

## **Raman Lidar Water Vapor Mixing Ratio and Temperature Value-Added Products**

R Newsom

C Sivaraman

November 2018



## **DISCLAIMER**

This report was prepared as an account of work sponsored by the U.S. Government. Neither the United States nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

# **Raman Lidar Water Vapor Mixing Ratio and Temperature Value-Added Products**

R Newsom  
C Sivaraman  
Both at Pacific Northwest National Laboratory

November 2018

Work supported by the U.S. Department of Energy,  
Office of Science, Office of Biological and Environmental Research

## **Acronyms and Abbreviations**

AGL	above ground level
AMF	ARM mobile facility
AOD	aerosol optical depth
ARM	Atmospheric Radiation Measurement
ENA	Eastern North Atlantic
FOV	field of view
IF	interference filter
NFOV	narrow field of view
PMT	photomultiplier tube
QC	quality control
RL	Raman lidar
RMS	root mean square
RR	rotational Raman
SGP	Southern Great Plains
SNR	signal-to-noise ratio
UTC	Coordinated Universal Time
VAP	value-added product
WFOV	wide field of view
WVMR	water vapor mixing ratio

# Contents

Acronyms and Abbreviations .....	iii
1.0 Background.....	1
1.1 ARM Raman Lidar Overview .....	1
2.0 Input Data .....	4
3.0 Algorithm and Methodology .....	7
3.1 MR.....	8
3.1.1 Uncertainty .....	11
3.1.2 Calibration.....	12
3.1.3 FOV Merging .....	13
3.2 TEMP .....	14
3.2.1 Uncertainty .....	15
3.2.2 Calibration and Overlap Correction .....	16
3.3 CAL.....	20
3.3.1 Offline Analysis .....	21
4.0 Output Data and Quality Control.....	22
4.1 MR.....	22
4.2 TEMP .....	23
4.3 CAL.....	24
5.0 Summary.....	26
6.0 Example Plots.....	26
7.0 References .....	29
Appendix A – Molecular Transmission .....	A.1

## Figures

1	The ARM Raman lidar container. ....	1
2	Layout of the ARM Raman lidar optics. ....	3
3	Atmospheric Raman backscatter spectrum due excitation at a wavelength of 355 nm. a) Vibro-rotational lines for N <sub>2</sub> , O <sub>2</sub> and H <sub>2</sub> O (Weitkamp 2005) and b) pure rotational Raman (RR) lines for N <sub>2</sub> and O <sub>2</sub> with the RR1 and RR2 filter transmissions superimposed. ....	3
4	Inputs to WVMR and TEMP. ....	4
5	Examples of the return signals from RR1 (dotted) and RR2 (dashed) for the SGP Raman lidar on 18 October 2009 near 05:00 UTC. ....	15
6	Example showing the result of a linear regression analysis to determine the TEMP calibration coefficients. ....	18
7	Example showing the initial (red dots) and final (black curve) estimates of the temperature overlap function. ....	19
8	Examples showing temporal variability of a) the overlap function, and b) the calibration coefficients. ....	20
9	Examples of baseline calibration profiles (red) for a) the WFOV and b) the NFOV for the period from 1 May through 31 July 2018. ....	22
10	a) WVMR and b) WVMR uncertainty from the SGP RL for 21 August 2017. ....	27
11	Comparisons between the radiosonde (black) and the RL (red) WVMR for four sounding times on 21 August 2017. ....	27
12	Output from TEMP for the SGP RL on 21 August 2017 showing a) unfiltered temperature, b) temperature uncertainty, and c) filtered temperature in which samples with relative uncertainties greater than 5% have been removed. ....	28
13	Comparisons between the radiosonde (black) and the RL (red) temperature measurements for four sounding times on 21 August 2017. ....	29

## Tables

1	ARM Raman lidar specifications. ....	2
2	ARM Raman lidar interference filter specifications for the H <sub>2</sub> O, N <sub>2</sub> and RR channels. ....	2
3	Input variables and datastreams for CAL. ....	5
4	Input variables and datastreams for MR. ....	6
5	Input variables and datastreams for TEMP. ....	7
6	Primary variables in the rlprofmr2news.c0 (MR) datastream. ....	23
7	Primary variables in the rlproftemp2news.c0 (TEMP) datastream. ....	24
8	Primary variables in the sgprlprofmrtempcalC1.c0 (CAL) datastream. ....	24

## 1.0 Background

This report describes the methods used to estimate profiles of water vapor mixing ratio (WVMR) and temperature from the U.S. Department of Energy’s Atmospheric Radiation Measurement (ARM) user facility Raman lidars (RL). This includes a description of the equations and the calibration methods that are used, as well as the output datastreams and variable names. We begin with an overview of the ARM RL systems, with a focus on the detection channels used for the WVMR and temperature measurements.

### 1.1 ARM Raman Lidar Overview

As of 2018, the ARM facility (Mather and Voyles 2013; Stokes and Schwartz 1994) operates three RLs. The systems are deployed at the Southern Great Plains (SGP) observatory in Oklahoma (Sisterson et al. 2016), the Eastern North Atlantic (ENA) observatory in the Azores, and the third ARM Mobile Facility (AMF3) at Oliktok Point in Alaska. All three systems are semi-autonomous, land-based, laser remote-sensing systems that provide height- and time-resolved measurements of water vapor mixing ratio, temperature, aerosol backscatter, extinction, and linear depolarization ratio from about 200m to greater than 10km AGL (Turner et al. 2002).

All three ARM RLs utilize nearly identical designs (Goldsmith et al. 1998; Turner et al. 2016). Each system is housed in an environmentally controlled shipping container, as shown in Figure 1. The laser, which is transmitted through the light baffle on top of the container, operates at a wavelength of 355 nm with a pulse repetition frequency of 30Hz. The pulse width and energy are 5 ns and 300 mJ, respectively. Table 1 lists technical specifications for the transmitter and receiver, and Table 2 lists the specifications for the interference filters used in the measurement of WVMR and temperature.



**Figure 1.** The ARM Raman lidar container. The laser beam exits through the light baffle on the top of the container.

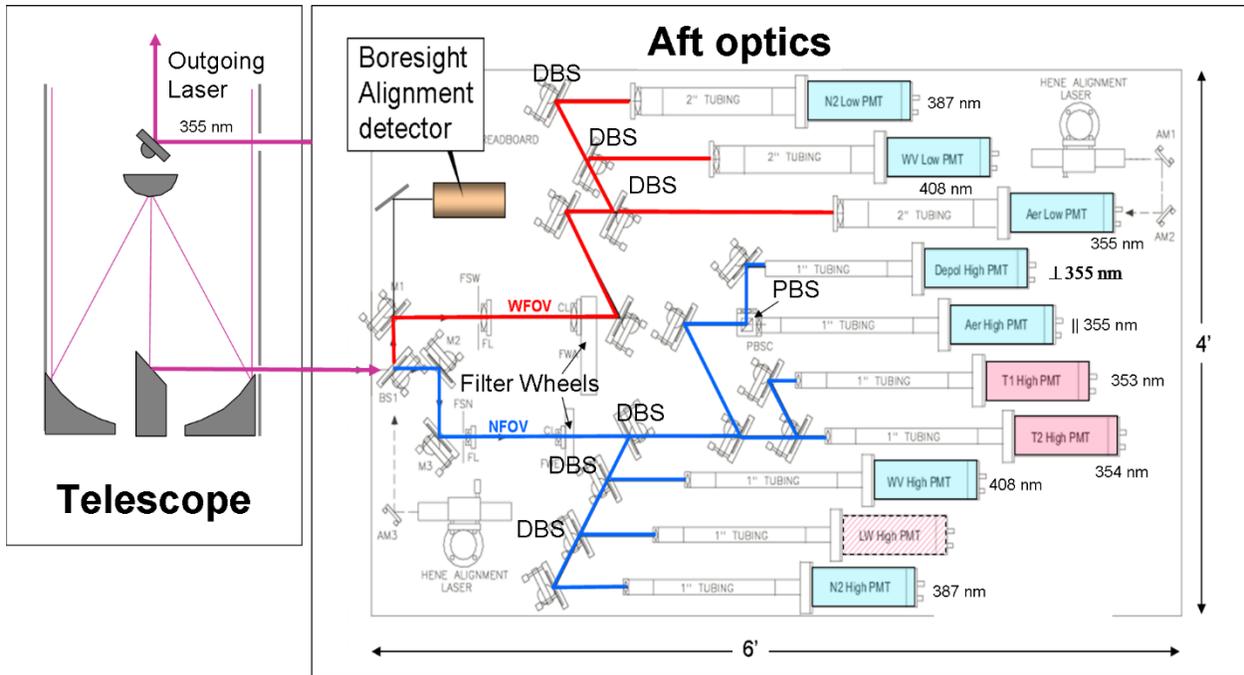
**Table 1.** ARM Raman lidar specifications.

<b>Transmitter</b>	
Laser	Continuum model 9030, frequency-Tripled Nd:YAG
Transmit wavelength	354.7 nm
Pulse energy	~300 mJ
Pulse width	~5 ns
Pulse repetition frequency	30 Hz
<b>Receiver</b>	
Telescope diameter	61 cm
WFOV	2 mrad
NFOV	0.3 mrad
PMTs	Electron Tube 9954B
Data acquisition	Licel transient data recorders
Pulse accumulation time	10s
Range resolution	7.5m

**Table 2.** ARM Raman lidar interference filter specifications for the H<sub>2</sub>O, N<sub>2</sub> and RR channels.

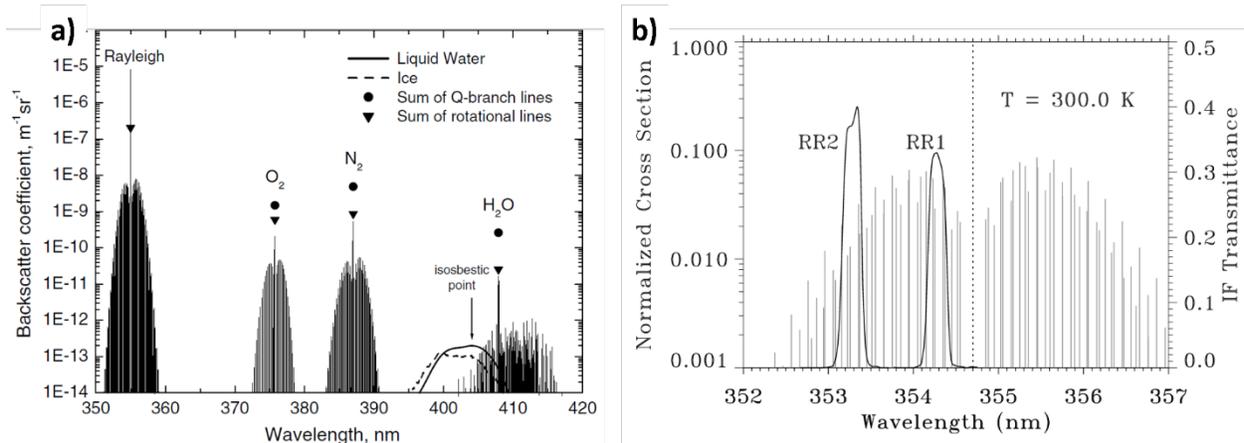
<b>Detection Channel</b>	<b>Center Wavelength (nm)</b>	<b>FWHM (nm)</b>
WFOV & NFOV H <sub>2</sub> O	407.5	0.27
WFOV & NFOV N <sub>2</sub>	386.7	0.31
NFOV RR1	354.27	0.22
NFOV RR2	353.27	0.21

The layout of the RL receiver is illustrated in Figure 2. The outgoing beam enters the telescope and is then redirected into the atmosphere by a turning mirror. Backscattered light from the atmosphere enters the telescope and is directed into the receiver system. This signal is then split between a narrow-field-of-view (NFOV) path (blue) and a wide-field-of-view (WFOV) path (red). The WFOV (2 mrad) path includes three detection channels for H<sub>2</sub>O, N<sub>2</sub> and unpolarized elastic. The NFOV (0.3 mrad) path includes six detection channels. This includes H<sub>2</sub>O, N<sub>2</sub>, co- and cross-polarized elastic, and two pure rotational Raman (RR) channels. All nine detection channels use Electron Tubes 9954B photomultiplier tubes (PMTs). The signals from the PMTs are acquired using transient data recorders from Licel GbR (Berlin, Germany). These data recorders provide simultaneous measurements of both analog photomultiplier voltage and photon counts at a height resolution of 7.5m and a time resolution of 10s.



**Figure 2.** Layout of the ARM Raman lidar optics. The atmospheric return enters the aft optics where it is split between a narrow-field-of-view path (RED) and a wide-field-of-view path (BLUE). The signals from each of the photomultiplier tubes are routed into separate Licel data recorders.

Figure 3a shows the Raman backscatter spectrum of the atmosphere for an incident wavelength of 355 nm at standard temperature and pressure (Weitkamp 2005). The strongest returns are associated with the elastic Rayleigh backscatter. To either side of the Rayleigh line are a number of lines caused by rotational transitions in the vibrational ground states of the  $O_2$  and  $N_2$  molecules. These lines are referred to as pure rotational Raman (RR) lines. The features marked as  $O_2$ ,  $N_2$ , and  $H_2O$  all exhibit prominent vibrational transitions, surrounded by much weaker rotational transitions.



**Figure 3.** Atmospheric Raman backscatter spectrum due to excitation at a wavelength of 355 nm. a) Vibrational lines for  $N_2$ ,  $O_2$  and  $H_2O$  (Weitkamp 2005) and b) pure rotational Raman (RR) lines for  $N_2$  and  $O_2$  with the RR1 and RR2 filter transmissions superimposed. The spectra assume standard temperature and pressure.

The H<sub>2</sub>O detection channel senses the Raman backscatter from atmospheric H<sub>2</sub>O at 407.5 nm, and the N<sub>2</sub> channel senses the Raman backscatter from atmospheric nitrogen at 386.7 nm due to excitation at the laser wavelength. Profiles of WVMR are estimated from the ratio of the H<sub>2</sub>O signal to the N<sub>2</sub> signal, as described in Section 3.1.

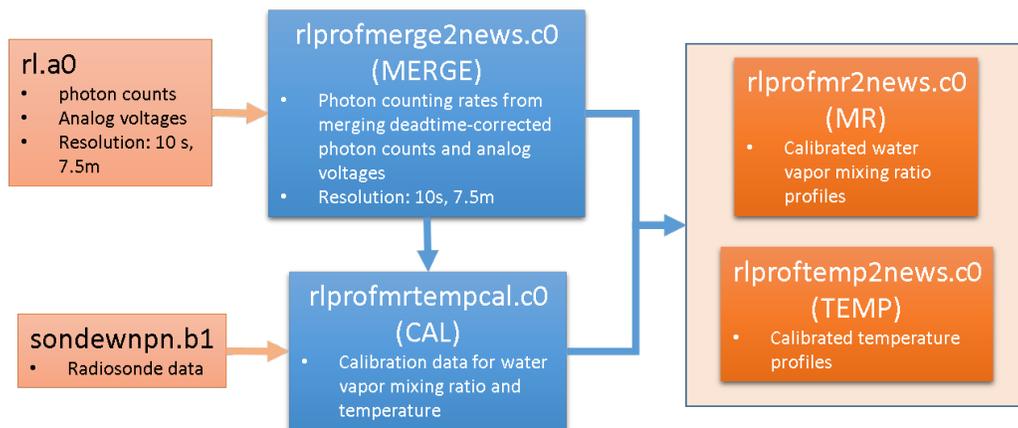
Figure 3b shows a close-up of the pure rotational Raman (RR) lines on either side of the transmit wavelength at 355nm. The positions and strength of the RR lines are temperature sensitive. Also shown are the transmission curves for the two RR channels. These channels use very narrow-bandwidth interference filters (IFs) to measure the energy content in two different portions of the rotational Raman spectrum. Temperature measurement is obtained using the ratio of the two RR signals, which depends nonlinearly on the atmospheric temperature.

The return signals for all channels are affected by incomplete overlap between the outgoing beam and the receiver FOV. The wide FOV (WFOV) achieves complete overlap at an altitude of 800m, while the narrow FOV (NFOV) achieves complete overlap at an altitude of 4km. The overlap functions can change over time due to subtle changes in beam alignment. The WFOV overlap functions generally tend to be more stable and exhibit less temporal variation than the NFOV.

During the daytime, solar radiation causes the random uncertainty in the RL measurements to increase. For the ARM RLs, solar sensitivity increases with wavelength and the size of the FOV. Thus the WFOV H<sub>2</sub>O is the most sensitive to solar and the RR channels are the least sensitive to solar radiation. For this reason, the RL's WVMR measurements exhibit a significant increase in uncertainty during the daytime. Whereas the temperature measurements are far less impacted.

## 2.0 Input Data

The algorithms that compute water vapor mixing ratio and temperature rely on several datastreams for input, as illustrated in Figure 4. Both MR and TEMP read in data from the MERGE and CAL value-added products (VAPs). CAL receives its input from the radiosonde datastream “sondewnpn.b1” and the MERGE VAP. The MERGE VAP receives input from the raw Raman lidar datastream “rl.a0.”



**Figure 4.** Inputs to WVMR and TEMP. The ARM datastream name is listed at the top of each box with the abbreviated name in listed below in parentheses.

The output from MERGE includes photon counting rate data from each of the nine detection channels. These data have the same resolution as the raw RL output (10s and 7.5m). Data from the radiosonde and from the MERGE VAP are used by CAL to generate data for the MR and TEMP calibration procedures. Both MR and TEMP receive input from the MERGE and CAL VAPs. Details concerning MERGE are given by Newsom et al. (2017).

Tables 3, 4, and 5 list variables and corresponding datastreams that are used as inputs to CAL, MR, and TEMP, respectively.

**Table 3.** Input variables and datastreams for CAL.

<b>Datastream</b>	<b>Variable Name</b>	<b>Description</b>	<b>Units</b>
rlprofmerge2news.c0	base_time	seconds since 1970-1-1 0:00:00 0:00	sec
	time_offset	Time offset from base_time	sec
	height_high	Height above ground level	km
	Cbh	Cloud base height	km
	shots_summed	Number of laser pulses accumulated during 10s integration	Unitless
	Filter	Flag indicating which filter is being used	Unitless
	water_counts_high	NFOV H2O photon counting rate	MHz
	water_counts_low	WFOV H2O photon counting rate	MHz
	nitrogen_counts_high	NFOV N2 photon counting rate	MHz
	nitrogen_counts_low	WFOV N2 photon counting rate	MHz
	t1_counts_high	RR1 photon counting rate	MHz
	t2_counts_high	RR2 photon counting rate	MHz
Sondewnpn.b1	base_time	seconds since 1970-1-1 0:00:00 0:00	sec
	time_offset	Time offset from base_time	sec
	Alt	Height above mean sea level	m
	Pres	Pressure	kPa
	Tdry	Dry bulb temperature	C
	Rh	Relative humidity	%

**Table 4.** Input variables and datastreams for MR.

<b>Datastream</b>	<b>Variable Name</b>	<b>Description</b>	<b>Units</b>
rIprofmerge2news.c0	base_time	seconds since 1970-1-1 0:00:00 0:00	sec
	time_offset	Time offset from base_time	sec
	height_high	Height array	km
	shots_summed	Number of laser pulses accumulated during 10s integration	unitless
	filter	Flag indicating which filter is being used	unitless
	cbh	Cloud base height	km
	Elastic_counts_low	Low channel merged elastic counting rate	MHz
	nitrogen_counts_high	NFOV nitrogen channel photon counting rate	MHz
	nitrogen_counts_low	WFOV nitrogen channel photon counting rate	MHz
	water_counts_high	NFOV H2O channel photon counting rate	MHz
	water_counts_low	WFOV H2O channel photon counting rate	MHz
Rlprofmrtempcal.c0	base_time	seconds since 1970-1-1 0:00:00 0:00	Sec
	time_offset	Time offset from base_time	Sec
	height	Height array	Km
	Cbh_calib	Cloud base height	Km
	temp_sonde	Temperature from sonde	K
	pres_sonde	Pressure from sonde	Mb
	mr_sonde	Water vapor mixing ratio from sonde	g kg-1
	h2o_lo_bkg	Photon counting rate for the WFOV H2O channel	MHz
	mr_uncal_hi	NFOV uncalibrated water vapor mixing ratio from the lidar	unitless
	mr_uncal_hi_err	Uncertainty in mr_uncal_hi	Unitless
	mr_uncal_lo	WFOV uncalibrated water vapor mixing ratio from the lidar	Unitless
	mr_uncal_lo_err	Uncertainty in mr_uncal_lo	Unitless

**Table 5.** Input variables and datastreams for TEMP.

Datastream	Variable Name	Description	Units
r profmerge2news.c0	base_time	seconds since 1970-1-1 0:00:00 0:00	sec
	time_offset	Time offset from base_time	sec
	height_high	Height array for the NFOV	km
	height_low	Height array for the WFOV	Km
	cbh	Cloud base height	Km
	shots_summed	Number of laser pulses accumulated during 10s integration	unitless
	filter	Flag indicating which filter is being used	unitless
	water_counts_low	WFOV H2O channel photon counting rate	MHz
	elastic_counts_high	High channel merged elastic counting rate	MHz
	t1_counts_high	RR1 (Low-J) channel photon counting rate	MHz
	t2_counts_high	RR2 (High-J) channel photon counting rate	MHz
R profmrtempcal.c0	base_time	seconds since 1970-1-1 0:00:00 0:00	Sec
	time_offset	Time offset from base_time	Sec
	height	Height array	Km
	cbh	Cloud base height	Km
	temp_sonde	Temperature from sonde	K
	pres_sonde	Pressure from sonde	Mb
	mr_sonde	Water vapor mixing ratio from sonde	g kg-1
	h2o_lo_bkg	Background photon counting rate for the WFOV H2O channel	MHz
	rr_ratio_hi	Ratio of the RR1 (Low-J) photon counting rate to the RR2 (High-J) photon counting rate to the	unitless
	rr_ratio_err	Uncertainty in rr_ratio_hi	Unitless

### 3.0 Algorithm and Methodology

This section describes the equations and methodologies used in MR, TEMP, and CAL. We note that all of these algorithms run independent of each other. The only requirement is that CAL run at least a couple of days ahead of MR and TEMP.

### 3.1 MR

WVMR is defined to be the mass of water vapor contained in a volume divided by the mass of dry in that same volume. This can be expressed as

$$r = \frac{M_{H_2O} N_{H_2O}}{M_{N_2} N_{N_2} + M_{O_2} N_{O_2}}, \quad (3.1.1)$$

where  $M_x$  and  $N_x$  is the molecular weight and number density for species x, respectively. With  $M_{H_2O} = 18.015 \text{ g mole}^{-1}$ ,  $M_{N_2} = 28.013 \text{ g mole}^{-1}$ ,  $M_{O_2} = 31.999 \text{ g mole}^{-1}$ , and assuming the atmosphere consists of roughly 80%  $N_2$  and 20%  $O_2$  by volume, equation (3.1.1) becomes

$$r \cong \frac{1}{2} \left( \frac{N_{H_2O}}{N_{N_2}} \right) \quad (3.1.2)$$

Estimates of the number density ratio  $N_{H_2O} / N_{N_2}$  are obtained from the RL's  $H_2O$  and  $N_2$  channel measurements. The remainder of this section describes how this is done.

We start with the lidar equation for the  $H_2O$  channel. This equation, which describes the range dependence of the return signal power, is given by

$$P_{H_2O}(z) = k_{H_2O} O_{H_2O}(z) \beta(\lambda_o : \lambda_{H_2O}, z) T_{mol}(\lambda_o, z) T_{aer}(\lambda_o, z) T_{mol}(\lambda_{H_2O}, z) T_{aer}(\lambda_{H_2O}, z) + P_{H_2O}^{bkg} \quad (3.1.3)$$

where

$k_{H_2O}$  = System constant that depends on the receiver telescope area and the total receiver optical efficiency

$O_{H_2O}(z)$  = Overlap function

$\lambda_o$  = Transmit wavelength (354.7 nm)

$\lambda_{H_2O}$  =  $H_2O$  channel center wavelength (407.5 nm)

$\beta(\lambda_o : \lambda_{H_2O}, z)$  = Raman backscatter coefficient at  $\lambda_{H_2O}$  due to excitation at  $\lambda_o$ .

$T_{mol}(\lambda_o, z)$  = Molecular transmission at  $\lambda_o$

$T_{aer}(\lambda_o, z)$  = Aerosol transmission at  $\lambda_o$

$$T_{mol}(\lambda_{H_2O}, z) = \text{Molecular transmission at } \lambda_{H_2O}$$

$$T_{aer}(\lambda_{H_2O}, z) = \text{Aerosol transmission at } \lambda_{H_2O}$$

and

$$P_{H_2O}^{bkg} = \text{Background power level (e.g., solar).}$$

Similarly, the range dependence of the return signal for the N<sub>2</sub> channel can be expressed as

$$P_{N_2}(z) = k_{N_2} O_{N_2}(z) \beta(\lambda_o : \lambda_{N_2}, z) T_{mol}(\lambda_o, z) T_{aer}(\lambda_o, z) T_{mol}(\lambda_{N_2}, z) T_{aer}(\lambda_{N_2}, z) + P_{N_2}^{bkg} \quad (3.1.4)$$

Taking the ratio of the background-subtracted H<sub>2</sub>O signal to the background-subtracted N<sub>2</sub> signal yields

$$\frac{P'_{H_2O}}{P'_{N_2}} = \left( \frac{k_{H_2O} O_{H_2O}(z)}{k_{N_2} O_{N_2}(z)} \right) \frac{\beta(\lambda_o : \lambda_{H_2O}, z) T_{mol}(\lambda_{H_2O}, z) T_{aer}(\lambda_{H_2O}, z)}{\beta(\lambda_o : \lambda_{N_2}, z) T_{mol}(\lambda_{N_2}, z) T_{aer}(\lambda_{N_2}, z)} \quad (3.1.5)$$

where

$$P'_{H_2O} = P_{H_2O} - P_{H_2O}^{bkg} \quad (3.1.6)$$

and

$$P'_{N_2} = P_{N_2} - P_{N_2}^{bkg} \quad (3.1.7)$$

are the background-subtracted signals. The backscatter terms in equation (3.1.5) can be expressed as

$$\beta(\lambda_o : \lambda_{H_2O}, z) = N_{H_2O}(z) \left. \frac{\partial \sigma_{H_2O}}{\partial \Omega} \right|_{\pi} \quad (3.1.8)$$

and

$$\beta(\lambda_o : \lambda_{N_2}, z) = N_{N_2}(z) \left. \frac{\partial \sigma_{N_2}}{\partial \Omega} \right|_{\pi}, \quad (3.1.9)$$

where  $N_{H_2O}$  and  $N_{N_2}$  are molecular number densities for atmospheric H<sub>2</sub>O and N<sub>2</sub>, respectively. The

Raman differential backscatter cross-sections for H<sub>2</sub>O and N<sub>2</sub> are denoted  $\left. \frac{\partial \sigma_{H_2O}}{\partial \Omega} \right|_{\pi}$  and  $\left. \frac{\partial \sigma_{N_2}}{\partial \Omega} \right|_{\pi}$ ,

respectively. We note that all of the range dependence in the backscatter coefficients is contained in the

molecular number density. The differential backscatter cross-sections can be regarded as constants, i.e., range independent. Substituting equations (3.1.8) and (3.1.9) into equation (3.1.5) and solving for  $N_{H_2O} / N_{N_2}$  gives

$$\frac{N_{H_2O}}{N_{N_2}} = \left( \frac{k_{N_2} \sigma_{N_2}}{k_{H_2O} \sigma_{H_2O}} \right) \left( \frac{O_{N_2}(z)}{O_{H_2O}(z)} \right) \left( \frac{P'_{H_2O}}{P'_{N_2}} \right) \frac{T_{mol}(\lambda_{N_2}, z) T_{aer}(\lambda_{N_2}, z)}{T_{mol}(\lambda_{H_2O}, z) T_{aer}(\lambda_{H_2O}, z)} \quad (3.1.10)$$

Substituting this expression into equation (3.2) yields

$$r(z) = C(z) \left( \frac{P'_{H_2O}}{P'_{N_2}} \right) \frac{T_{mol}(\lambda_{N_2}, z) T_{aer}(\lambda_{N_2}, z)}{T_{mol}(\lambda_{H_2O}, z) T_{aer}(\lambda_{H_2O}, z)} \quad (3.1.11)$$

where we have defined the calibration profile,  $C(z) \equiv k_{N_2} \sigma_{N_2} O_{N_2}(z) / (2k_{H_2O} \sigma_{H_2O} O_{H_2O}(z))$ . We note that this expression for  $C(z)$  is not used in practice. Instead, the calibration profile is determined empirically.

The molecular transmissions are readily calculated using radiosonde data together with semi-empirical formulas from Bucholtz (1995) for the Rayleigh cross-section (see Appendix A). MR processes one 24-hour period at a time. For any given 24-hour period, the algorithm interpolates the radiosonde temperature, pressure, and humidity profiles to the time and height grid used by the RL's MR algorithm. The interpolated temperature and pressure fields are then used to compute the molecular number density field, which enables calculation of the molecular extinction coefficient and thus the transmission, as described in Appendix A.

Estimating the aerosol transmission terms is more problematic since this requires reasonably accurate estimates of aerosol extinction. In principle, aerosol extinction can be computed from RL's  $N_2$  channel using the technique of Ansmann et al. (1990). But the retrieved extinction profiles are often quite noisy, and strongly affected by overlap. This can introduce undesired artifacts in the WVMR without extremely careful quality control of the retrieved extinction profiles.

Additionally, the ratio of the aerosol transmissions  $T_{aer}(\lambda_{N_2}, z) / T_{aer}(\lambda_{H_2O}, z)$  is typically quite close to unity for low-to-moderate aerosol optical depths (AOD). Whiteman et al. (1992) have shown that for very hazy conditions (AOD = 1) the aerosol transmission ratio decreases by about 5% from the surface to 7km AGL. Under more typical conditions with AOD=0.2, the aerosol transmission ratio decreases by only about 1% from the surface to 7 km AGL. Given this fact as well as the difficulty of accurately estimating aerosol extinction, the effects of aerosol transmission are ignored in the MR algorithm.

Ignoring the effects of aerosol transmission, equation (3.1.11) can be rewritten as

$$r(z) = C(z)r_o(z) \quad (3.1.12)$$

where

$$r_o \equiv \left( \frac{T_{mol}(\lambda_{N_2}, z)}{T_{mol}(\lambda_{H_2O}, z)} \right) \left( \frac{P'_{H_2O}(z)}{P'_{N_2}(z)} \right) \quad (3.1.13)$$

is the uncalibrated WVMR.

### 3.1.1 Uncertainty

Uncertainty estimates are obtained by propagating the effects of shot noise through the calculations. The uncertainty in the calibration profile and in the transmission terms are not included. From equations (3.1.12) and (3.1.13), the uncertainty in WVMR is given by

$$\delta r = C(z)\delta r_o(z) \quad (3.1.1.1)$$

where

$$\delta r_o = \frac{T_{mol}(\lambda_{N_2}, z)}{T_{mol}(\lambda_{H_2O}, z)} \left( \left( \frac{\delta P'_{H_2O}}{P'_{H_2O}} \right)^2 + \left( \frac{\delta P'_{N_2}}{P'_{N_2}} \right)^2 \right)^{1/2} \quad (3.1.1.2)$$

$$\delta P'_{H_2O} = \left( (\delta P_{H_2O})^2 + (\delta P_{H_2O}^{bkg})^2 \right)^{1/2}, \quad (3.1.1.3)$$

and

$$\delta P'_{N_2} = \left( (\delta P_{N_2})^2 + (\delta P_{N_2}^{bkg})^2 \right)^{1/2}. \quad (3.1.1.4)$$

RL signals from the MERGE VAP are given in terms of photon counting rates. The relationship between photon counts and photon counting rate for any detection channel is given by

$$P = \frac{c}{2\Delta r N_{shots}} p \quad (3.1.1.5)$$

Where  $c$  is the speed of light ( $3 \times 10^8 \text{ ms}^{-1}$ ),  $\Delta r$  is the range gate length,  $N_{shots}$  is the number of laser pulses accumulated during the pulse integration period, and  $p$  is the number of photon counts accumulated

within the range cell  $\Delta r$  during the pulse integration period. Photon counts obey Poisson statistics, so the uncertainty in  $p$  is simply  $\sqrt{p}$ . Thus, the uncertainty in the photon counting rate is

$$\begin{aligned}\delta P &= \frac{c}{2\Delta r N_{shots}} \sqrt{p} \\ &= \sqrt{\frac{c}{2\Delta r N_{shots}} P}\end{aligned}\tag{3.1.1.6}$$

The expression above gives the uncertainty in the photon counting rate signal in terms of the photon counting rate, gate size, and the number of laser shots accumulated.

### 3.1.2 Calibration

The RL's WVMR measurements are calibrated using radiosonde data. We note that all ARM RLs are deployed at sites where radiosondes are launched either twice or four times daily, depending on the site. The calibrated WVMR is given by

$$r(z, t) = \alpha(t) C_o(z) r_o(z, t)\tag{3.1.2.1}$$

where  $\alpha(t_s)$  is a time-dependent scale factor, and  $C_o(z)$  is the so-called baseline calibration profile, which is independent of time. Baseline calibration profiles are stored in a configuration file for later retrieval at runtime by MR. They are determined prior to runtime through off-line analysis of the CAL output, as described in Section 3.3.

In equation (3.1.2.1), we model the temporal variability of the calibration profile as  $\alpha(t)C_o(z)$ , i.e., as a time-dependent scaling of a height-dependent function. During development of MR it became apparent that we needed to allow for some temporal variation in the calibration profile in order to reduce biases arising from diurnal variations in system alignment and sensitivity. The simple approach adopted here is robust and works reasonably well under most circumstances.

MR is configured to process one 24-hour period at a time. For a given 24-hour period, MR reads in all available radiosonde profiles and the corresponding profiles of uncalibrated mixing ratio,  $r_o$ , from the CAL VAP. As explained in Section 3.3, CAL computes profiles of  $r_o$  for both the WFOV and NFOV using a 30-minute averaging interval that is centered on the radiosonde launch times.

By equating the RHS of equation (3.1.2.1) to the radiosonde measurements, we can estimate the time-dependent scale factor at the radiosonde launch times using

$$\alpha(t_s) = \text{median} \left( \frac{r_{\text{sonde}}(z, t_s)}{C_o(z)r_o(z, t_s)} \right) \text{ for } z_{\min} \leq z < z_{\max} \text{ and } \delta r_o(z, t_s) / r_o(z, t_s) \leq 0.25 \quad (3.1.2.2)$$

where the median is computed over  $z$  for a given radiosonde launch time,  $t_s$ . The range of  $z$  is restricted to include the highest signal-to-noise-ratio (SNR) portion of the profile. For the NFOV we set  $(z_{\min}, z_{\max}) = (0.5\text{km}, 4.0\text{km})$ , and for the WFOV we set  $(z_{\min}, z_{\max}) = (0.3\text{km}, 2.0\text{km})$ . In addition to the height range restriction, the median is computed using only those measurements for which the relative error is less than 25%, i.e.,  $\delta r_o(z, t_s) / r_o(z, t_s) \leq 0.25$ . Once  $\alpha(t_s)$  is determined, we compute the mean percent difference between the calibrated WVMR profile and the radiosonde profile, i.e.,

$$\Delta(t_s) = \frac{|r_{\text{sonde}}(z, t_s) - \alpha(t_s)C_o(z)r_o(z, t_s)|}{r_{\text{sonde}}(z, t_s)} \quad (3.1.2.3)$$

The procedure above is repeated to obtain  $\alpha(t_s)$  and  $\Delta(t_s)$  for all radiosondes launched during a given 24-hour period. The final step in the calibration procedure involves interpolating the scale factor to the RL's time grid. As an additional quality control measure, the interpolation is performed using only those values of  $\alpha(t_s)$  such that  $\Delta(t_s) \leq 0.2$ .

### 3.1.3 FOV Merging

The calibration procedure described in Section 3.1.2 is carried out separately for the NFOV and the WFOV. Once the calibrated WFOV and NFOV WVMR data are determined, they are combined to form the merged WVMR using the following expressions

$$r_{\text{Merged}}(z, t) = w(z)r_{\text{WFOV}}(z, t) + (w(z) - 1)r_{\text{NFOV}}(z, t) \quad (3.1.3.1)$$

and

$$\delta r_{\text{Merged}} = \left( w^2 \delta r_{\text{WFOV}}^2 + (w - 1)^2 \delta r_{\text{NFOV}}^2 \right)^{1/2} \quad (3.1.3.2)$$

where the weighting function,  $w(z)$ , is defined as

$$w(z) = \begin{cases} 1 & z < z_{\text{WFOV}} \\ 1 - \frac{z - z_{\text{WFOV}}}{z_{\text{NFOV}} - z_{\text{WFOV}}} & z_{\text{WFOV}} \leq z \leq z_{\text{NFOV}} \\ 0 & z > z_{\text{NFOV}} \end{cases} \quad (3.1.3.3)$$

The height range for the merging operation, i.e., the merge range, is given by  $z_{WFOV}$  and  $z_{NFOV}$ . The weighting function varies linearly from 1 to 0 over this range. The merge range is specified in the MR configuration file with typical values of  $z_{WFOV} = 0$  and  $z_{NFOV} = 1.2$  km.

The results of equations (3.1.3.1) and (3.1.3.2) are reported in the MR VAP as “mr\_merged” and “mr\_merged\_err”, respectively (see Table 6 in Section 4.1).

## 3.2 TEMP

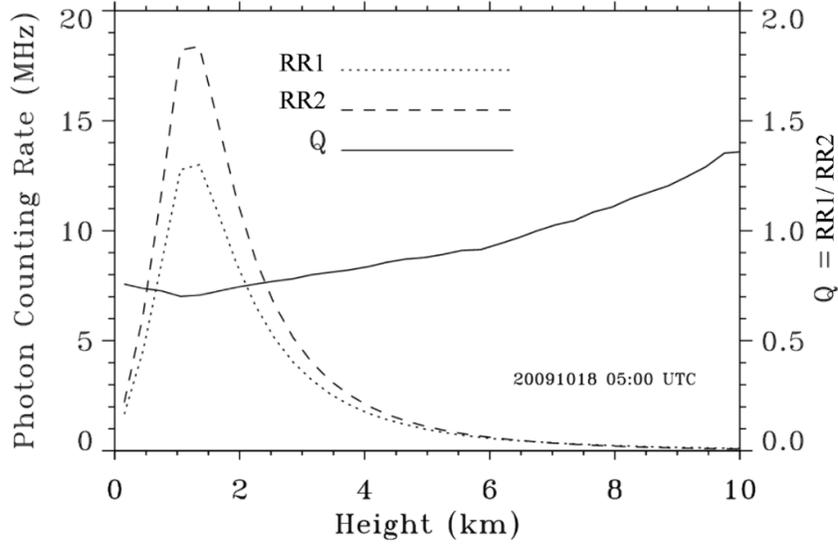
The ratio of the RR signals,  $Q$ , is well approximated by (Behrendt and Reichardt, 2000)

$$Q(z) \equiv \frac{P'_{RR1}(z)}{P'_{RR2}(z)} = O(z) \exp(a + b / T') \quad (3.2.1)$$

where  $P'_{RR1} = P_{RR1} + P_{RR1}^{bkg}$  and  $P'_{RR2} = P_{RR2} + P_{RR2}^{bkg}$  are the background-subtracted RR signals,  $O(z)$  is the ratio of the RR1 overlap function to the RR2 overlap function,  $z$  is the height above the lidar,  $a$  and  $b$  are height-independent calibration coefficients, and  $T' = T / 300K$  is the non-dimensional temperature, where  $T$  is in units of Kelvin.

The ratio of the RR signals largely depends on the temperature of the scattering volume and the effect of incomplete overlap between the outgoing beam and the field of view of the receiver. Differential transmission effects are negligible because the wavelength difference between the two RR channels is small.

Figure 5 shows representative samples of the signals measured by the two RR channels as well as the corresponding signal ratio,  $Q$ . Both signals reach their maximum value at a height of about 1.5 km AGL. In the region where overlap effects are negligible (above 4 km), the altitude variation in the signal ratio is mainly affected by the variation in the atmospheric temperature. In this region the signal ratio increases with altitude as the energy in the high-J rotational states decreases relative to the low-J states due to decreasing temperature.



**Figure 5.** Examples of the return signals from RR1 (dotted) and RR2 (dashed) for the SGP Raman lidar on 18 October 2009 near 05:00 UTC. The solid curve shows the RR signal ratio,  $Q = RR1 / RR2$ .

Solving equation (3.2.1) for the temperature gives

$$T(z) = \frac{(300K)b}{\ln(Q(z)/O(z)) - a} \quad (3.2.2)$$

Unknown quantities in this equation include the calibration coefficients,  $a$  and  $b$ , and the overlap function  $O(z)$ . The goal of the calibration procedure (Section 3.2.2) is to estimate these quantities.

### 3.2.1 Uncertainty

The uncertainty in the RL's temperature estimates are computed by applying standard error analysis to equation (3.2.2). The result is given by

$$\left(\frac{\delta T}{T}\right)^2 = \left(\frac{T}{300K}\right)^2 \left[ \left(\frac{\delta Q}{bQ}\right)^2 + \left(\frac{\delta a}{b}\right)^2 \right] + \left(\frac{\delta b}{b}\right)^2 \quad (3.2.1.1)$$

where  $\delta a$  and  $\delta b$  are the uncertainties in the calibration coefficients (see Section 3.2.2),

$$\left(\frac{\delta Q}{Q}\right)^2 = \left(\frac{\delta P_{RR1}}{P_{RR1}}\right)^2 + \left(\frac{\delta P_{RR2}}{P_{RR2}}\right)^2, \quad (3.2.1.2)$$

$$\delta P'_{RR1} = \left( (\delta P_{RR1})^2 + (\delta P_{RR1}^{bkg})^2 \right)^{1/2}, \quad (3.2.1.3)$$

and

$$\delta P'_{RR2} = \left( (\delta P_{RR2})^2 + (\delta P_{RR2}^{bkg})^2 \right)^{1/2} \quad (3.2.1.4)$$

RL signals from the MERGE VAP are given in terms of photon counting rates. As explained in Section 3.1.1, the uncertainty in the photon counting rate for any channel  $x$  is given by

$$\delta P_x = \sqrt{\frac{c}{2\Delta r N_{shots}} P_x}, \quad (3.2.1.5)$$

where  $c$  is the speed of light ( $3 \times 10^8 \text{ ms}^{-1}$ ),  $\Delta r$  is the range gate length,  $N_{shots}$  is the number of laser shots accumulated during the pulse integration period, and  $P_x$  is the photon counting rate for detection channel  $x$ .

### 3.2.2 Calibration and Overlap Correction

The calibration procedure in TEMP uses radiosonde measurements to estimate the calibration coefficients  $a$  and  $b$ , and the overlap function. We note that all ARM RLs are deployed at sites where radiosondes are launched either 2 or 4 times daily, depending on the site. For a given 24-hour period, the TEMP algorithm reads in all available radiosonde data and the corresponding profiles of the RR signal ratio,  $Q$ , from the CAL VAP. As explained in Section 3.3, CAL computes profiles of  $Q$  using a 30-minute averaging interval that is centered on the radiosonde launch times. Independent estimates of the calibration coefficients and overlap functions are determined for each radiosonde launched during a given 24-hour period. This enables us to represent the temporal variability with a resolution determined by the frequency of radiosonde launches.

#### 3.2.2.1 Calibration Coefficients

Above the minimum height for complete overlap the overlap function is unity. The calibration coefficients can then be determined from equation (3.2.1) for  $z > 4 \text{ km}$  by replacing  $T$  with the radiosonde temperature. In this case, equation (3.2.1) can be recast as

$$y = a + bx \quad \text{for} \quad z \geq 4 \text{ km} \quad (3.2.2.1.1)$$

where

$$y = \ln Q \quad (3.2.2.1.2)$$

and

$$x = 300K / T_{sonde} \quad (3.2.2.1.3)$$

The calibration coefficient are determined by minimizing the following cost function with respect to  $a$  and  $b$ :

$$L = \sum_i \frac{(y_i - a - bx_i)^2}{(\delta y_i)^2} \quad (3.2.2.1.4)$$

where

$$\delta y = \delta Q / Q \quad (3.2.2.1.5)$$

The summation in (3.2.2.1.4) is carried out over samples that occur within a prescribed height range. The lower limit of the height range is given by the minimum height for complete overlap, 4 km, and the upper limit is set to 10 km. In addition to the height range restriction, we also reject samples for which  $\delta y > 0.1$ . Applying the standard least squares methods results in the following 2x2 linear system:

$$\mathbf{A} \begin{pmatrix} a \\ b \end{pmatrix} = \mathbf{f} \quad (3.2.2.1.6)$$

where

$$\mathbf{A} = \begin{pmatrix} \sum_i 1 / \delta y_i^2 & \sum_i x_i / \delta y_i^2 \\ \sum_i x_i / \delta y_i^2 & \sum_i x_i^2 / \delta y_i^2 \end{pmatrix} \quad \text{and} \quad \mathbf{f} = \begin{pmatrix} \sum_i y_i / \delta y_i^2 \\ \sum_i x_i y_i / \delta y_i^2 \end{pmatrix} \quad (3.2.2.1.7)$$

The calibration coefficients are obtained from the solution of equation (3.2.2.1.6). Estimates of the uncertainties in the calibration coefficients are given by the diagonal elements of the inverse of  $\mathbf{A}$ ,

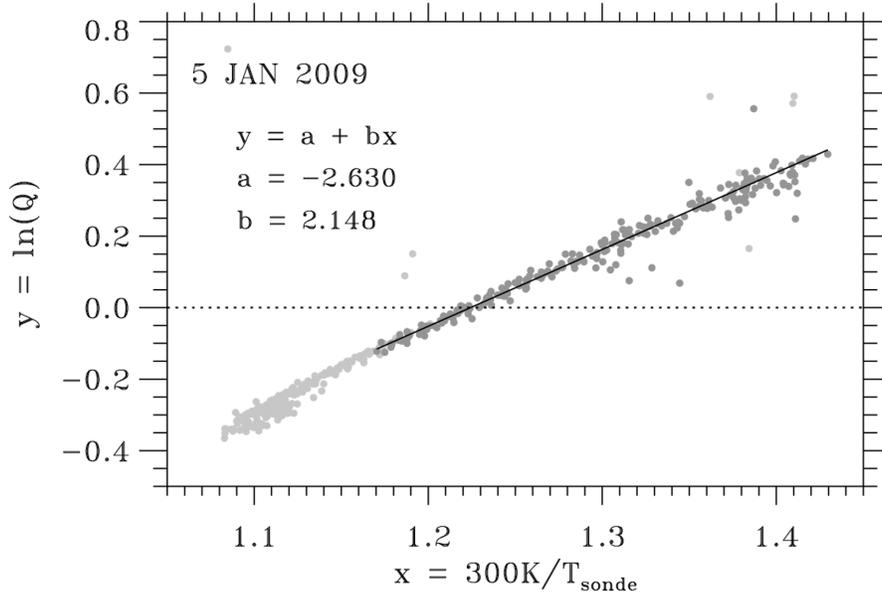
$$\delta a = (\mathbf{A}^{-1})_{11} \quad (3.2.2.1.8)$$

and

$$\delta b = (\mathbf{A}^{-1})_{22} \quad (3.2.2.1.9)$$

The quality of the linear fit is assessed by computing the RMS difference and the correlation between  $y$  and  $a + bx$ . Estimates of  $a$  and  $b$  are deemed to be valid if  $\text{RMS} < 0.1$  and the correlation  $> 0.7$ .

Figure 6 shows an example of a linear fit used to determine the calibration coefficients. We note that the dark gray dots represent valid samples used in the regression analysis. The light gray points represent samples that were not used in the fitting process because they were either outside the prescribed height range or  $\delta y > 0.1$  or both.



**Figure 6.** Example showing the result of a linear regression analysis to determine the TEMP calibration coefficients. This example was taken from the SGP RL on 5 January 2009.

### 3.2.2.2 Overlap Function

If the determination of the calibration coefficients is successful, then an initial estimate of the overlap function is obtained from

$$O_o(z) = Q(z) / \exp(a + bx) \quad \text{for all } z. \quad (3.2.2.2.1)$$

This result can exhibit significant variability with height due to measurement noise. In theory, the overlap function should be a smoothly varying function of height that asymptotically approaches one as  $z$  becomes large. To suppress random variability with height, the initial estimate from equation (3.2.2.2.1) is smoothed using a 5-point boxcar average, i.e.,

$$O_s(z_i) = \frac{1}{5} \sum_{j=i-2}^{i+2} O_o(z_j) \quad (3.2.2.2.2)$$

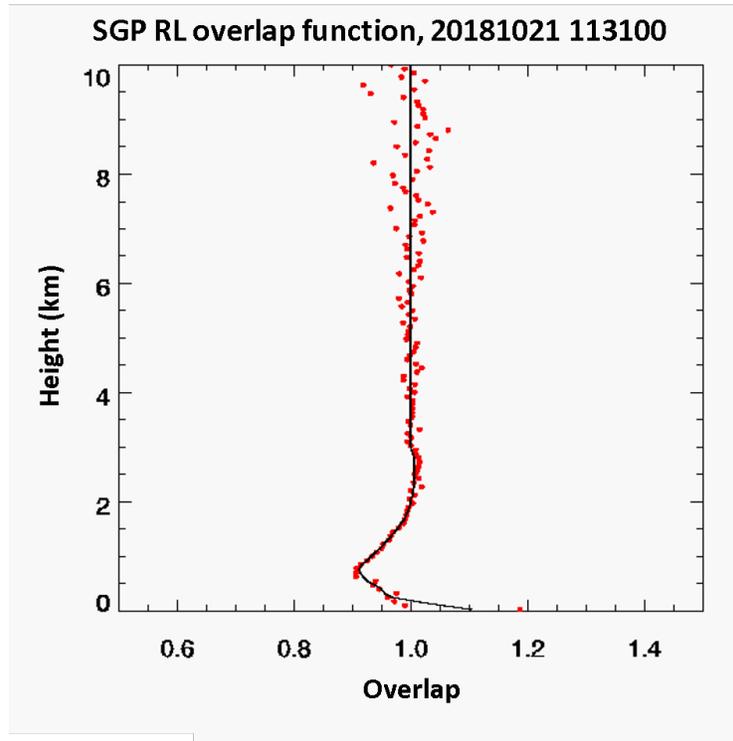
Then to force  $O_s(z)$  to a constant value of one at large  $z$ , we compute

$$O(z) = 1 + g(z)(O_s(z) - 1) \quad (3.2.2.2.3)$$

where  $g(z)$  is a function that varies smoothly from 1 for  $z < 1.5$ km, to 0 for  $z > 4$  km.

Figure 7 shows an example of the initial and final estimates of the overlap function for a single sounding time. This example was obtained from the SGP RL on 21 October 2018 at 11:31 UTC. The red dots

represent the initial estimate as given by equation (3.2.2.2.1), and the black curve represents the final estimates from equation (3.2.2.2.3).

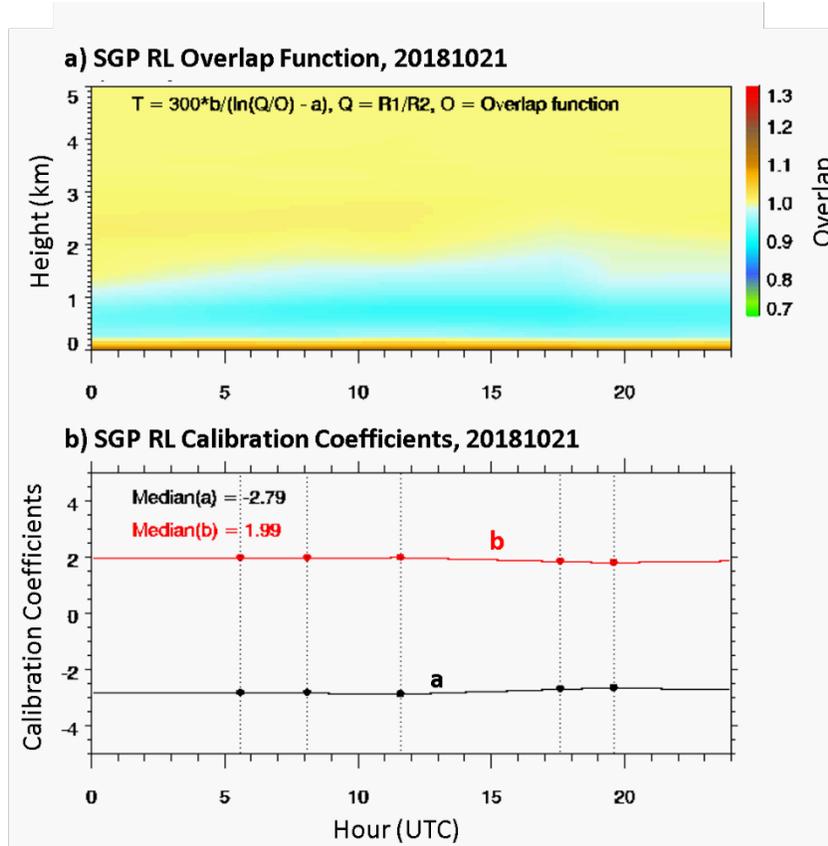


**Figure 7.** Example showing the initial (red dots) and final (black curve) estimates of the temperature overlap function. This example was taken from the SGP RL on 21 October 2018 at 11:31 UTC.

### 3.2.2.3 Time-Dependent Calibration

TEMP processes one 24-hour period at a time. The procedures described in sections 3.2.2.1 and 3.2.2.2 are repeated for all valid radiosonde profiles acquired during that 24-hour period. This gives independent estimates of the calibration coefficients and the overlap functions at each radiosonde launch time. The final step in the calibration procedure involves interpolating these results to the time and height grid of the RL.

Figure 8 shows representative examples of the overlap function and calibration coefficients over the course of one 24-hour processing period (21 October 2018). As the figure suggests, the diurnal variation in both the overlap function and the calibration coefficients is fairly small.



**Figure 8.** Examples showing temporal variability of a) the overlap function, and b) the calibration coefficients. These examples were taken from the SGP RL on 21 October 2018. The vertical dotted lines in panel b indicate the sounding times.

### 3.3 CAL

CAL handles some of the overhead required for calibrating the RL's WVMR and temperature measurements. Its sole function is to compute profiles of the uncalibrated WVMR and the RR signal ratio at the times of the radiosonde launches. That information is then used by MR and TEMP to perform calibration.

Like MR and TEMP, CAL is configured to process one 24-hour period at a time. It first reads in all available radiosonde profiles from the site for a given 24-hour period. It then reads in the MERGE data and computes profiles of the uncalibrated mixing ratio,  $r_o^{WFOV}$ , and  $r_o^{NFOV}$ , and the NFOV rotational Raman signal ratio,  $Q$ , by averaging the photon counting rate data over a 30-minute time period centered on the radiosonde launch time (with a 60-m range bin). This results in 30-minute-averaged profiles of  $r_o^{WFOV}$ ,  $r_o^{NFOV}$  and  $Q$  for each radiosonde launched during a given 24-hour period.

Table 8 lists all the variables that appear in the CAL VAP. The CAL variables that are used by MR are `mr_uncal_hi`, `mr_uncal_hi_err`, `mr_uncal_lo`, and `mr_uncal_lo_err`. The CAL variables that are used by TEMP are `rr_ratio_hi` and `rr_ratio_hi_err`.

We stress that CAL does not actually compute any calibration coefficients. Instead, it merely computes profiles of 30-minute-averaged MERGE data at the radiosonde times, which provides the information necessary for MR and TEMP to do their calibrations. Alternatively, the logic used in CAL could be incorporated into MR and TEMP. However, this would complicate the algorithms somewhat because the calibration parameters are determined using a coarse temporal resolution (30 min) whereas MR and TEMP are typically run with a finer temporal resolution (10 min or less).

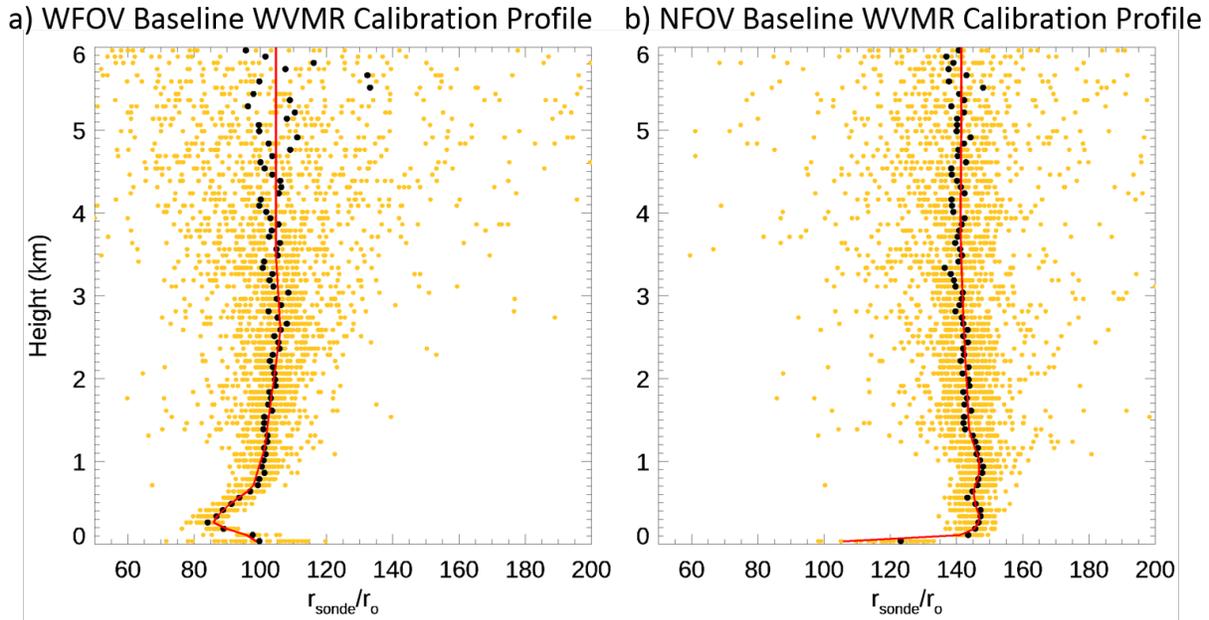
One of the benefits of CAL is its ability to run completely independent of the MR and TEMP. This allows CAL to run days, weeks, or even months ahead of MR and TEMP. This enables offline analysis of the CAL output to estimate the baseline calibration profiles that are used by MR, as described in Section 3.3.1.

### 3.3.1 Offline Analysis

As described in Section 3.1.2, MR uses a prescribed function that we refer to as the baseline calibration profile,  $C_o(z)$ . Baseline calibration profiles are stored in a configuration file and then retrieved by MR at runtime. Each profile in the configuration file represents the mean calibration taken over a period of between roughly two to six months. The idea is that the configuration file captures the long-term variability in the calibration. We note that MR will not run if the configuration file contains no baseline calibration profile for the specified run date.

Baseline calibration profiles are determined through offline analysis of the CAL output. The first step in this analysis is to read in Cal data for the time period to be analyzed, which is typically two to six months in duration. Profiles of  $r_{sonde} / r_o^{WFOV}$  and  $r_{sonde} / r_o^{NFOV}$  are then computed for each radiosonde launched during the analysis period. From these data we compute the median values of  $r_{sonde} / r_o^{WFOV}$  and  $r_{sonde} / r_o^{NFOV}$  as a function of height. Smooth curves are then hand-drawn through the median profiles.

Figure 9 shows examples of baseline calibration profiles for the SGP RL for the period from 1 May through 30 July 2018. The yellow dots represent the values of  $r_{sonde} / r_o$ , and the black dots represent the median value of  $r_{sonde} / r_o$  at a specific height. The baseline calibration profile is the red curve, which is obtained by hand-drawing a smooth curve through the median values.



**Figure 9.** Examples of baseline calibration profiles (red) for a) the WFOV and b) the NFOV for the period from 1 May through 31 July 2018. Black dots represent the median value at each level. The red curve is hand-drawn through the median values.

## 4.0 Output Data and Quality Control

In this section we describe the output of MR and TEMP. We note that the primary output variables generally contain a mix of good- and poor-quality measurements, and it is the end user’s responsibility to perform quality control on these data. Uncertainty estimates are provided for each of the primary variables for just this purpose. In this section we provide listings of the primary variables, and describe how to use the uncertainty estimates to identify and filter out poor-quality measurements.

### 4.1 MR

Table 6 lists the variables included in the MR VAP (rlprofmr2news.c0). The output contains calibrated and uncalibrated WVMR data for the WFOV, NFOV, and combined FOVs. The primary variable of interest to the end user is “mr\_merged” and its corresponding uncertainty “mr\_merged\_err.” Our recommended quality control process consists of filtering out measurements with relative uncertainties exceeding a prescribed maximum value. This is accomplished by rejecting measurements such that

$$\text{mr\_merged\_err}/\text{mr\_merged} > \text{relative\_uncertainty\_threshold},$$

where “relative\_uncertainty\_threshold” is the maximum permissible relative uncertainty. This threshold can be set based on the user’s needs. For the WVMR field we typically set this threshold at 25%, i.e.,  $\text{relative\_uncertainty\_threshold}=0.25$ . This is effective at removing much of the noisiest data, as illustrated in Figure 10.

**Table 6.** Primary variables in the rlprofmr2news.c0 (MR) datastream. All time tags correspond to the middle of the averaging interval. Likewise, the height array gives the height from the ground to the center of the range gate.

Variable	Description	Units
base_time	seconds since 1970-1-1 0:00:00 0:00	sec
time_offset	Time offset from base_time	sec
height	Height array	km
cbh	Cloud base height	km
time_sonde	Flag indicating sonde launch	km
temp_sonde	Temperature from radiosonde	K
pres_sonde	Pressure from radiosonde	mb
mr_sonde	Water vapor mixing ratio from radiosonde	g kg-1
n2_trans_mol	One-way transmission due to N2	unitless
h2o_trans_mol	One-way transmission due to H2O	unitless
mr_lo_cal	Calibration profile for the WFOV water vapor mixing ratio	unitless
mr_lo	Calibrated WFOV water vapor mixing ratio	g kg-1
mr_lo_err	Uncertainty in the calibrated WFOV water vapor mixing ratio	g kg-1
mr_hi_cal	Calibration profile for the NFOV water vapor mixing ratio	unitless
mr_hi	Calibrated NFOV water vapor mixing ratio	g kg-1
mr_hi_err	Uncertainty in the calibrated NFOV water vapor mixing ratio	g kg-1
mr_merged	Calibrated water vapor mixing ratio from merged WFOV and NFOV	g kg-1
mr_merged_err	Uncertainty calibrated water vapor mixing ratio from merged WFOV and NFOV	g kg-1

## 4.2 TEMP

Table 7 lists the variables included in the TEMP VAP (rlproftemp2news.c0). The output contains calibrated and uncalibrated NFOV temperature data. We note that there are no WFOV rotational Raman detection channels in the ARM RLs, so FOV merging is not necessary in TEMP. The primary variable of interest to the end user is “temperature” and its corresponding uncertainty “temperature\_error.” Our recommended quality control process consists of filtering out measurements with relative uncertainties exceeding a prescribed maximum value. This is accomplished by rejecting measurements such that

$$\text{temperature\_err/ temperature} > \text{relative\_uncertainty\_threshold},$$

where “relative\_uncertainty\_threshold” is the maximum permissible relative uncertainty. This threshold can be set based on the user’s needs. For the temperature field we typically set this threshold at 5%, i.e., relative\_uncertainty\_threshold=0.05. This is effective at removing much of the noisiest data, as illustrated in Figure 12.

**Table 7.** Primary variables in the rlproftemp2news.c0 (TEMP) datastream. All time tags correspond to the middle of the averaging interval. Likewise, the height array gives the height from the ground to the center of the range gate.

Variable	Description	units
base_time	seconds since 1970-1-1 0:00:00 0:00	sec
time_offset	Time offset from base_time	sec
height	Height array	km
temperature	Calibrated temperature	K
temperature_error	Uncertainty in calibrated temperature	K
rot_raman_ratio	RR signal ratio (RR1/RR2)	unitless
rot_raman_ratio_error	Uncertainty in the RR1/RR2 ratio	unitless
a_coef	calibration coefficient	unitless
b_coef	calibration coefficient	unitless
olap_function	Overlap function estimate	unitless
temp_sonde	Temperature from radiosonde	K
pres_sonde	Pressure from radiosonde	mb
mr_sonde	water vapor mixing ratio from radiosonde	g kg <sup>-1</sup>
time_sonde	Flag indicating sonde launch	unitless
cbh	cloud base height	km

### 4.3 CAL

CAL computes 30-minute averaged profiles of the uncalibrated WVMR and the rotational Raman signal ratio at the times of the radiosonde launches. That information is then used by MR and TEMP to perform calibration. The CAL output is also used in offline analysis to estimate the baseline calibration profiles for the MR, as described in Section 3.3. Table 8 provides a list of primary variable names in CAL output files.

**Table 8.** Primary variables in the sgprlprofmrtempcalC1.c0 (CAL) datastream. All time tags correspond to the middle of the averaging interval. Likewise, the height array gives the height from the ground to the center of the range gate.

Variable	Description	Units
base_time	seconds since 1970-1-1 0:00:00 0:00	sec
time_offset	Time offset from base_time	sec
height	Height array	km
cbh	cloud base height	km
temp_sonde	Temperature from radiosonde	K
pres_sonde	Pressure from radiosonde	mb

Variable	Description	Units
mr_sonde	water vapor mixing ratio from radiosonde	g kg <sup>-1</sup>
n2_lo	Background-subtracted WFOV N2 channel photon counting rate	MHz
n2_lo_err	Uncertainty in background-subtracted WFOV N2 channel photon counting rate uncertainty	MHz
n2_lo_bkg	WFOV N2 channel background photon counting rate	MHz
n2_lo_bkg_err	Uncertainty in WFOV N2 channel background photon counting rate	MHz
n2_hi	Background-subtracted NFOV N2 channel photon counting rate	MHz
n2_hi_err	Uncertainty in background-subtracted NFOV N2 channel photon counting rate uncertainty	MHz
n2_hi_bkg	NFOV N2 channel background photon counting rate	MHz
n2_hi_bkg_err	Uncertainty in NFOV N2 channel background photon counting rate	MHz
h2o_lo	Background-subtracted WFOV H2O channel photon counting rate	MHz
h2o_lo_err	Uncertainty in background-subtracted WFOV H2O channel photon counting rate uncertainty	MHz
h2o_lo_bkg	WFOV H2O channel background photon counting rate	MHz
h2o_lo_bkg_err	Uncertainty in WFOV H2O channel background photon counting rate	MHz
h2o_hi	Background-subtracted NFOV H2O channel photon counting rate	MHz
h2o_hi_err	Uncertainty in background-subtracted NFOV H2O channel photon counting rate uncertainty	MHz
h2o_hi_bkg	NFOV H2O channel background photon counting rate	MHz
h2o_hi_bkg_err	Uncertainty in NFOV H2O channel background photon counting rate	MHz
t1_hi	Background-subtracted NFOV RR1 channel photon counting rate	MHz
t1_hi_err	Uncertainty in background-subtracted NFOV RR1 channel photon counting rate uncertainty	MHz
t1_hi_bkg	NFOV RR1 channel background photon counting rate	MHz
t1_hi_bkg_err	Uncertainty in NFOV RR1 channel background photon counting rate	MHz
t2_hi	Background-subtracted NFOV RR2 channel photon counting rate	MHz
t2_hi_err	Uncertainty in background-subtracted NFOV RR2 channel photon counting rate uncertainty	MHz
t2_hi_bkg	NFOV RR2 channel background photon counting rate	MHz
t2_hi_bkg_err	Uncertainty in NFOV RR2 channel background photon counting rate	MHz
n2_trans_mol	One-way transmission due to N2	unitless
h2o_trans_mol	One-way transmission due to H2O	unitless
mr_uncal_hi	NFOV uncalibrated WVMR	unitless
mr_uncal_hi_err	Uncertainty in NFOV uncalibrated WVMR	unitless
mr_uncal_lo	WFOV uncalibrated WVMR	unitless
mr_uncal_lo_err	Uncertainty in WFOV uncalibrated WVMR	unitless
rr_ratio_hi	NFOV RR1/RR2	unitless
rr_ratio_hi_err	Uncertainty in NFOV RR1/RR2	unitless

## **5.0 Summary**

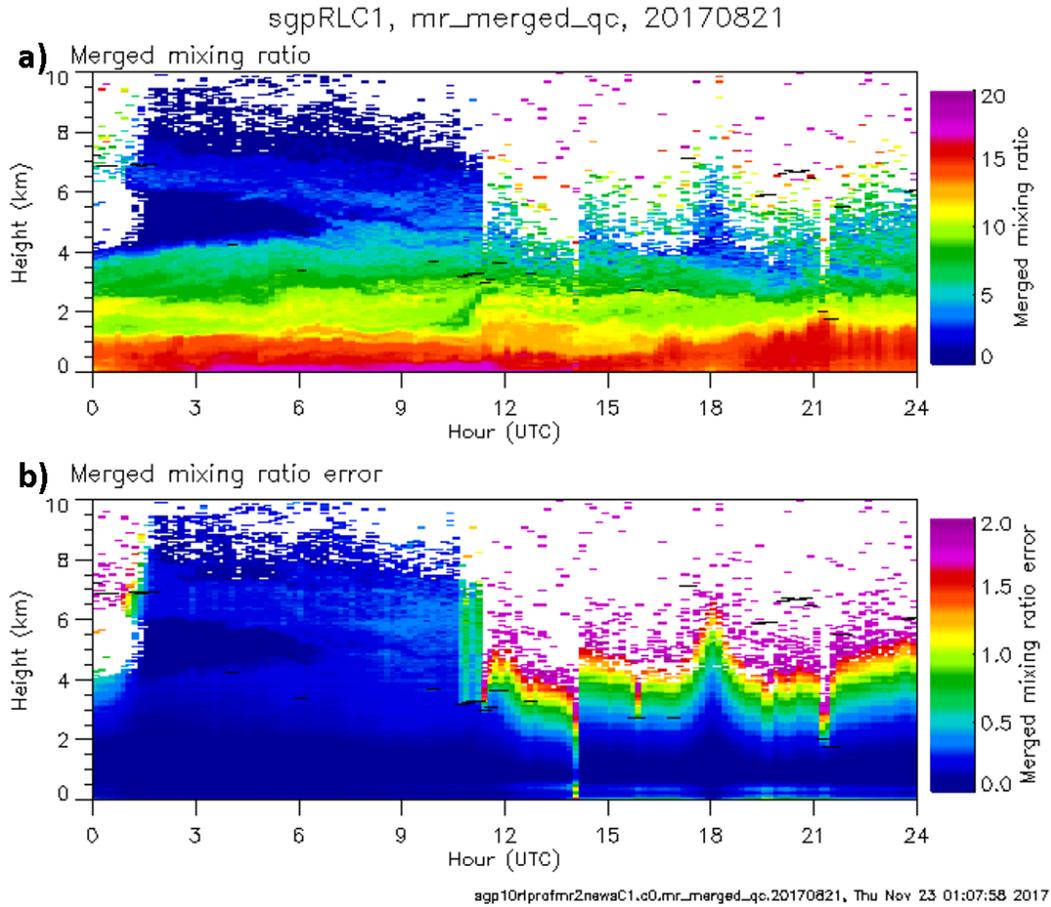
This report documents the methods and equations used to derive profiles of WVMR and temperature from the raw photon counting data acquired by the ARM Raman lidars. Section 1 provided an overview of the Raman lidar system, including the details of the transmitter and the various detection channels. The algorithm flow, input datastreams and variable names are given in Section 2. Section 3 described the theory underlying the measurements and the methods used for calibrating WVMR and temperature. Finally, Section 4 provides a comprehensive listing of the variables included in the output datastreams. This included a description of the recommended quality control procedures that end users should apply when analyzing the WVMR and temperature fields from the ARM RLs.

As stated in Section 3, the MR and TEMP algorithms perform no quality control (QC) on the final output. Instead, it is the responsibility of the end user to perform any necessary QC. Uncertainty estimates for WVMR and temperature are included in the output for just this reason. We recommend filtering out measurements with relative uncertainties exceeding a prescribed threshold value. This threshold can be set based on the user's needs. However, we find that a threshold value of 25% works well for WVMR, and a threshold of 5% works well for the temperature field.

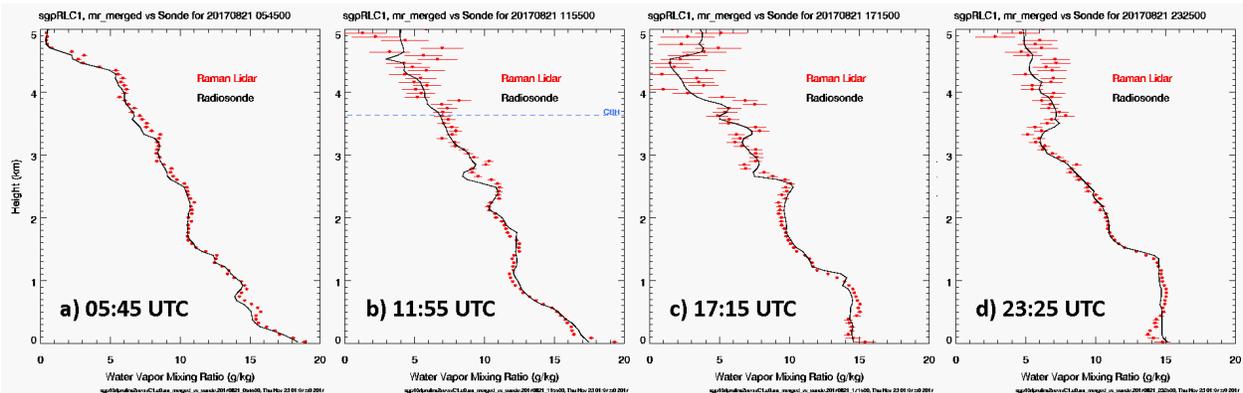
## **6.0 Example Plots**

Figure 10 shows an example of WVMR (and its uncertainty) derived from the SGP RL for 21 August 2017. We note these data have been quality controlled to remove samples with relative uncertainties greater than 25%. The diurnal variation in sensitivity is quite apparent. At night (roughly 1 to 11 UTC) the maximum height for valid measurements is roughly 8km. During the daytime this height is reduced to about 4km due to the influence of solar radiation.

Figure 11 shows comparisons between the RL and the radiosonde measurements of WVMR for the same day shown in Figure 10. We note that the 05:45 UTC sounding occurs shortly before solar midnight. The soundings at 11:55 and 23:25 UTC correspond to morning and evening periods, respectively; and the 17:15 UTC sounding occurs shortly before solar noon.

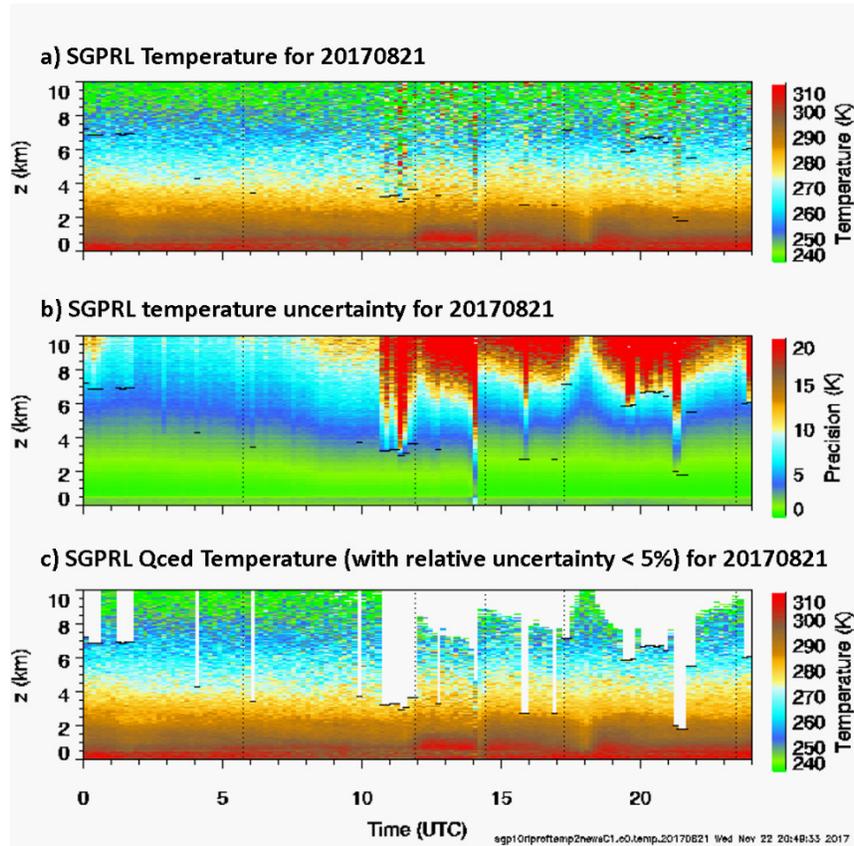


**Figure 10.** a) WVMR and b) WVMR uncertainty from the SGP RL for 21 August 2017. These data have been quality controlled by removing samples with relative uncertainties greater than 25%. Black dashes indicate cloud base heights.



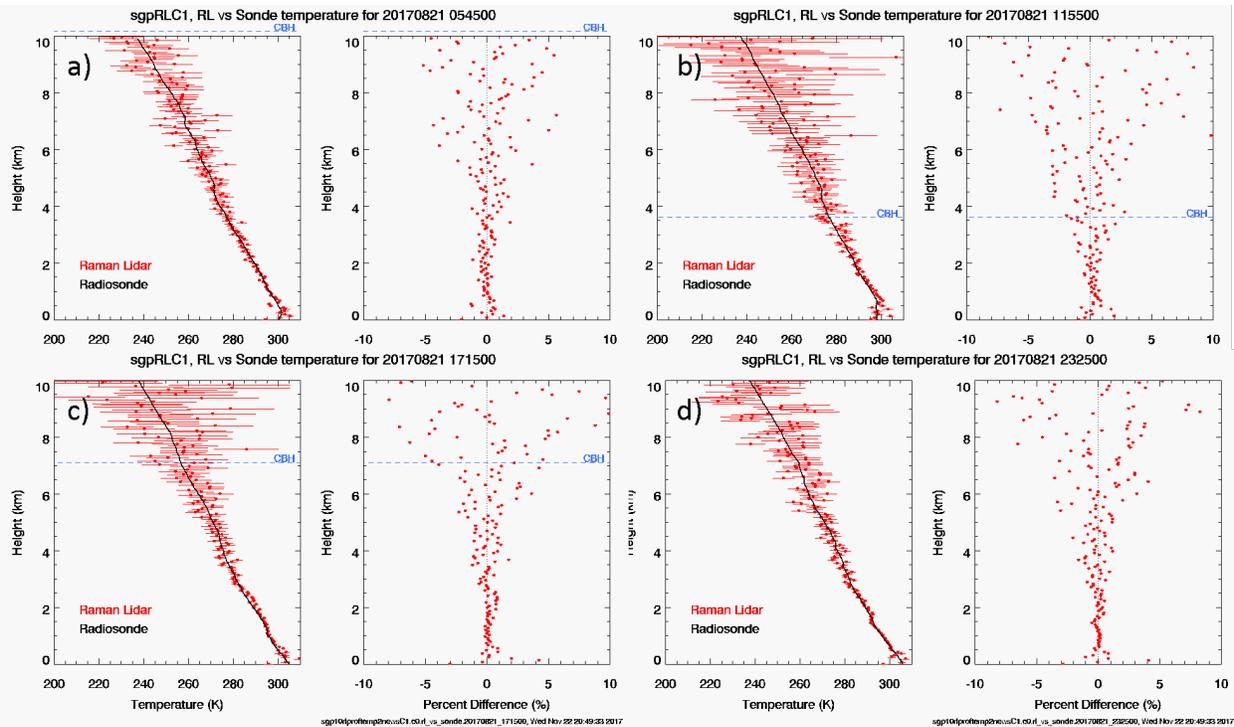
**Figure 11.** Comparisons between the radiosonde (black) and the RL (red) WVMR for four sounding times on 21 August 2017. The sounding times are a) 05:45 UTC, b) 11:55 UTC, c) 17:15 UTC, and d) 23:25 UTC.

Figure 12 shows an example of the temperature field (and its uncertainty) derived from the SGP RL for 21 August 2017. We note these data have been quality controlled to remove samples with relative uncertainties greater than 5%. The diurnal variation in sensitivity is much less apparent for the temperature than for the WVMR. This is because the RR signals are far less sensitive to solar than either the H<sub>2</sub>O or N<sub>2</sub> signals.



**Figure 12.** Output from TEMP for the SGP RL on 21 August 2017 showing a) unfiltered temperature, b) temperature uncertainty, and c) filtered temperature in which samples with relative uncertainties greater than 5% have been removed.

Figure 13 shows comparisons between the RL and the radiosonde measurements of temperature for the same day shown in Figure 12. We note that the 05:45 UTC sounding occurs shortly before solar midnight. The soundings at 11:55 and 23:25 UTC correspond to morning and evening periods, respectively. The 17:15 UTC sounding occurs shortly before solar noon.



**Figure 13.** Comparisons between the radiosonde (black) and the RL (red) temperature measurements for four sounding times on 21 August 2017. The sounding times are a) 05:45 UTC, b) 11:55 UTC, c) 17:15 UTC, and d) 23:25 UTC.

## 7.0 References

- Albert A, M Riebesell, and C Weitkamp. 1990. "Measurement of atmospheric aerosol extinction profiles with a Raman lidar." *Optics Letters* 15(13): 746–748, <https://doi.org/10.1364/OL.15.000746>
- Behrendt, A, and J Reichardt. 2000. "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(9): 1372–1378, <http://doi.org/10.1364/AO.39.001372>
- Bucholtz A. 1995. "Rayleigh-scattering calculations for the terrestrial atmosphere." *Applied Optics* 34(15): 2765–2773, <http://doi.org/10.1364/AO.34.002765>
- Goldsmith JEM, FH Blair, SE Bisson, and DD Turner. 1998. "Turn-key Raman lidar for profiling atmospheric water vapor, clouds and aerosols." *Applied Optics* 37(21): 4979–4990, <http://doi.org/10.1364/AO.37.004979>
- Mather, JH, and JW Voyles. 2013. "The Arm Climate Research Facility: A Review of Structure and Capabilities." *Bulletin of the American Meteorological Society* 94(3): 377–392, <https://doi.org/10.1175/BAMS-D-11-00218.1>
- Miloshevich, LM, H Vömel, DN Whiteman, and T Leblanc. 2009. "Accuracy assessment and correction of Vaisala RS92 radiosonde water vapor measurements." *Journal of Geophysical Research – Atmospheres* 114, D11305, <https://doi.org/10.1029/2008JD011565>

Newsom RK, DD Turner, B Mielke, MF Clayton, R Ferrare, and C Sivaraman. 2009. "Simultaneous analog and photon counting detection for Raman lidar." *Applied Optics* 48(20): 3903–3914, <https://doi.org/10.1364/AO.48.003903>

Newsom, RK, DD Turner, and JEM Goldsmith. 2013. "Long-term evaluation of temperature profiles measured by an operational Raman lidar." *Journal of Atmospheric and Oceanic Technology* 30(8): 1616–1634, <https://doi.org/10.1175/JTECH-D-12-00138.1>

Newsom RK, J Goldsmith, and C Sivaraman. 2017. Raman Lidar MERGE Value-Added Product. [DOE/SC-ARM-TR-189](https://doi.org/10.1175/DOE/SC-ARM-TR-189).

Stokes GM and SE Schwartz, 1994. "The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the Cloud and Radiation Test Bed." *Bulletin of the American Meteorological Society* 75(7): 1201–1221, [https://doi.org/10.1175/1520-0477\(1994\)075<1201:TARMPP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1994)075<1201:TARMPP>2.0.CO;2)

Sisterson DL, RA Pepler, TS Cress, PJ Lamb, and DD Turner. 2016. "The ARM Southern Great Plains (SGP) site" in *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years*, Providence, RI, USA: American Meteorological Society, pp. 6.1-6.14.

Turner DD, RA Ferrare, LA Heilman Brasseur, WF Feltz, and TP Tooman. 2002. "Automated retrievals of water vapor and aerosol profiles from an operational Raman lidar." *Journal of Atmospheric and Oceanic Technology* 19(1): 37–49, [https://doi.org/10.1175/1520-0426\(2002\)019<0037:AROWVA>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<0037:AROWVA>2.0.CO;2)

Turner, DD, JEM Goldsmith, and RA Ferrare. 2016. *Development and applications of the ARM Raman lidar. The Atmospheric Radiation Measurement Program: The First 20 Years, Meteorological Monograph 57*, American Meteorological Society, 18.1-18.15, <https://doi.org/10.1175/AMSMONOGRAPHS-D-15-0026.1>

Weitkamp, Claus. 2005. *Lidar, Range-Resolved Optical Remote Sensing of the Atmosphere. Lidar, Range-Resolved Optical Remote Sensing of the Atmosphere*, Edited by Claus Weitkamp. Berlin, Germany: Springer, pp. 252.

Whiteman, DN, SH Melfi, and RA Ferrare. 1992. "Raman lidar system for the measurement of water vapor and aerosols in the Earth's atmosphere." *Applied Optics* 31(16): 3068–3082, <https://doi.org/10.1364/AO.31.003068>

## Appendix A

### Molecular Transmission

The one-way molecular transmission between the surface and height,  $z$ , is given by

$$T_{mol}(z) = \exp\left(-\int_0^z \gamma_{mol}(z') dz'\right)$$

where  $\gamma_m(z)$  is the molecular extinction coefficient. We assume that transmission loss is due exclusively to scattering, i.e., we ignore absorption. In that case, the molecular extinction coefficient is given by

$$\gamma_{mol}(z) = \sigma N_{mol}(z)$$

where  $N_m(z)$  is the molecular number density and  $\sigma$  is the total Rayleigh scattering cross-section. The molecular number density profile is computed from radiosonde data using the ideal gas law, i.e.,

$$N_{mol}(z) = \frac{p(z)}{kT(z)}$$

where  $k=1.38064852 \times 10^{-23}$ , J K<sup>-1</sup> is the Boltzmann constant, and  $p$  and  $T$  are the radiosonde pressure and temperature, respectively.

The total Rayleigh scattering cross-section is given by (Bucholtz 1995)

$$\sigma = \frac{24\pi^3 (n_{stp}^2 - 1)^2}{\lambda^4 N_{stp}^2 (n_{stp}^2 + 2)^2} \left( \frac{6 + 3\delta}{6 - 7\delta} \right)$$

where  $N_{stp} = 2.54743 \times 10^{25}$  m<sup>-3</sup> is the molecular number density (for all species) at standard temperature and pressure, and  $\delta$  is the molecular depolarization factor. The refractive index of the atmosphere at standard temperature and pressure is well approximated by

$$n_{stp} = 1 + 10^{-8} \left( \frac{5.791}{2.380 \times 10^{-4} - \lambda^{-2}} + \frac{0.169}{5.736 \times 10^{-5} - \lambda^{-2}} \right)$$

where the wavelength,  $\lambda$ , is in units of nanometers.

The depolarization factor,  $\delta$ , accounts for the anisotropy of the air molecules and is wavelength dependent. Bucholtz (1995) presents tabulated values of  $\delta$  as a function of wavelength. The MR algorithm uses this table to obtain interpolated values of  $\delta$  at the wavelength corresponding to the H<sub>2</sub>O and N<sub>2</sub> detection channel. This gives  $\delta=0.0295$  for the H<sub>2</sub>O channel ( $\lambda=407.5$  nm), and  $\delta=0.0296$  for the N<sub>2</sub> channel ( $\lambda=386.7$  nm).



[www.arm.gov](http://www.arm.gov)

U.S. DEPARTMENT OF  
**ENERGY**

---

Office of Science