study we

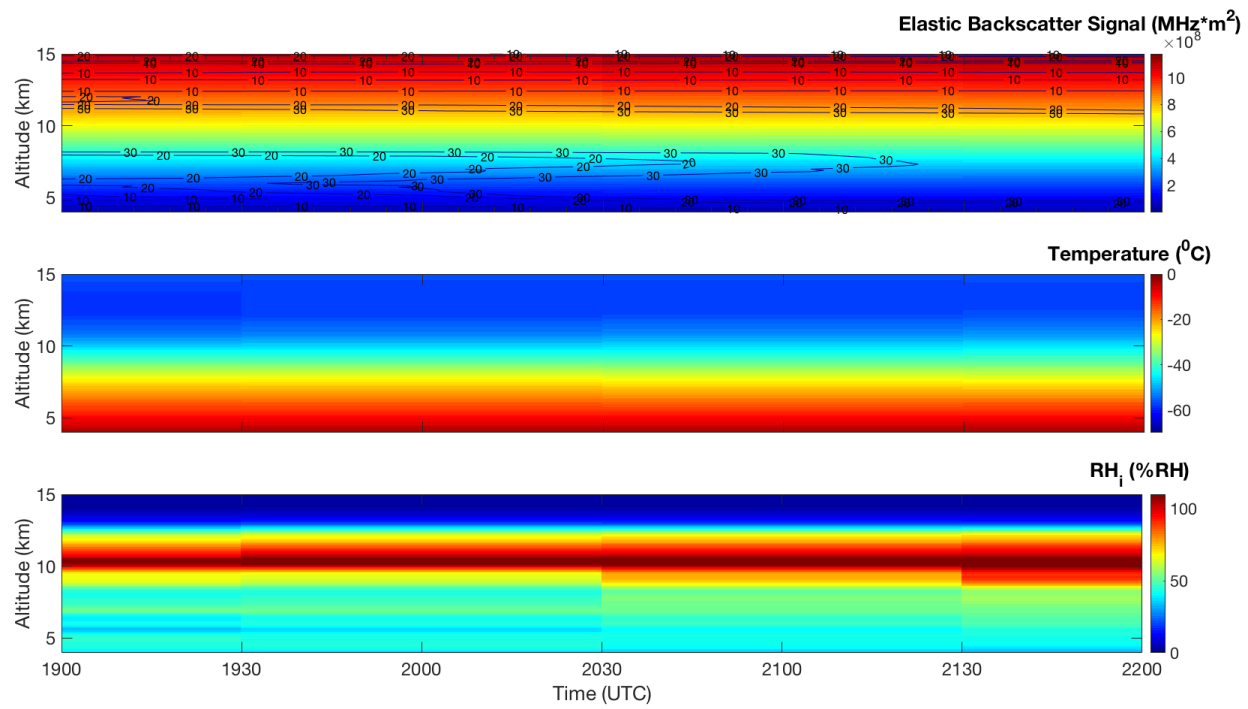


Figure 4.6: Top panel: Elastic backscatter signal of the RALMO measurements made on 01 September 2011 from 1900-2200 UT. The contour lines indicate the lidar ratio values. Middle Panel: Temperature retrievals for the 3 hours of measurements, Bottom Panel: Relative humidity over ice retrievals.

detected a few ISS layers that last for more than 2 hours. One such region is shown in Fig. 4.6.

As shown in the bottom panel of Fig. 4.6,  $RH_i$  over 100%RH were found in between the 10 to 11 km region from 1900-2200 UT. Temperatures of the ISS region stayed around -55 to -45°C. From the top panel of Fig. 4.6 we observed the lidar ratio values in the 10 to 12 km are over 20 sr. Thus, this particular ISS layer has occurred inside of a cloud. To draw more specific conclusion in future, further investigations on the lifetime of ISS layers will be performed with a continuous data set.

## 4.4 Discussion and conclusions

We have detected ice supersaturation layers over Payerne, Switzerland using an OEM scheme from the RALMO lidar nighttime measurements assimilated into ERA5 reanalysis. From a total of 165 measurements with 30 min temporal and 90 m vertical resolution in the June to November 2011 period, we found the frequency of ISS region occurrence over Payerne to be 27%. Our findings agree with the results from aircraft measurements of ISS by Jensen et al. (2001). The thickness for each ISS layer detected over the 82.5 hours periods of measurements are determined, and 5% of the ISS layers are thicker than 3 km with about 75% having a vertical thickness less than 1.5 km. Our findings agree with Dickson et al. (2009) who found roughly 5% of the ISS layers they measured to be greater than 3 km thick, and more than 80% of the ISS layers less than 1.5 km in vertical extend. Most previous studies, i.e. Jensen et al. (2001); Treffeisen et al. (2007); Rädcl and Shine (2007); Spichtinger et al. (2003) agree that shallow ISS layers form more frequently than do thick ones.

A theoretical explanation about why ISS layers are often only a few hundred meters thick are presented by Gierens et al. (2012). To explain these layers requires the assumption that ISS originated mainly from lifting in the atmosphere (Gierens et al., 2012). Gierens et al. (2012) further states that a humid air parcel with initial humidity below saturation that gets lifted adiabatically, first reaches ice saturation at a certain altitude and then reaches liquid saturation somewhat higher up. Thus, the thickness of ice saturation and liquid saturation approximately 200 to 500 m. According to Gierens et al. (2012), thickness of the ISS layers mainly decreases with temperature; thus explains thermodynamically why ISS layers are often

only a few hundred meters thick.

We have detected a few ISS layers occurring at the same time at different heights. Therefore, further statistics of ISS layers are calculated based on the total number of bins we had in both height and time of the analysis. Our study contains a total of 23760 bin layers and 598 of the bins contained  $RH_i$  over 100%RH ( 2.5%). From the total 598 ISS bin layers 67% of layers occur in temperature below  $-38^{\circ}\text{C}$ . The maximum  $RH_i$  detected in the study is about 162%RH at  $-32^{\circ}\text{C}$ . The aircraft-based studies by Jensen et al. (2005) and Popp et al. (2007) have made several ISS measurements where  $RH_i$  reached up to 230%RH in a clear sky condition and 230-250%RH in cloudy conditions respectively. The highest values of  $RH_i$ , over 1000%RH, by far given in Lübken et al. (2009) and Fan et al. (2013) are related to ice clouds at altitudes of 80 to 85 km at temperatures around  $-130^{\circ}\text{C}$ .

The heights and temperatures of the ISS layers are correlated with each other. In the troposphere temperature decreases with height; at about 10 km it reaches approximately  $-60^{\circ}\text{C}$ . Between 10 to 15 km is where we find the typical tropopause, at this height the temperatures still remains below  $-60^{\circ}\text{C}$  range. Thus, we expect ISS layers to occur mostly in the upper tropospheric region. We found about 44% of layers occur in 8 to 10 km altitude range at temperatures between  $-60$  to  $-40^{\circ}\text{C}$ . Less than 1% of the layers are found to exist above 12 km and this could be due to the limitations of the ERA5-reice water vapor data set in the upper troposphere or limitation of number profiles we consider in the pilot study.

Most ISS data sets cannot distinct between clear and cloudy conditions due to instrumental limitations (Gierens et al., 2012). Based on the estimated lidar ratio profiles we can distinguish ISS layers in clear and cloudy conditions. Bins with a LR between 10 to 40 sr are labelled as clouds; otherwise it is clear. The results showed 51% of the ISS bin layers occur in clear conditions and from total bin layers 46% of ISS layers that occur below  $-38^{\circ}\text{C}$  were found to be in clear sky conditions. Thus, we have some first evidence that ISS do not occur preferentially within ice clouds.

Even though ISS layers have a small effect on the radiation budget of the atmosphere, they still play an essential role in cloud formation. Thus, these layers have a considerable impact on modelling Earth's weather and climate system.

Our pilot study, which utilizes a limited set of nighttime lidar measurements, demonstrates



the ability of RALMO to detect ISS layers. Even though studies of ISS made using Raman lidar measurements are available our  $RH_i$  retrievals are far more accurate as they are assimilated into ERA5 reanalysis to provide an optimal estimate of the current state of the atmosphere. Our next step will be to extend our study to detect the frequency of ISS occurrence for the entire RALMO data-set from 2008 on wards. We are also interested in investigating the temporal evolution of the ISS layers as well as seasonal variability of the ISS layers. We are also interested in extending our work to detecting ISS layers to the day time measurements.

## Bibliography

- Ansmann, A., Wandinger, U., Riebesell, M., Weitkamp, C., and Michaelis, W.: Independent measurement of extinction and backscatter profiles in cirrus clouds by using a combined Raman elastic-backscatter lidar, *Applied optics*, 31, 7113–7131, 1992.
- Cattrall, C., Reagan, J., Thome, K., and Dubovik, O.: Variability of aerosol and spectral lidar and backscatter and extinction ratios of key aerosol types derived from selected Aerosol Robotic Network locations, *Journal of Geophysical Research: Atmospheres*, 110, 2005.
- Chen, W.-N., Chiang, C.-W., and Nee, J.-B.: Lidar ratio and depolarization ratio for cirrus clouds, *Applied Optics*, 41, 6470–6476, 2002.
- Comstock, J. M., Ackerman, T. P., and Turner, D. D.: Evidence of high ice supersaturation in cirrus clouds using ARM Raman lidar measurements, *Geophysical research letters*, 31, 2004.
- Dickson, N., Gierens, K., Rogers, H., and Jones, R.: Vertical spatial scales of ice supersaturation and probability of ice supersaturated layers in low resolution profiles of relative humidity, in: *Proceedings of the International Conference on Transport, Atmosphere and Climate (TAC-2)*, Aachen and Maastricht, pp. 22–25, 2009.
- Fan, Z. Y., Sica, R. J., Walker, K. A., Boone, C. D., and Bernath, P.: Hemispheric Differences in Mesopause-region Temperature Associated with the Presence of Polar Mesospheric Clouds Determined from ACE-FTS Satellite Measurements, A. Deepak Publishing, 2013.

- Fusina, F., Spichtinger, P., and Lohmann, U.: Impact of ice supersaturated regions and thin cirrus on radiation in the midlatitudes, *Journal of Geophysical Research: Atmospheres*, 112, 2007.
- Gierens, K., Spichtinger, P., and Schumann, U.: Ice supersaturation, in: *Atmospheric Physics*, pp. 135–150, Springer, 2012.
- Haarig, M., Engelmann, R., Ansmann, A., Veselovskii, I., Whiteman, D. N., and Althausen, D.: 1064 nm rotational Raman lidar for particle extinction and lidar-ratio profiling: cirrus case study, *Atmospheric Measurement Techniques*, 9, 4269–4278, 2016.
- Immler, F., Treffeisen, R., Engelbart, D., Krüger, K., and Schrems, O.: Cirrus, contrails, and ice supersaturated regions in high pressure systems at northern mid latitudes, *Atmospheric Chemistry and Physics*, 8, 1689–1699, 2008.
- Jensen, E., Smith, J., Pfister, L., Pittman, J., Weinstock, E., Sayres, D., Herman, R., Troy, R., Rosenlof, K., Thompson, T., et al.: Ice supersaturations exceeding 100% at the cold tropical tropopause: implications for cirrus formation and dehydration, *Atmospheric Chemistry and Physics*, 5, 851–862, 2005.
- Jensen, E. J., Toon, O. B., Vay, S. A., Ovarlez, J., May, R., Bui, T., Twohy, C. H., Gandrud, B. W., Pueschel, R. F., and Schumann, U.: Prevalence of ice-supersaturated regions in the upper troposphere: Implications for optically thin ice cloud formation, *Journal of Geophysical Research: Atmospheres*, 106, 17 253–17 266, 2001.
- Lübken, F.-J., Berger, U., and Baumgarten, G.: Stratospheric and solar cycle effects on long-term variability of mesospheric ice clouds, *Journal of Geophysical Research: Atmospheres*, 114, 2009.
- Mahagammulla Gamage, S., Sica, R. J., Martucci, G., and Haeefe, A.: Retrieval of Temperature From a Multiple Channel Pure Rotational Raman-Scatter Lidar Using an Optimal Estimation Method, *Atmospheric Measurement Techniques Discussions*, 2019, 1–34, <https://doi.org/10.5194/amt-2019-107>, URL <https://www.atmos-meas-tech-discuss.net/amt-2019-107/>, 2019a.

- Mahagammulla Gamage, S., Sica, R. J., Martucci, G., and Haefele, A.: 1DVar re-analysis of 10 years of ERA5 assimilating a Raman lidar for temperature and humidity, *Remote Sensing*, 11, 2019b.
- O'Connor, E. J., Illingworth, A. J., and Hogan, R. J.: A technique for autocalibration of cloud lidar, *Journal of Atmospheric and Oceanic Technology*, 21, 777–786, 2004.
- Popp, P., Marcy, T., Watts, L., Gao, R., Fahey, D., Weinstock, E., Smith, J., Herman, R., Troy, R., Webster, C., et al.: Condensed-phase nitric acid in a tropical subvisible cirrus cloud, *Geophysical Research Letters*, 34, 2007.
- Rädel, G. and Shine, K.: Evaluation of the use of radiosonde humidity data to predict the occurrence of persistent contrails, *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 133, 1413–1423, 2007.
- Sakai, T., Nagai, T., Nakazato, M., Mano, Y., and Matsumura, T.: Ice clouds and Asian dust studied with lidar measurements of particle extinction-to-backscatter ratio, particle depolarization, and water-vapor mixing ratio over Tsukuba, *Applied Optics*, 42, 7103–7116, 2003.
- Spichtinger, P., Gierens, K., Leiterer, U., and Dier, H.: Ice supersaturation in the tropopause region over Lindenberg, Germany, *Meteorologische Zeitschrift*, 12, 143–156, 2003.
- Spichtinger, P., Gierens, K., and Dörnbrack, A.: Formation of ice supersaturation by mesoscale gravity waves, *Atmospheric Chemistry and Physics*, 5, 1243–1255, 2005.
- Treffeisen, R., Krejci, R., Ström, J., Engvall, A.-C., Herber, A., and Thomason, L.: Humidity observations in the Arctic troposphere over Ny-Ålesund, Svalbard based on 15 years of radiosonde data, *Atmospheric Chemistry and Physics*, 7, 2721–2732, 2007.
- Young, S. A.: Analysis of lidar backscatter profiles in optically thin clouds, *Applied optics*, 34, 7019–7031, 1995.

# Chapter 5

## Conclusions and future work

### 5.1 Summary

I have applied an OEM to retrieve temperature and relative humidity using PRR and Vibrational Raman lidar back-scatter measurements. I used Raman lidar measurements from the Meteoswiss/EPFL Raman lidar for Meteorological Observations (RALMO), located in Payerne, Switzerland (46°48' N, 6°56' E). The RALMO transmitting system consists of a Nd:YAG laser Ultraviolet laser of 354.7 nm and four telescopes that collect backscattered photons. The telescopes deliver by backscattered light to the two polychromators of the RALMO. One polychromator is attached to a photomultiplier tube that detects the pure rotational Raman back-scattering due to atmospheric oxygen and nitrogen molecules in the high (JH) and low (JL) energy transitions. The second polychromator is connected to a photomultiplier tube that detects the rotational-vibrational Raman signals of nitrogen and water vapor (wavelengths of 386.7 and 407.5 nm, respectively).

I have presented a first-principle forward model based on the Raman lidar equation that reproduces the RALMO raw PRR measurements. In the first project in this thesis, I have successfully retrieved temperature, geometrical overlap, particle extinction, lidar constants, background counts, and dead time using multiple RALMO PRR measurements using both the analog and digital channels. The OEM has shown great advantages over the traditional temperature calculation for the PRR measurements. The OEM uses raw PRR measurements without noise filtering or any other data pre-processing, it does not require a calibration function for

temperature but allows us to calibrate using one single temperature reference measurement instead of multiple points over an extended temperature range. It can retrieve multiple other atmospheric and instrumental parameters along with temperature, provides a full uncertainty budget with both random and systematic uncertainties, provides the vertical resolution of the retrievals, and retrieval cutoff height that specify up to which height the retrieved profile is primarily due to the measurements and not the *a priori* temperature profile. It also provides averaging kernels, which are important for comparing correlative measurements and for validating measurements. Using four case studies it has been demonstrated that the OEM can retrieve temperature from RALMO PRR measurements in both day and nighttime, clear and cloudy (thin clouds and aerosol layers) sky conditions as evidenced by the excellent agreement of the temperatures with coincident radiosonde temperature measurements. My OEM temperature retrieval scheme is also applicable to any similar Raman lidar system.

The second project presented uses measurements from both the PRR and vibrational Raman channels of RALMO to retrieve relative humidity. Previous studies of relative humidity using Raman lidar measurements required estimation of temperature and water vapor mixing ratio profiles separately, which were then combined to form a relative humidity product. The forward model I developed retrieves relative humidity with respect to water directly from the Raman lidar measurements, without the need of a separate determinations of temperature or mixing ratio. Thus, my relative humidity OEM scheme does not require traditional water vapor calibrations to determine relative humidity. Our relative humidity OEM scheme was further enhanced by assimilating lidar data into ERA5 reanalysis data. ERA5 is the fifth generation of the ECMWF atmospheric reanalysis of the global climate and it combines model data with observations from all over the world. The accuracy of the ERA5 reanalysis on a particular day is higher than a forecast model. Thus, it can be used to provide accurate *a priori* information for retrieving temperature and relative humidity. ERA5 provides hourly temperature, specific humidity, relative humidity and data for many other atmospheric parameters on 37 pressure levels from the surface to 31 km. The relative humidity data product from ERA5 is calculated over water for temperatures above zero, over ice below freezing temperature (below 0°C), and over a mixture of water and ice between these temperatures. Thus, the direct output of ERA5 relative humidity cannot be used in our OEM scheme to retrieve relative humidity over water

for all heights. Therefore, we re-derived a relative humidity over water ( $RH_w$ ) product based on the ERA5 specific humidity and temperature data to use as an *a priori* profile for relative humidity in our OEM scheme. The new OEM scheme with ERA5 temperature and re-derived relative humidity is called ERA5-reRH scheme.

We compared ERA5 temperature and  $RH_w$  data with 56 sonde measurements that were launched from Payerne, Switzerland during period of 2004-2015 at times between 0600-0900 UT, 1300-1500 UT, and 1800-2100 UT. We did not use routine sonde measurements made at 1100 and 2300 UT from Payerne as most sonde observations made at standard times are assimilated in the ERA5 data. From the comparison results we observed a systematic bias in both ERA5 temperature and  $RH_w$  with respect to the sonde measurements. ERA5-derived relative humidity is greater than the sonde measurements. ERA5 temperatures showed a warm bias varying between 0 to 5 K up to about 12 km and the  $RH_w$  showed a dry bias of 0 to 15%RH in the same height range. Thus, we applied a bias correction to the ERA5 temperature and relative humidity data used in our ERA-reRH OEM scheme. We also used the standard deviations of the differences between ERA5 and the sonde measurements (the model uncertainty) to estimate correlation lengths allowing me to construct improved *a priori* covariance matrices for temperature and  $RH_w$ .

In order to be able to demonstrate the quality of the retrievals after combining Raman lidar with ERA5, we repeated the same processing using the U. S. standard atmospheric model (hereafter referred to as RALMO retrieval scheme) instead of ERA5. In comparison to both day and nighttime ERA5 data and RALMO retrievals, ERA5-reRH retrievals show the best agreement with the coincident sonde measurements. The ERA5-reRH relative humidity retrievals are improved compared to the temperature retrievals, especially in daytime. Overall, the lidar measurements have a significant impact on the ERA5-reRH relative humidity retrievals up to 6 km during day and up to about 11 km during night. Also, full uncertainty budgets showed the total ERA5-reRH temperature uncertainties for both day and night are less than 1 K for all heights. The ERA5-reRH relative humidity uncertainty is less than 4%RH for nighttime and for daytime it goes up as high as 7%RH. Our main objective of implementing ERA5-reRH is to improve the quality relative humidity retrievals to obtain an optimal estimate of the true state was achieved successfully.

The third project of this thesis presents a preliminary study of detecting the ice supersaturation (ISS) layers over Payerne using the 12 nighttime RALMO measurements that were made from June-November in 2011. ISS is a frequent phenomenon which occurs in the atmospheric regions where the temperature is below  $0^{\circ}\text{C}$ . These layers can be found inside ice crystal clouds or even outside the clouds. Typically, ISS layers can last for minutes to hours. In our work we refer to  $RH_i$  values that are over 100% in temperatures below  $0^{\circ}\text{C}$  as ISS. The forward model used in ERA-reRH OEM scheme was reformed by changing the physics of the relative humidity estimation over water to ice by changing the saturation vapor pressure calculation over water to ice to be able to retrieve relative humidity over ice ( $RH_i$ ) for all heights. The bias corrected ERA5 relative humidity was also calculated with respect to ice to include detection of ISS layers. The new ISS OEM scheme is referred to as ERA5-reice.

The ERA5-reice OEM scheme was applied to RALMO nighttime measurements. The frequency of an ISS layer occurrence over Payerne for the months of June-November, 2011 is found to be 27%. About 5% of the detected ISS layers are vertically spread over 3 km in height. About 75% of the ISS layers were less than 1.5 km thick. In some profiles we detected multiple ISS layers. Thus, each height bin in with  $RH_i$  greater than 100%RH is considered as a single ISS bin layer. From the total 598 ISS bin layers 67% of layers occur in temperatures below  $-38^{\circ}\text{C}$ . The maximum  $RH_i$  detected in the study is about 162%RH at  $-32^{\circ}\text{C}$ . About 44% of the ISS layers were present in the heights between 8-10 km. We also classified the presence of ISS layers in clear and cloudy conditions using the lidar ratio values estimated from ERA5-reice retrieved particle extinction and aerosol scattering ratio estimated from the RALMO elastic and PRR measurements. The results showed 51% of the ISS layers occur in clear conditions.

## 5.2 Conclusion

Accurate knowledge of atmospheric temperature and relative humidity trends, their associated uncertainties and physical significance are crucial to understand Earth's climate system. Often, trends are estimated based on long-term measurements made by a single or multiple instruments, sometimes at only a few locations. Space-based measurements have a large global

coverage, but have poor vertical and temporal resolutions. Observations can be combined with climate models to improve the interpretation of any trends detected. However, the estimated trends can be affected by numerous factors, such as the temporal and spatial resolution of the measurements, the uncertainties of the measurements, observational times, observational environment, methods and instruments used to make the measurements, data assimilation errors, etc. Thus, the need of making accurate atmospheric measurements and knowing the associated uncertainties is crucial. Developing new techniques to accurately retrieve or assimilate atmospheric parameters from the existing data sets is also essential. The unifying aspect of this thesis is the new information about the atmosphere that will be discovered by applying the optimal estimation method (OEM) to both the RALMO measurements obtained over the last decade and to other lidar systems which adopt our analysis methods.

This thesis provides the means of accurate estimates of the atmospheric temperature and water vapor (expressed as a relative humidity) from Raman lidar measurements. I have introduced a first principle optimal estimation method (OEM) for temperature retrievals from pure Rotational Raman (PRR) lidar measurements that works for both clear and cloudy sky conditions (provided the clouds or aerosol layers are thin enough that the lidar's laser can penetrate them) and can be applied to any Raman lidar system with PRR measurements. The new OEM technique allows retrievals of temperature and multiple other atmospheric and instrumental parameters such as particle extinction, lidar constant, and background simultaneously. Each retrieval parameter has its own estimate uncertainty budget, including both random and systematic uncertainties. Knowledge of the uncertainties is critical for atmospheric change studies. Knowledge of the uncertainties also allows a better understanding of contribution of each model parameter to the retrieval parameters. The OEM retrieved atmospheric parameters and their uncertainties can be used for trend analysis. Similarly, changes which occur in the instrument over time can be evaluated using the OEM retrieved instrumental parameters. The new OEM temperature retrieval method was tested on PRR temperature measurements from the MeteoSwiss Raman Lidar for Meteorological Observations system in different sky conditions, compared to temperature calculated using the traditional PRR calibration formulas, and validated with coincident radiosonde temperature measurements.

For the first time I have introduced an OEM scheme for direct retrieval of relative humidity



from Raman lidar measurements. This new method uses both PRR and vibrational Raman lidar measurements from both day and night time measurements. The relative humidity retrieval method does not require separate temperature or water vapor mixing ratio estimates to calculate relative humidity. The OEM relative humidity retrieval scheme can retrieve atmospheric parameters relative humidity, temperature, and particle extinction, in addition to instrumental parameters such as lidar constants, overlap, dead time, and background signal.

Reanalysis allows the reconstruction of past atmospheric data sets by assimilating global observations into existing weather and climate models. I used the OEM as a data assimilation technique to assimilate Raman lidar measurements into ERA5 reanalysis data set. I identified the biases in the ERA5 temperature and relative humidity data sets using measurements from 56 special radiosondes. My conclusion from this work is that users of ERA5 reanalysis temperature and relative humidity data must correct for the respective biases as described in Chapter 4.

The new data assimilation OEM technique introduced in thesis was used for detecting ice supersaturation layers in the upper troposphere. This new technique differentiates between ice supersaturation layers which occur in different temperature ranges, height ranges, as well as sky conditions. Unlike other existing lidar studies of ice supersaturation, my method provide a complete uncertainty estimate of the relative humidity values of each detected ice supersaturation layer.

In the near future, the OEM technique that assimilate Raman lidar measurements into ERA5 reanalysis will be used to reprocess the existing RALMO Raman lidar data set, and hopefully other Raman lidar groups will adopt this methodology. The improved temperature and relative humidity retrievals and their uncertainties can be used to improve our understanding of climate and weather.

The major outcomes of my thesis are the following.

- First principle optimal estimation method (OEM) temperature retrievals from pure Rotational Raman (PRR) lidar measurements that works for clear and cloudy sky conditions.
- Application of the OEM for direct retrieval of relative humidity using both PRR and vibrational Raman lidar measurements from both day and night times.

- Assimilation of Raman lidar measurements into the European Weather Centre for Medium-range Weather Forecast Re-analysis (ERA5) reanalysis to obtain improved estimates of relative humidity over water and over ice.

## 5.3 Future work

For future work I recommend including RALMO measurements from the two elastic channels (analog and digital) and the measurements from the additional near range telescope to the OEM schemes I have introduced. The new information will allow the temperature retrievals to be more accurate as the OEM will get additional information from the elastic measurements and the retrievals are expected to reach higher altitudes. I recommend adapting the developed OEM schemes in the MeteoSwiss routine temperature and relative humidity data products. Additionally, elastic and near field telescope measurements will also improve the particle extinction and overlap function retrievals from all the RALMO Raman channels.

Future direction for the ISS layer detection study is to apply OEM algorithm ERA5-reice to RALMO nighttime measurements from 2008-2019 and to provide a comprehensive statistical analysis. Also, after a validation of ERA5-reice daytime retrievals against coincident sonde or other available sources measurements an effort to detect ISS layers from RALMO daytime measurements will be made.

Future work should also include using the methodology I have developed to re-process the entire RALMO data set and search for trends in the temperature and relative humidity time series. The full uncertainty budgets provided with the OEM-retrieved temperature and relative humidity profiles allows accurate trend estimates. The averaging kernels of the re-processed RALMO data sets should be used for inter-comparisons with measurements from satellites or other instruments. Also, by assimilating lidar retrievals with those measurements I can further improve the accuracy of the results.

# Appendix A

## Jacobians of the three types of PRR temperature forward models

The Jacobian represent the sensitivity of the forward model to the change of the state vector (here temperature).

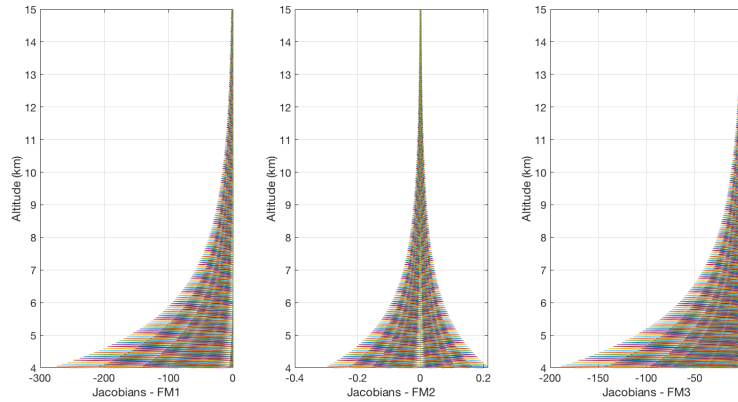


Figure A.1: Left Panel: Jacobians for FM1 where the density is replaced by ideal gas law given in Eq.2.1. Middle Panel: Jacobians for FM2. FM2 is same as given in Eq.2.1 where the only temperature dependent term in the forward model will be the differential cross-section. Right Panel: Jacobians for the FM3 where the density in the FM2 is replaced by using the ideal gas law and using differential cross-section term as a model parameter.

## Appendix B

# Conventional method of calculating relative humidity from Raman lidar measurements

Direct determination of relative humidity profiles from the Raman lidar measurements is not possible. One has to calculate temperature from pure-Rotational Raman(PRR) lidar channels and water vapor mixing ratios from water vapor Raman lidar channels simultaneously to calculate relative humidity. The Raman method for water vapor mixing ratio measurements uses two Raman signals, one of which is the back-scatter signal from the water vapor (406 nm), and the other one is the back-scatter signal from the nitrogen (386 nm) that is considered as the reference signal. Usually, a single line from the Stokes vibration-rotation of the water vapor and nitrogen spectra are detected by each channel. The water vapor mixing ratios ( $m_{wv}(z)$ ) relative to dry air can be obtained by forming the signal ratio of water vapor and nitrogen measurements that are corrected for non-linearity and background (Whiteman et al., 1992).

The traditional method of estimating water vapor mixing ratio from the Raman measurements require a determination of a calibration constant that can be determined by comparison with *in situ* water vapor measurements. Hence, the lidar water vapor mixing ratio profiles depend on the accuracy of the measurements made by the second instrument. Often, the relative humidity measurements (later converted to water vapor mixing ratios for comparison) from the coincident radiosonde are used to determine the water vapor calibration constant. Accuracy

of the water vapor calibration constant depends on the sonde measurements, and improved calibration techniques are given in the literature (Whiteman et al., 1992; Ferrare et al., 1995).

Traditional Raman lidar temperature calculation method originally proposed by Cooney (1972), also requires a ratio of signals from two PRR channels (JH higher quantum number channel, JL lower quantum number channel) that has opposite temperature dependency. Unlike water vapor and nitrogen, most PRR channels in temperature Raman lidars detect multiple back-scattered wavelengths from both nitrogen and oxygen molecules in both Stokes and anti-Stokes portions of the PRR spectrum. Thus, a suitable temperature dependent function (often called the calibration function) for the ratio of the PRR measurements needs to be approximated to calculate the temperature (Cooney, 1972; Behrendt, 2005). In practice, temperature measurements from the radiosondes that coincide with lidar measurements in space and time are used to determine the calibration coefficients of the approximated calibration function. Often, a simple form of calibration function with only two calibration coefficients. Work by Behrendt (2005), Zuev et al. (2017) and Gerasimov and Zuev (2016) shows that the uncertainty of the calculated temperature from the Raman lidar measurements depends on the approximated calibration function and as well as the accuracy of temperature measurements made by the second instrument.

Water vapor mixing ratio profiles calculated from the vibrational Raman lidar measurements can be combined with the temperature profiles calculated from the PRR lidar measurements to calculate relative humidity over water. The calculation requires estimation of vapor pressure and saturated vapor pressure. Pressure required to find the water vapor pressure, can be obtained either from routine sonde measurements or from standard atmospheric models. The saturation vapor pressure is obtained by the log inverse of Eq.1.4.

Summary of the process of estimating relative humidity from the Raman lidar measurements using the conventional method is given below.

- Raw Raman lidar measurements used in the water vapor and temperature calculations need to be corrected for non-linearity and background noise and glued.
- An estimation of a calibration constant to calculate water vapor mixing ratio using the ancillary water vapor measurements is required.

- An approximation of a calibration function and determination of calibration function coefficients are required to calculate temperature from the Raman lidar measurements.
- The water vapor pressure can be estimated using the water vapor mixing ratio profiles from the Raman lidar measurements and ancillary pressure measurements in Eq.1.35.
- The saturated vapor pressure can be calculated using Raman lidar temperature measurements in Eq.1.4.
- The calculated water vapor pressure and the saturated vapor pressure profiles can be used to calculate the relative humidity over water using the Eq.1.6.

The conventional method of calculating relative humidity from the Raman lidar measurements does not provide a complete uncertainty budget or a cutoff height that the calculated relative humidity can be trusted. Also, to improve the accuracy of both water vapor mixing ratios and temperature calculations frequent calibration is required.

## **Bibliography**

Behrendt, A.: Temperature measurements with lidar, in: Lidar, pp. 273–305, Springer, 2005.

Cooney, J.: Measurement of atmospheric temperature profiles by Raman backscatter, Journal of applied meteorology, 11, 108–112, 1972.

Ferrare, R., Melfi, S., Whiteman, D., Evans, K., Schmidlin, F., and Starr, D. O.: A comparison of water vapor measurements made by Raman lidar and radiosondes, Journal of Atmospheric and Oceanic Technology, 12, 1177–1195, 1995.

Gerasimov, V. V. and Zuev, V. V.: Analytical calibration functions for the pure rotational Raman lidar technique, Optics express, 24, 5136–5151, 2016.

Whiteman, D., Melfi, S., and Ferrare, R.: Raman lidar system for the measurement of water vapor and aerosols in the Earth's atmosphere, Applied Optics, 31, 3068–3082, 1992.

Zuev, V. V., Gerasimov, V. V., Pravdin, V. L., Pavlinskiy, A. V., and Nakhtigalova, D. P.: Tropospheric temperature measurements with the pure rotational Raman lidar technique using nonlinear calibration functions, *Atmospheric Measurement Techniques*, 10, 315–332, 2017.

## **Appendix C**

### **1D Var retrieval of relative humidity from Raman lidar and U. S standard atmospheric model**

In this section we present two case studies from the RALMO day and nighttime measurements using the 1D Var retrieval of relative humidity from Raman lidar and U.S standard atmospheric model. Due to the high variability of the atmospheric humidity and for inter-comparison purposes, we use 15 min time integrated RALMO measurements that coincident with the radiosonde measurements. To demonstrate the robust nature of the OEM we then present a 24 hours time series of the OEM retrieved relative humidity, temperature, and particle extinction using the RALMO measurements with 15 min time and 30 m vertical resolution. To reduce the processing time of the OEM we have restricted the analog measurements up to 6 km altitude (analog measurements above 6 km go to noise) and digital measurements from 2 to 28 km. The grid spacing of the retrieval grid is set to 90 m at all heights and the raw data have been co-added to 30 m bins.



## C.1 Case 1: Nighttime, clear sky, 09 September 2011 2200 - 2215 UT

A coincident radiosonde from Payerne is launched at 2200 UT; it is used for the determination of the model parameters. Fig. C.1 shows the raw data in the four PRR channels, two water vapor channels, and two nitrogen channels.

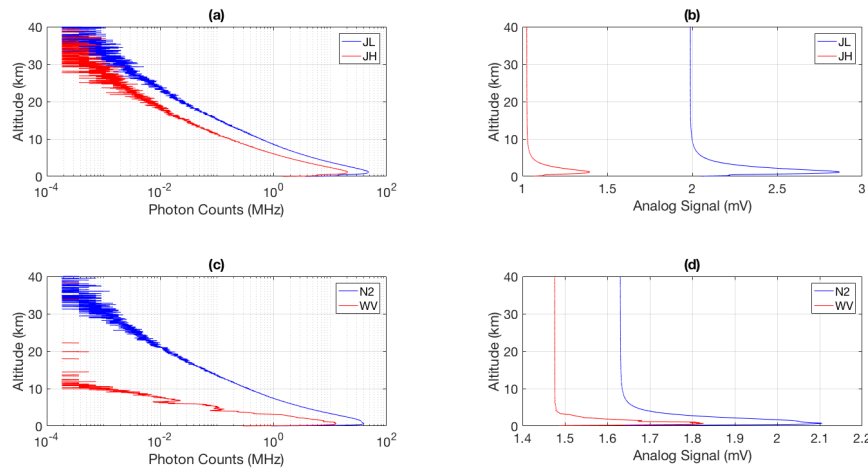


Figure C.1: Count rate for 15 min of RALMO measurements from 2200 UT on 09 September 2011, a clear night. (a) digital channels (blue curve, JL; red curve, JH). (b) analog channels (blue curve, JL; red curve, JH). (c) digital channels (blue curve, N2; red curve, WV). (d) analog channels (blue curve, N2; red curve, WV)

Standard OEM diagnostics are shown in Fig. C.2, C.3 and C.4 revealing a healthy retrieval. The retrieved profiles reach a maximum height (measurement response  $> 0.9$ ) of 25 km and 12 km for temperature and relative humidity, respectively (Fig. C.3). The vertical resolution increases with height and reaches 300 m (700 m) at the profile top for temperature (relative humidity), as illustrated in Fig. C.4.

Fig. C.5 shows the OEM retrieved temperature (left panel) and relative humidity (right panel) from the RALMO clear nighttime measurements. For comparison, the coincident sonde temperature and relative humidity measurements (blue curves) and the *a priori* profiles (green curves) are shown in each panel. In this case study, the OEM retrieved temperatures and relative humidity profiles agree with the coincide sonde measured temperatures and relative humidity

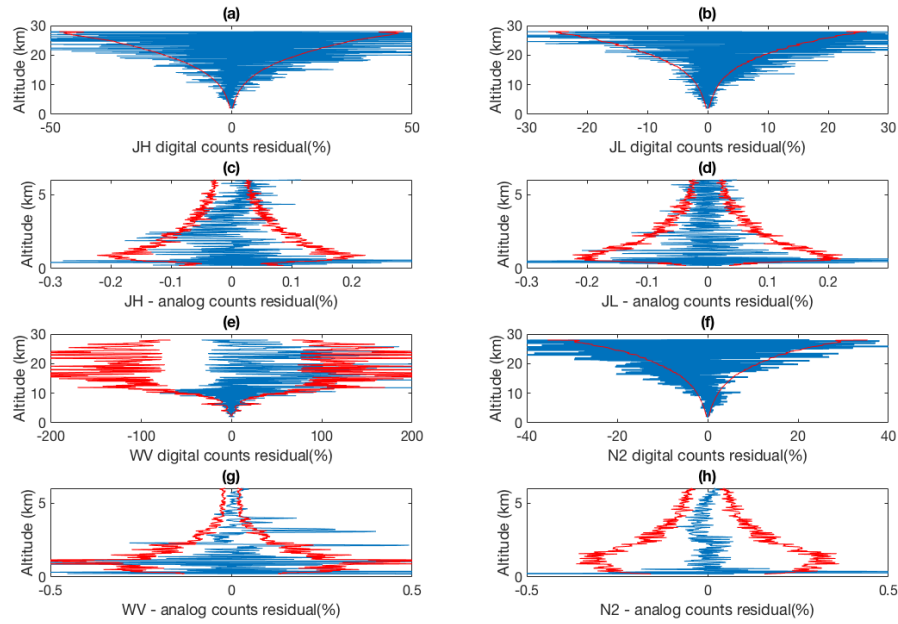


Figure C.2: Residuals between the forward model and clear nighttime RALMO measurements on 09 September 2011 for the eight channels (blue curves). The red curves show the standard deviation of the measurements. (a)JH digital (b) JL digital (c) JH analog (d) JL analog (e)WV digital (f) N2 digital (c) WV analog (d) N2 analog.

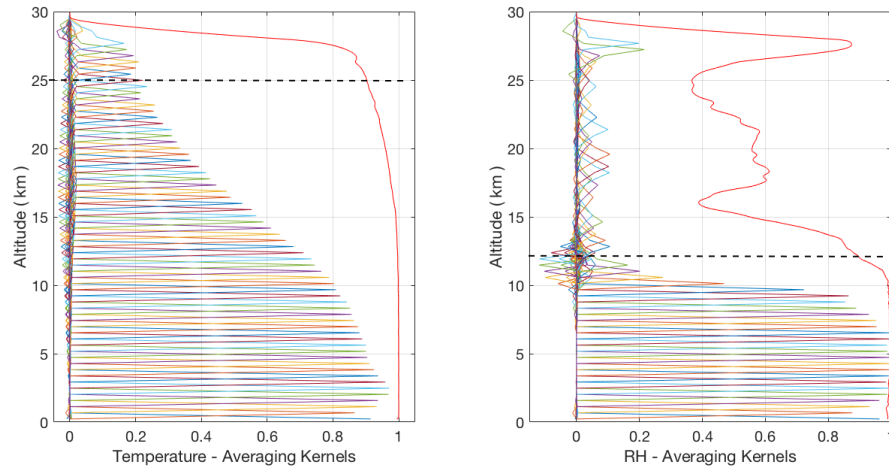


Figure C.3: Averaging kernels for temperature retrievals (left panel) and relative humidity (right panel) from the clear nighttime RALMO measurements on 09 September 2011. The red curves shows the response functions. The horizontal dashed line shows the height below which the retrieval is due primarily to the measurement and not the *a priori* profiles. For clarity averaging kernels for every fifth altitude bin the retrieval grid are shown.

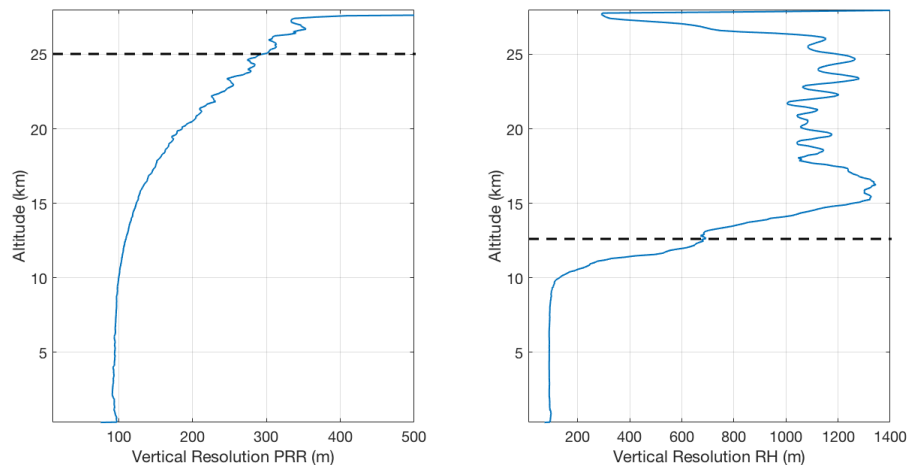


Figure C.4: Vertical resolution for temperature retrievals (left panel) and relative humidity (right panel) from the clear nighttime RALMO measurements on 09 September 2011. The horizontal dashed line shows the height below which the retrieval is due primarily to the measurement and not the *a priori* profiles.

within the statistical uncertainty up to the cutoff height of each retrieval.

One of the main features of the OEM is its full uncertainty budget that contains both random and systematic uncertainties. Fig. C.6 shows the uncertainty budget for the temperature (left panel) and relative humidity (right panel) retrievals. Measurement noise is the dominant source of uncertainty for both temperature and relative humidity followed by calibration. The total uncertainty in the upper troposphere is 1 K and 8% at 8 km.

## C.2 Case 2: Daytime, clear sky, 10 September 2011 1000 - 1015 UT

For the second case study measurements from the coincident radiosonde from Payerne that was launched at 1000 UT is used to determine the model parameters. Raw measurements from all 8 Raman channels of the RALMO are shown in Fig. C.7. The residuals of all the 8 channels (not shown) agree within the square root of the measurement covariance, indicating the goodness of the forward model to generate measurements that represent the RALMO daytime measurements.

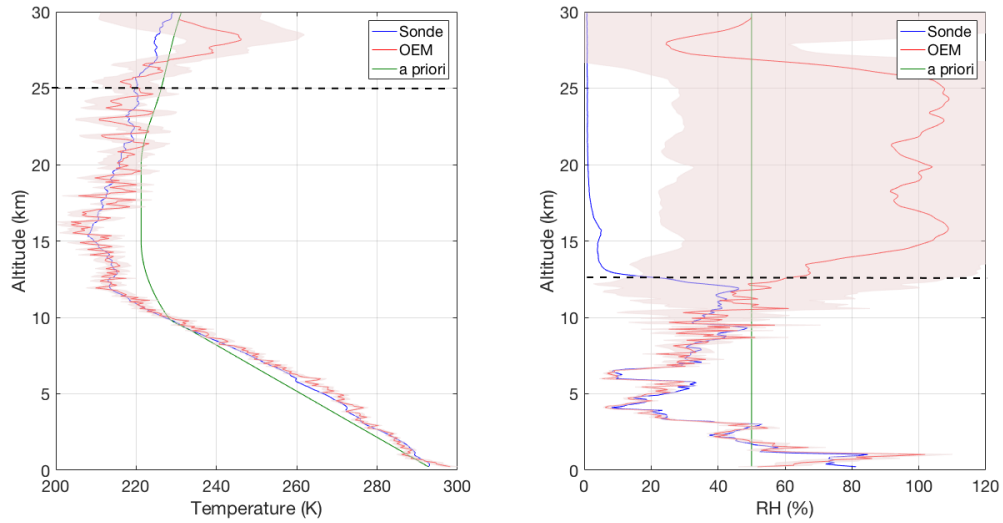


Figure C.5: Left panel: Retrieved temperature profile and the statistical uncertainty (red curve and shaded area) using the OEM from the clear nighttime RALMO measurements on 09 September 2011. The blue curve is the radiosonde measurement. The sonde was launched at 2200 UT. The green curve is the *a priori* temperature profile used by the OEM. The horizontal dashed line shows the height below which the retrieval is due primarily to the measurement and not the *a priori* temperature profile. Right panel: Retrieved relative humidity profile and the statistical uncertainty (red curve and shaded area). The blue curve is the radiosonde measurement. The green curve is the *a priori* temperature profile used by the OEM. The horizontal dashed line shows the height below which the retrieval is due primarily to the measurement and not the *a priori* relative humidity profile.

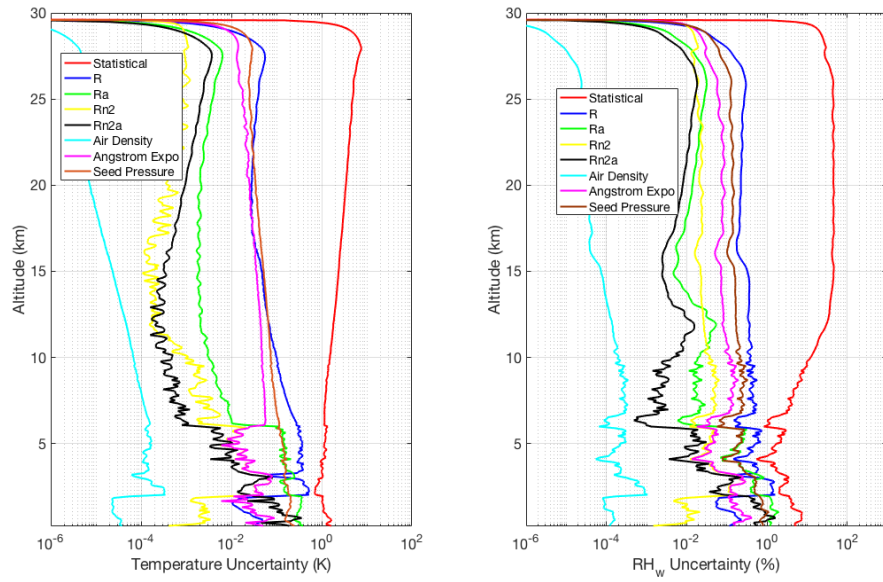


Figure C.6: Left panel: Temperature, Right panel: Relative humidity, random and systematic uncertainties due to the forward model parameters for the temperature retrievals from the clear nighttime RALMO measurements on 09 September 2011. In both panels, statistical uncertainties (red curve),  $R_{PRR,d}$  digital coupling constant for PRR (blue curve),  $R_{PRR,a}$  analog coupling constant for PRR (green curve),  $R_{wv,d}$  digital coupling constant for wv/n2 (yellow curve),  $R_{wv,a}$  analog coupling constant for wv/n2 (black curve), air density (cyan curve), Angstrom exponent (purple curve), and seed pressure (brown curve).

The OEM retrieved profiles reach a maximum height (measurement response > 0.9) of 12 km and 7 km for temperature and relative humidity, respectively (Fig. C.8). Unlike nighttime measurements, daytime digital measurements go to background about 10 km and the analog measurements go to background even before that (Fig. C.7). Hence, it is understandable that the retrievals tend to get more information from the *a priori* as it go beyond the height where the signals go to background.

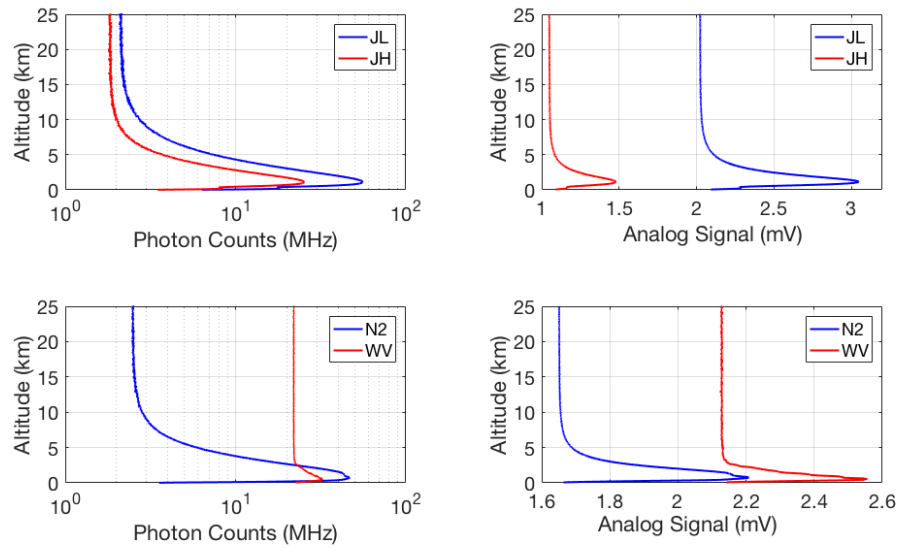


Figure C.7: Count rate for 15 min of RALMO measurements from 1000 UT on 10 September 2011, a cloudy daytime. (a): digital channels (blue curve, JL; red curve, JH). (b): analog channels (blue curve, JL; red curve, JH). (c) digital channels (blue curve, N2; red curve, WV). (d): analog channels (blue curve, N2; red curve, WV)

The OEM retrieved temperature and relative humidity from the RALMO clear daytime measurements are shown in Fig. C.9 left and right panels respectively. The coincident sonde temperature and relative humidity measurements (blue curves) and the *a priori* profiles (green curves) are shown in each panel. Similar to the nighttime case study, the OEM retrieved temperature profile agree with the coincident sonde measured temperature within the statistical uncertainty up to the cutoff height of each retrieval. Relative humidity retrievals also agree with the sonde measurements up to about 6 km within the statistical uncertainty.

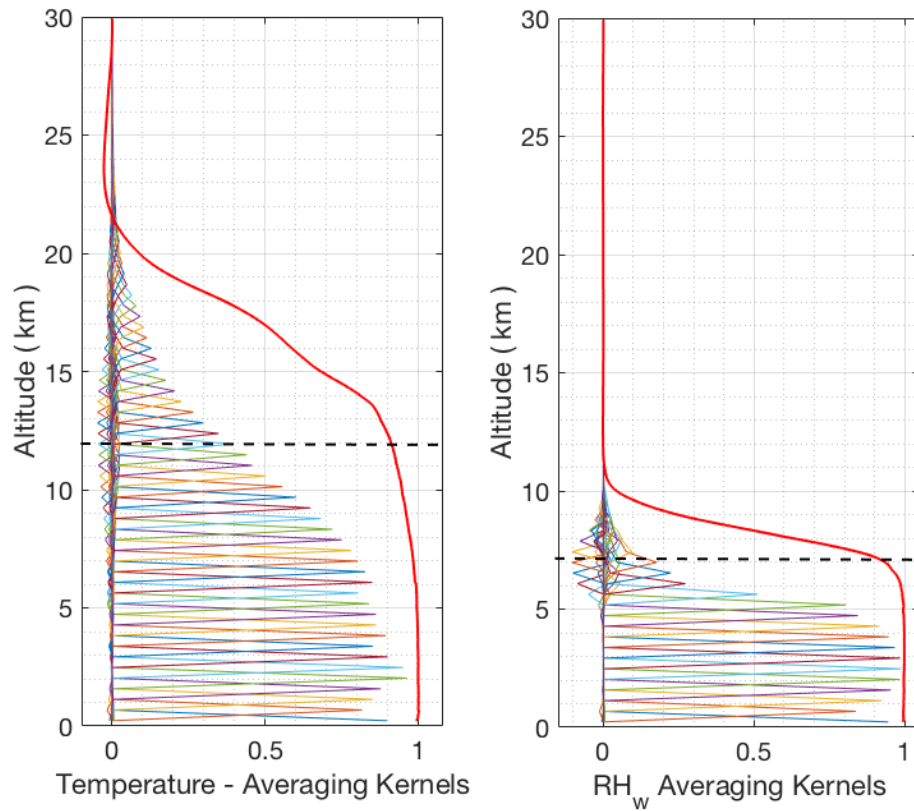


Figure C.8: Averaging kernels for temperature retrievals (left panel) and relative humidity (right panel) from the cloudy daytime RALMO measurements on 10 September 2011. The red curves shows the response functions. The horizontal dashed line shows the height below which the retrieval is due primarily to the measurement and not the *a priori* profiles. For clarity averaging kernels for every fifth altitude bin the retrieval grid are shown.

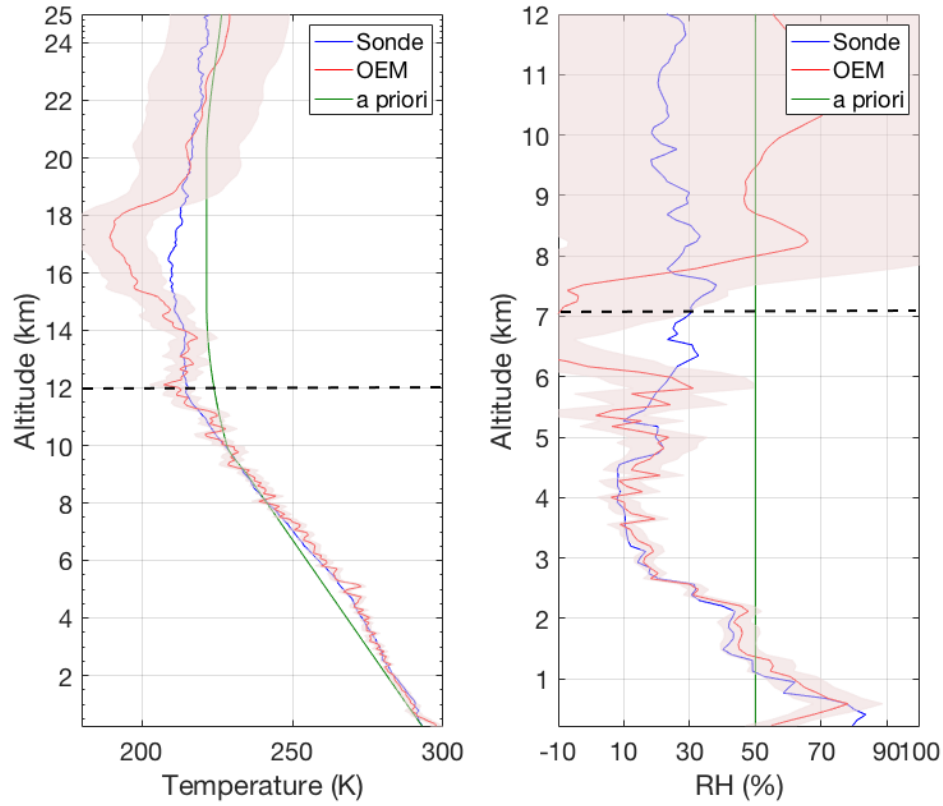


Figure C.9: Left panel: Retrieved temperature profile and the statistical uncertainty (red curve and shaded area) using the OEM from the cloudy daytime RALMO measurements on 10 September 2011. The blue curve is the radiosonde measurement. The sonde was launched at 1000 UT. The green curve is the *a priori* temperature profile used by the OEM. The horizontal dashed line shows the height below which the retrieval is due primarily to the measurement and not the *a priori* temperature profile. Right panel: Retrieved relative humidity profile and the statistical uncertainty (red curve and shaded area). The blue curve is the radiosonde measurement. The green curve is the *a priori* temperature profile used by the OEM. The horizontal dashed line shows the height below which the retrieval is due primarily to the measurement and not the *a priori* relative humidity profile.



### C.3 Temperature, relative humidity and particle extinction time series

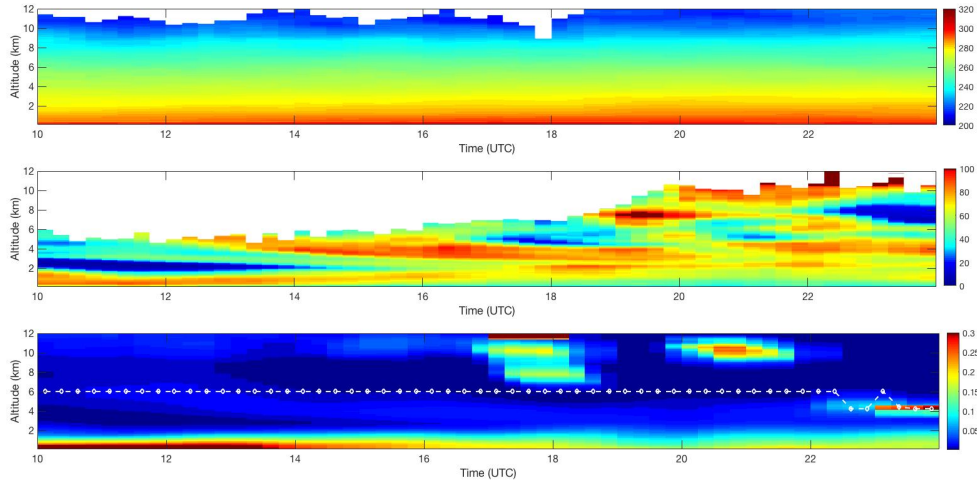


Figure C.10: The OEM retrieved temperature profiles from RALMO measurements on 02 August 2011, from 1000-2359 UT with 15 min temporal and 90 m vertical resolutions.

We present a 14 hour time series of OEM retrieved temperature, relative humidity, and particle extinction measurements from 02 August, 2011( C.10). We have used the nighttime sonde measurements to determine the coupling constants required for the OEM. Estimations of lidar constants are done with respect to the day and nighttime sonde measurements. For the lidar measurements made between 0000-0500 UT, sonde measurements made by the sonde launched on 01 August 2011 are considered. For lidar measurements made between 0500-1700 UT, same day daytime sonde measurements are considered; for lidar measurements made between 1700-2359 UT sonde measurements from the sonde launched on same day nighttime are considered.

The RALMO measurements on 02 August 2011, from 1000-2359 UT are used with 15 min temporal and 30 m vertical resolutions. The retrieval grid resolution is 90 m. In the first plot of the Fig. C.10 we have shown the temperature variation for the 14 hours starting from 10 UT. The white area indicates the retrievals that are above the cutoff height for that retrieval. For comparison purposes, we have only shown temperature retrievals up to 12 km even though

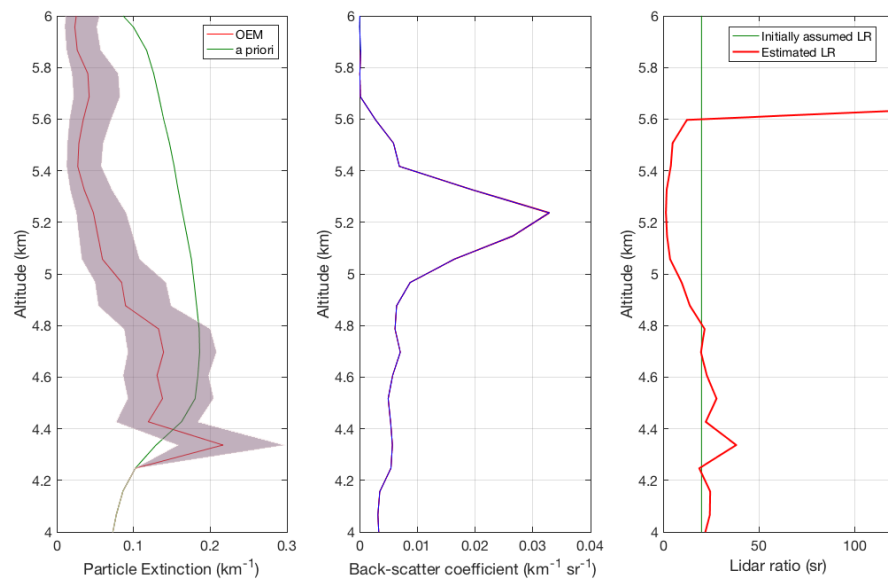


Figure C.11: Left: Retrieved particle extinction (red) and the *a priori* particle extinction used in the OEM (green). Center: Back-scatter coefficient calculated from the RALMO measurements 2245-2300 UT, on 02 August 2011 with of a cloud present about 4 km altitude. Right: Lidar ratio used to determine *a priori* particle extinction (green) and the estimated lidar ratio using the OEM-retrieved particle extinction (red).

the OEM retrieved nighttime temperatures reach up to 25 km. For 02 August 2011, we observe increase in temperature at night than the day time. It is possible that nighttime warming can add more moisture to the air. This can be seen in the second plot in Fig. C.10 where we have shown the relative humidity time series. We observe a greater increase in relative humidity overnight than the day time. Another feature we observe in the relative humidity time series is the dry layer present from 10-14 UT around 2-3 km in height. Beyond 14 UT we, observe a growth of high relative humidity layer around 4-6 km that can be due to the particle extinction layer we observe same time in the third plot shown in the Fig. C.10. Even though we cannot differentiate the thin layer particle extinction is due to a thin cloud or an aerosol layer. We can still conclude the high relative humidity that is present after 14 UT is because of the particle extinction layer present at the same time. The white dash-line in the particle extinction plot in the Fig. C.10 indicates the transition height. Below this transition height we have constrained the particle extinction retrieval. Hence, below the white dash-line the particle extinction retrievals that are shown depend more on the particle extinction profiles calculated using the RALMO elastic and PRR measurements. In our OEM algorithm we only recognize clouds that have aerosol scattering ratios above 2. After 22 UT in the time series we observe a case where we consider to be a cloud that appears below the complete overlap region of the RALMO. Hence, the transition height has changed accordingly.

In Fig. C.11 we present the OEM retrieved particle extinction profile for the measurements from 2245-2300 UT where the low-level cloud is present. The transition height for this retrieval is around 4.3 km. The red curve with the red shaded area in the first plot in Fig. C.11 shows the OEM-retrieved particle extinction and the green curve is the *a priori* particle extinction estimated using the RALMO PRR and elastic measurements. As stated in Section 2.4.3, we have assumed a lidar ratios to estimate an *a priori* extinction. Fig. C.11 shows that the OEM-retrieved particle extinction is much smaller than the *a priori*. The plot on the right in Fig. C.11 shows that the initially assumed lidar ratios, to estimate the *a priori* particle extinction (green curve) is much larger than the the estimated lidar ratio using the OEM-retrieved particle extinction (red curve). The back-scatter coefficient profile estimated from RALMO measurements is shown with the statistical uncertainty in the center plot in Fig. C.11.

## C.4 Summary

We have applied an OEM algorithm to retrieve relative humidity (over water), temperature, particle extinction and multiple other parameters from the Raman lidar measurements in any clear and cloudy conditions. The *a priori* temperature profiles are obtained from the U.S. standard atmospheric model with a 35 K standard deviation. The relative humidity *a priori* profile of a constant value of 50% for all heights was chosen and the 100% standard deviation was considered. The OEM retrievals are not well constrained as the two *a priori* profiles of used for temperature and relative humidity are less accurate and the associated uncertainties are very high. The retrievals are discussed for two different cases that represent day and nighttime clear sky conditions. The time series presented shows the ability of the OEM to retrieve relative humidity and other parameters in different sky conditions. The retrievals from both day and nighttime measurements agree well with the coincident sonde relative humidity and temperature measurements. In our work presented here we consider the retrievals are valid up to a cutoff height where the response function become 0.9. The cutoff heights for nighttime temperature and relative humidity are 25 km and 12 km respectively. Daytime cutoff heights for temperature and relative humidity are 12 km and 7 km, respectively. Thus, 1D Var OEM retrievals of temperature and relative humidity from Raman lidar presented here depend on the Raman lidar measurements up until the cutoff height.

## Appendix D

# Estimated bias and the covariance for relative humidity over ice data in the ERA5 reanalysis

The *a priori*  $RH_i$  used in the ISSR study, is calculated from the ERA5 specific humidity and temperature data. To find the *a priori* uncertainty we followed a similar exercise as given in Chapter 3.

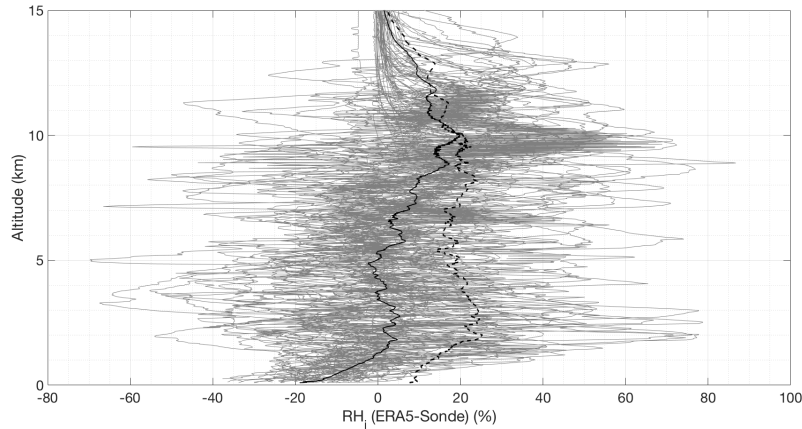


Figure D.1: Bias in the ERA5 relative humidity over ice measurements to the observational measurements from sonde. 56 sonde measurements from year 2008-2011 are shown. Solid black curve is the bias in percentage and black dash curve is the standard deviation.

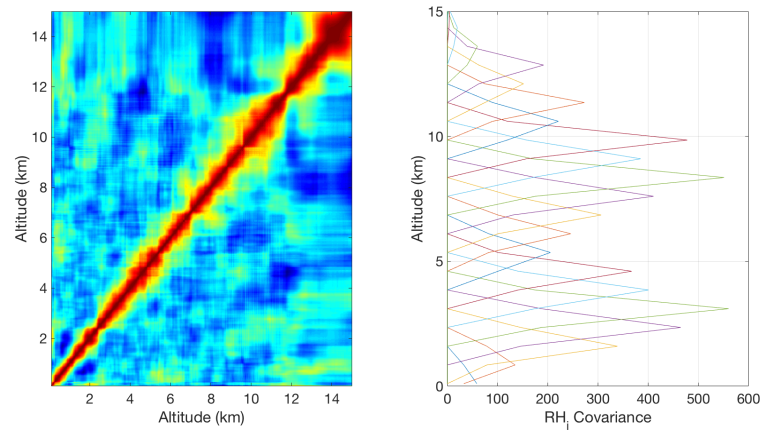


Figure D.2: Left Panel: Correlation of the bias corrected ERA5 relative humidity over ice. Right Panel: Covariance matrix for *a priori* relative humidity over ice estimated assuming a tent function with a correlation length of 750 m and standard deviation of the ERA5 relative humidity over ice bias.

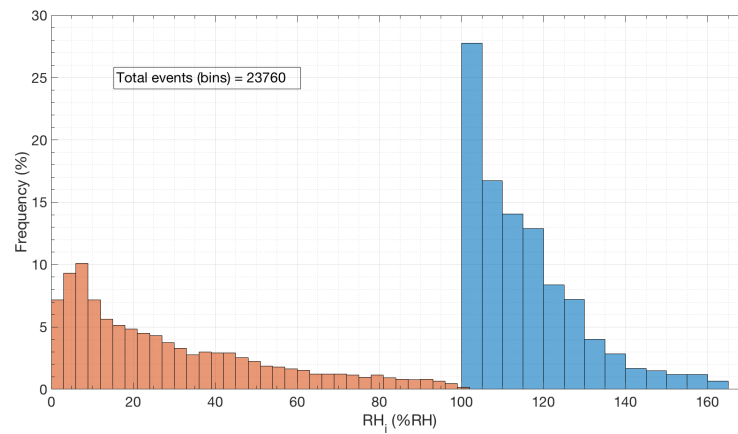


Figure D.3: ERA5-reIce retrieved  $RH_i$  from the RALMO measurements made from June-November, 2011

# Curriculum Vitae

**Name:** Shayamila Mahagammulla Gamage

**Post-Secondary Education and Degrees:** The University of Western Ontario  
Physics  
2013 - 2015 M.Sc.

University of Western Ontario  
London, ON  
2015 - 2019 Ph.D.

University of Sri Jayewardenepura  
Sri Lanka  
2008 - 2012 B.Sc.(Honours Physics,Mathematics and Statistics)

**Honours and Related Work Experience:** Lillian Margaret and Walter David Jackson Scholarship in Physics  
Teaching Assistant  
The University of Western Ontario  
2013 - 2019

## Publications:

1. Gamage, S. Mahagammulla, A. Haeefe, and R. J. Sica. "First application of the optimal estimation method to retrieve temperature from pure rotational raman scatter lidar measurements." *EPJ Web of Conferences*. Vol. 176. EDP Sciences, 2018.
2. Mahagammulla Gamage, S., Sica, R. J., Martucci, G., and Haeefe, A.: Retrieval of Temperature From a Multiple Channel Pure Rotational Raman-Scatter Lidar Using an Optimal Estimation Method, *Atmos. Meas. Tech. Discuss.*
3. S. Mahagammulla Gamage, R.J. Sica, A. Haeefe, G. Martucci, Application of ERA5 Reanalysis to Assimilation of Relative Humidity From Multiple Channel Raman-Scatter Lidar Measurements, 2019, (to be submitted)

4. S. Mahagammulla Gamage, R.J. Sica, A. Haeefe, G. Martucci, Ice supersaturation over Payerne, Switzerland using RALMO Raman lidar measurements, 2019, (to be submitted)

### **Selected Conference talks and posters:**

The European Geosciences Union meeting, 2018, Austria *"First Steps Towards Retrieving Relative Humidity from Lidar Measurements Using The Optimal Estimation Method (OEM): Lower Atmospheric OEM Temperature Retrievals from Pure Rotational Raman Lidar Measurements"*

The International Laser-Radar Conference, 2017, Romania *"First application of the Optimal Estimation Method (OEM) for retrieve temperatures from Pure Rotational Raman (PRR) scatter lidar measurements"*

American Meteorological Society's Ninth Symposium on Lidar Atmospheric Applications, January 2019 *Application of the Optimal Estimation Method (OEM) to Retrieve Temperature and Relative Humidity from Rotational Raman Lidar Back-scatter Measurements*