

# FIRST APPLICATION OF THE OPTIMAL ESTIMATION METHOD TO RETRIEVE TEMPERATURE FROM PURE ROTATIONAL RAMAN SCATTER LIDAR MEASUREMENTS

S. Mahagammulla Gamage<sup>1,\*</sup>, A. Haeefe<sup>2,1</sup> and R.J. Sica<sup>1,2</sup>

<sup>1</sup>*Department of Physics and Astronomy, The University of Western Ontario, Canada*

<sup>2</sup>*Federal Office of Meteorology and Climatology, MeteoSwiss, Payerne, Switzerland*

## ABSTRACT

We present the application of the Optimal Estimation Method (OEM) to retrieve atmospheric temperatures from pure rotational Raman (PRR) backscatter lidar measurements. A forward model (FM) is developed to retrieve temperature and tested using synthetic measurements. The OEM offers many advantages for this analysis, including eliminating the need to determine temperature calibration coefficients.

## 1 INTRODUCTION

The Raman lidar technique uses the pure rotational Raman (PRR) spectra of atmospheric molecules to measure the temperature. The PRR spectrum contains two branches called Stokes and anti-Stokes. Both branches have approximately the same amplitude and they are positioned symmetrically in wavenumber on either side of the excitation line. The traditional Raman lidar temperature retrieval method, introduced by Cooney in 1972, used the ratio of two signals from the lidar's PRR channels [1]. The traditional retrieval algorithm consists of smoothing raw lidar measurements, determining the lidar calibration function coefficients using ancillary measurements such as those from radiosondes, and then calculating temperatures and estimating the absolute and relative uncertainties [2].

The most important capability of the Optimal Estimation Method (OEM) is that it can directly retrieve the lidar constant along with the temperatures from the raw lidar measurements. The OEM also allows us to estimate both random and systematic uncertainties on a profile-by-profile basis.

We use lidar measurements from the Raman Lidar for Meteorological Observations (RALMO) located in Payerne, Switzerland (46°48'N, 6°56'E). RALMO is a fully automated lidar developed by the Federal Institute of Technology in Lausanne, EPFL, and operated by MeteoSwiss. It has two PRR channels that measure the Raman-shifted backscatter arising from rotational energy state transitions in atmospheric nitrogen and oxygen molecules due to excitation at the laser wavelength of 354.7 nm. The High Quantum number channel (JH) in the RALMO measures the backscattered signals from the energy exchange that occur in the high quantum states in both Stokes and anti-Stokes branches while the Low Quantum number channel (JL) measures the signals from the energy exchange occurring in the low quantum states. Detailed description of the RALMO PRR system is given in Dinoev et al.[3]. Implementation of the OEM will be presented in section 2.3. In section 3 retrievals from the synthetic measurements are presented. Conclusions and future work are given in section 4.

## 2 METHODOLOGY

### 2.1 Basics of the OEM

The OEM is a regularized matrix inverse method based on Bayes' theorem. It requires a knowledge of a forward model capable of predicting data if the model parameters are already known. The variables that need to be estimated are called state variables, and they are assembled into a state vector  $\mathbf{x}$ . Similarly, measurements are assembled into a measurement vector  $\mathbf{y}$ . Then the relationship between  $\mathbf{x}$  and  $\mathbf{y}$  is described by the forward model  $\mathbf{F}$ , which con-

tains all the physical and instrumental factors describing the measurements.

The general form of the forward model  $\mathbf{F}$  is,

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \boldsymbol{\varepsilon} \quad (1)$$

where  $\mathbf{b}$  is a parameter vector including all model variables and  $\boldsymbol{\varepsilon}$  is an uncertainty vector including contributions from random uncertainties in the measurements, in the forward model, and in the model parameters. The state vector and the model parameters describe the atmosphere and the instrument measuring it. The state vector contains all the parameters that are to be retrieved. The model parameters contain all other parameters that are needed to fully model the measurement but are not retrieved. In the presence of measurement uncertainties ( $\boldsymbol{\varepsilon} \neq 0$ ), a statistical estimate for the state vector ( $\hat{\mathbf{x}}$ , retrieved parameters) or the optimal solution of  $\mathbf{x}$ , can be obtained using an estimate of the *a priori* state vector ( $\mathbf{x}_a$ ) before the measurements were made.

The cost function associated with the OEM includes the measurement  $\mathbf{y}$ , and its covariance  $\mathbf{S}_y$ , the forward model, the retrieved state, model parameters, the *a priori* and its covariance  $\mathbf{S}_a$ :

$$\begin{aligned} \text{cost} = & [\mathbf{y} - \mathbf{F}(\mathbf{x}, \mathbf{b})]^T \mathbf{S}_y^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x}, \mathbf{b})] \\ & + [\mathbf{x} - \mathbf{x}_a]^T \mathbf{S}_a^{-1} [\mathbf{x} - \mathbf{x}_a]. \end{aligned} \quad (2)$$

If the forward model is non-linear the Marquardt-Levenberg method can be used iteratively to minimize the cost. Further details of the OEM are given in Rodgers [4].

## 2.2 Forward model

We use a forward model based on the Raman lidar equation, with the number density replaced by pressure and temperature via the ideal gas law [5]. The forward model for true counts from both PRR channels is:

$$N_{RR,t}(z) = \frac{C_{RR}^* P(z)}{z^2 kT(z)} \Theta(z) \sigma_{RR}(T, z) \times \Gamma_{atm}^2(z) + B_{RR}^* \quad (3)$$

$C_{RR}$  is the lidar calibration constant that includes transmitted photons, detector efficiency, height resolution and the area of the telescope.  $\Theta(z)$  is the overlap function,  $P(z)$  is pressure,  $T(z)$  is temperature.  $\Gamma_{atm}^2(z)$  is the atmospheric transmission,  $B_{RR}$  is the background noise and  $z$  is the altitude. The backscatter coefficient  $\sigma_{RR}(T, z)$  is given by:

$$\sigma_{RR}(T, z) = \sum_{i=O_2, N_2} \sum_{J_i} \Gamma_{RR}(J_i^*) \eta_i \left( \frac{d\sigma}{d\Omega} \right)_{\Pi}^{RR,i}(J_i^*) \quad (4)$$

$\Gamma_{RR}(J_i)$  is the transmission of the receiver at the wavelength of the RR line  $J_i$ ,  $\eta(i)$  is the relative volume abundance of  $N_2$  and  $O_2$ ,  $\left( \frac{d\sigma}{d\Omega} \right)_{\Pi}^{RR,i}(J_i)$  is the differential back-scatter cross section (depends on temperature), and the corresponding quantum state is  $J$ .

For each channel  $RR$  can be replaced by  $JL$  and  $JH$  and  $C_{RR}^*$ ,  $B_{RR}^*$  and  $J^*$  have different values. For a channel with a height-dependent background,  $B_{RR}$  would then be a function of altitude, and the additional coefficients would be included in the forward model.

The two lidar calibration constants  $C_{JL}$  and  $C_{JH}$  can be related by:

$$R = \frac{C_{JH}}{C_{JL}} \quad (5)$$

The ratio,  $R$ , is constant over long time periods where  $C_{JL}$  and  $C_{JH}$  are stable over hours only. It is also independent of temperature and allows us to retrieve only one lidar calibration constant.  $R$  can be derived using:

$$R = \left( \frac{N_{t,JH} - B_{JH}}{N_{t,JL} - B_{JL}} \right) \bigg/ \left( \frac{\sigma_{JH}(T, z)}{\sigma_{JL}(T, z)} \right) \quad (6)$$

Based on the assumption of non-paralyzable detectors, the true count profile can be transformed into the observed height profile using the dead-time correction in Eq.(7):

$$N_{true} = \frac{N_{obs}}{1 - N_{obs}\gamma} \quad (7)$$

Currently we retrieve temperature, background noise (that includes photomultiplier shot noise,

and sky background) and the lidar calibration constant  $C_{JL}$ . As the measurements and retrieval grid spacings are relatively small compared to the height of the atmosphere, we use linear interpolation between the two grids.

### 2.3 Implementation of the OEM

For this study we use the Qpack software package in Matlab implemented by Eriksson et al.[6]. Qpack is a part of the ARTS model used by the radiometer and microwave community. We use the Marquardt–Levenberg method in Qpack to do non-linear retrievals. This also allows us to retrieve parameters using an *a priori* considerably far from the true solution [7]. For the preliminary testing of the forward model we use the US standard atmospheric model as *a priori* knowledge. As discussed in Eriksson et al. [6], when there is some correlation between the various quantities the respective covariance matrix should have non zero off-diagonal elements. The correlation is provided by using correlation functions with a correlation length appropriate to the physical situation [7]. We use a tent function with a covariance length of 1 km for the temperature covariance.

We use measurements from both JL and JH channels in the measurement vector and the covariance matrix  $\mathbf{S}_y$  is a diagonal matrix. With the use of the Qpack OEM non-linear solver the optimal solution is obtained minimizing the cost.

## 3 RESULTS

We present the initial test results of the OEM using the synthetic measurements. The US standard atmospheric model temperature and pressure profiles were used in the forward model to generate the synthetic measurements shown in Fig. 1. Poisson noise and a constant background were added to the synthetic measurements to make those similar to the 4 hours of integrated nighttime RALMO measurements obtained on 28<sup>th</sup> June 2011.

The averaging kernel in Fig. 2, shows the sensitivity of the retrieved temperatures to the true

temperatures. We can retrieve real temperatures with 1 km vertical resolution up to about 15 km. However, the sensitivity decreases with the increasing height. The sum of the averaging kernel shown in red should equal to unity where the retrievals are accurate.

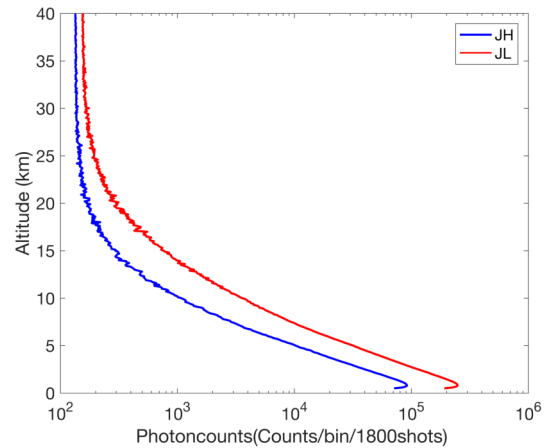


Figure 1: Synthetic measurements of JH and JL channels, generated using the forward model in Eq.(3).

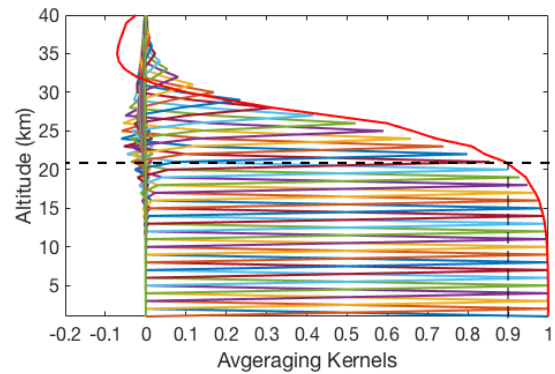


Figure 2: The averaging kernel for temperature. The red curve shows the sum of the kernels.

Fig. 3, shows the percent difference between the synthetic measurement and the retrievals. The difference between the measurements and model is within the Poisson photoncount uncertainty. The lidar calibration constant was retrieved with a percent difference less than 0.5% and the percent difference for the retrieved background is less than 3%.

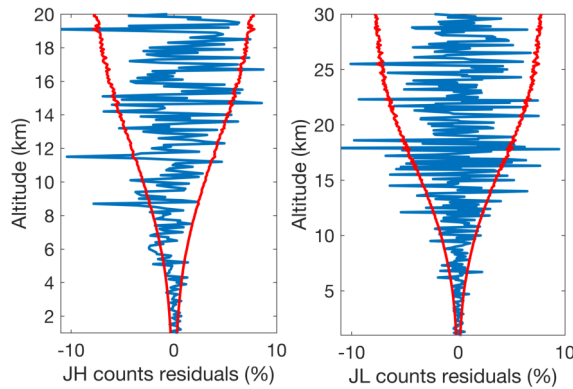


Figure 3: Residuals between the forward model and the measurements (blue curves). The red curves show the variance of the measurements.

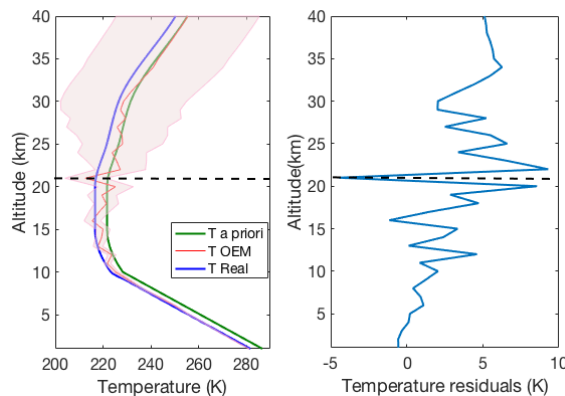


Figure 4: Figure on left shows the retrieved temperatures (red curve). The blue and green curves show the true and the *a priori* temperatures respectively. Figure on the right shows the temperature residuals between retrieved and true states. Dashed line shows the cutoff height of the retrievals

Fig. 4 (left), shows a comparison of the retrieved temperatures to the *a priori* and true temperatures. In Fig. 4 (right), the difference between the retrieved and the true temperature profiles are shown. The uncertainty and temperature residuals increase with height. The retrieved and the true temperatures up to 15 km agree within 3 K. As was seen in Fig. 3, the sensitivity of the retrievals decrease after 15 km and reach zero eventually. The retrieved temperatures start to fall back to the *a priori* state from 21 km height. Therefore, the cutoff height of the retrieval can be set around 21 km where the averaging kernels is 0.9.

These results were produced by the OEM solver after completing 5 iterations. The initial and the final costs of the retrievals are 5.2566 and 0.6037 respectively.

#### 4 CONCLUSIONS AND OUTLOOK

The OEM has been shown to be able to successfully retrieve temperatures from the synthetic, but realistic, pure rotational Raman-scattering-lidar measurements. The temperature retrievals agree within 3 K up to 15 km with the true profile. The retrieval uncertainty and vertical resolution increase as shown by the averaging kernels. Currently background noise and the lidar calibration constant for the JL channel can also be retrieved using the OEM with less than 3% uncertainty. We are currently in the process of extending the retrieval to include dead time, transmission and other parameters in addition to temperature, with the goal of then applying the retrieval to RALMO measurements and validating them.

**ACKNOWLEDGEMENTS:** We thank the Natural Sciences and Engineering Research Council (NSERC), Canadian Space Agency (CSA) and MeteoSwiss for supporting this project.

#### References

- [1] Cooney, J., 1972. Measurement of atmospheric temperature profiles by Raman backscatter. *Journal of applied meteorology*, 11(1), pp.108-112.
- [2] Zuev, Vladimir V., et al. Tropospheric temperature retrievals using nonlinear calibration functions in the pure rotational Raman lidar technique. *Atmos. Meas. Tech. Discuss.*, doi 10.
- [3] Dinoev, T. S., et al. Monitoring of Eyjafjallajökull ash layer evolution over payerne Switzerland with a Raman lidar. *Proceedings of the TECO (2010)*.
- [4] Rodgers, C.D., 2000. *Inverse methods for atmospheric sounding: theory and practice (Vol. 2)*. World scientific.
- [5] Behrendt A. Temperature measurements with lidar. *InLidar 2005 (pp. 273-305)*. Springer New York.
- [6] Eriksson, P., Jiménez, C. and Buehler, S.A., 2005. Qpack, a general tool for instrument simulation and retrieval work. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 91(1), pp.47-64.
- [7] Sica, R.J. and Haefele, A., 2015. Retrieval of temperature from a multiple-channel Rayleigh-scatter lidar using an optimal estimation method. *Applied optics*, 54(8), pp.1872-1889.