Project DEEPWAVE
DLR Rayleigh Lidar Data Report

Version 3
21 March 2016

Prepared by:
Bernd Kaifler (bernd.kaifler@dlr.de)
German Aerospace Center, Institute of Atmospheric Physics
Münchener Strasse 20, 82234 Wessling, Germany
phone: +49 8153 283519

Lidar team:
Bernd Kaifler, Natalie Kaifler, Benedikt Ehard, Markus Rapp

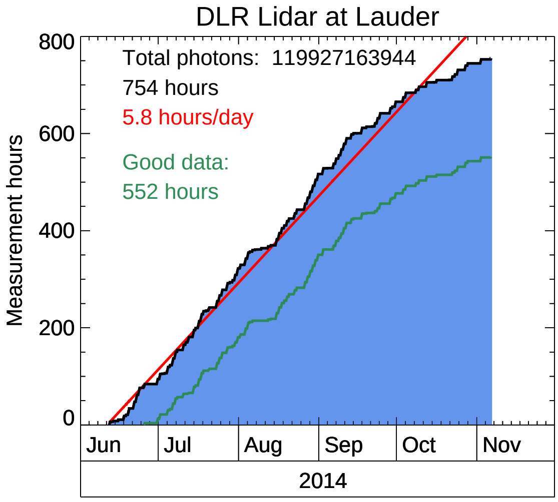

Figure 1: Lidar observations carried out with the DLR Rayleigh-/Raman lidar at Lauder, New Zealand

1 Data Set Overview

The DLR Rayleigh-/Raman lidar measures atmospheric backscatter profiles from 22 to approximately 90 km altitude. Temperature profiles are retrieved assuming only molecular scattering and hydrostatic equilibrium. Stratospheric aerosols may be present up to 32 km altitude, causing errors in the temperature retrieval below. These errors **are not included** in the standard error provided with the temperature data set. Use temperature profiles below 32 km with caution and contact the PI!

The lidar was set up at the atmospheric research station of the National Institute of Water and Atmospheric Research in Lauder (45.04°S, 169.68°E), New Zealand, in June 2014. Routine observations were carried out between 25 June and 3 November 2014.

In the period 19-25 June the lidar data acquisition system malfunctioned and retrieved temperature profiles show an artificial non-linear gradient above 60 km altitude. For this reason, data are not included in the initial submission to the data archive, but are available on request.

The lidar was run in total for 754 hours, with 552 hours of good temperature data (see Figure 1).

Please contact the PI before using the data for publications in any form.

2 Instrument Description

The DLR Rayleigh-/Raman lidar, also known as TELMA (Temperature Lidar for Middle Atmosphere Research), is a mobile lidar system installed in an 8 ft container. The laser is a diode-pumped Nd:YAG laser from Innolas Laser GmbH (SpitLight DPSS 250) generating 12 W at 532 nm and 100 Hz pulse repetition rate. The beam is expanded to 12 mm, resulting in 180 µrad beam divergence. A piezo actuated mirror is used to control beam pointing. The telescope comprises a 630 mm diameter f/2.45 parabolic mirror with a spot size of 70 µm and a superstructure made of carbon fibers and aluminum which supports an optical fiber in the focal point. The effective field of view is approximately 240 µrad. In the receiver the beam is collimated after exiting the fiber. A dichroic mirror separates the 608 nm light originating from vibrational Raman scattering. The remaining beam is split into two beams with a splitting ratio of 8:92 and directed to low noise detectors operated in single photon counting mode. Two 0.8 nm FWHM interference filters mounted in front of the detectors block background light. The far channel detector (92% signal) is a SPCM-AQRH-16 avalanche photo diode from EXCELITAS, the near channel detector is a H7421 high sensitivity photo multiplier tube from HAMAMATSU. The 608 nm beam passes through two 3 nm FWHM interference filters. The detector used for this beam is also a H7421. A mechanical chopper located immediately behind the fiber end blocks the intense light originating from the lower atmosphere and prevents saturation of the detectors. In addition, the avalanche photo diode of the far channel is gated and opens at 41 km.

Photon count signals from the three detectors are recorded with 2 ns resolution using a P7888 Multiscaler from FAST Comtec GmbH. This allows for flexible vertical and temporal binning during data analysis. Taking this to the extreme, backscatter profiles for each individual laser pulse can be retrieved, if desired.

3 Data Collection and Processing

High resolution temperature profiles are retrieved from raw photon count profiles using an iterative approach. First, the nightly mean photon count profile is formed by integrating all laser pulses in time with vertical resolution of 100 m. Following the standard procedure in lidar data analysis, the background is removed and the profile scaled to account for geometric factors. The resulting relative density profile is then integrated from top to bottom assuming hydrostatic equilibrium. We use SABER temperature data to seed the temperature integration between 95 and 107 km altitude, depending on the signal-to-noise ratio of the lidar density profile. The calculated temperature profile becomes largely independent of the seed temperature about 1-2 density scale heights below the seed altitude. Approximately one density scale height below the seed altitude, the uncertainty of calculated temperatures becomes dominated by photon noise rather than errors in the seed temperature. Integration of the near channel is seeded at approximately 52 km (the specific altitude depends on the signal-to-noise ratio) using temperature data from the far channel. Temperature profiles of both channels are merged into a single profile by computing a weighted average in the altitude region 44-49 km, where weights w are set according to the function $w(z) = 1 - \cos(\pi/(2\mathrm{n}) \cdot i)$, where n is the number of bins in the overlapping region, and i is the bin index. This function guarantees a smooth transition from the one profile to the other. Differences between the two profiles in the overlapping region are usually less than 1 K. See also section 3.2.2.

We estimate the uncertainty of retrieved temperature profiles caused by photon noise and errors in seed temperature based on Monte Carlo simulations. As photon counting is a Poisson process, the uncertainty of the photon counts within a given altitude bin is $\sqrt{N}$, where N is the number of detected photons in this bin. In order to estimate the noise in the temperature data, we create for each measured backscatter profile 500 simulated profiles by adding random noise scaled with $\sqrt{N}$ to the measured profile. For each simulated profile we compute temperature profiles as outlined above. In this process the seed temperature is also varied by adding random noise to SABER temperatures, assuming that the SABER temperature is within ± 15 K of the true value. The final temperature profile is then computed as the mean of all 500 simulated profiles, and the uncertainty is estimated from the standard deviation. Note that the hereby computed uncertainty estimates include only noise contributions from photon noise and errors in seed temperature. Errors resulting from e.g. aerosol scattering below 32 km altitude are **not** included.

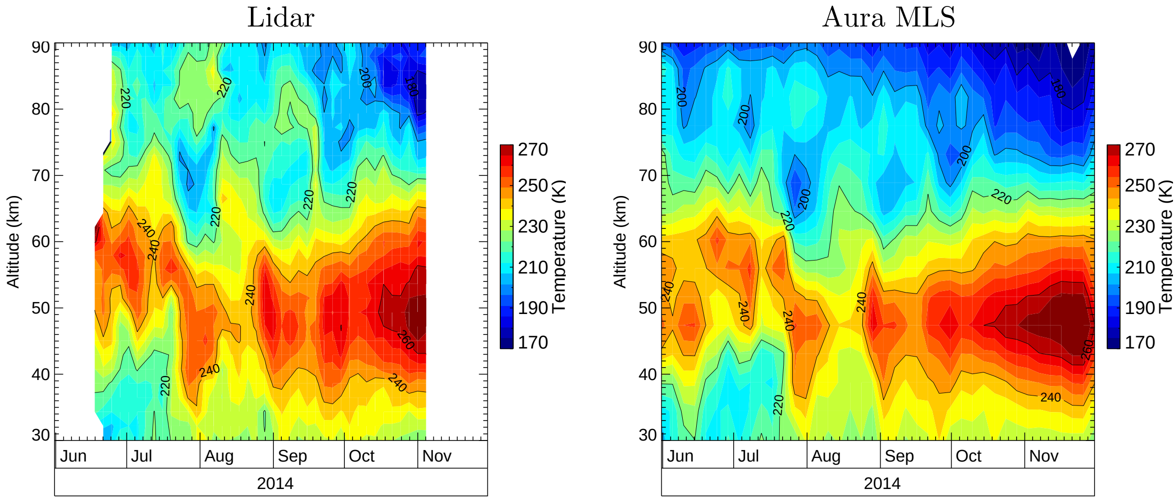

Figure 2: Nightly mean temperature profiles (left) and nearest (in space and time) temperature profiles retrieved from the MLS instrument onboard the Aura satellite.

After computing the nightly mean temperature profile, this profile is used to seed the temperature integration of two-hour binned lidar data. The two-hour temperature profiles are then used to seed one-hour binned data and so on. The sequence is: nightly mean, two-hour, one-hour, 30 min, 10 min. Temperature profiles with higher temporal resolution can be produced on request.

Before the temperature integration, lidar density profiles are smoothed in the vertical domain in order to increase the signal-to-noise ratio. Currently available are two data sets with 900 m and 2900 m vertical resolution. Data sets with other resolutions can be produced on request.

3.1 Data Discrimination

Raw photon count profiles (any combination of temporal and vertical resolution) are excluded from the analysis if any of the following criteria are met:

1. the background signal is greater than 1 kHz
2. the average lidar return signal is lower than 50 kHz between 30-35 km in case of the near channel and 45-50 km in case of the far channel
3. the signal-to-noise ratio is less than 5 below 75 km altitude (far channel)
4. the signal-to-noise ratio is less than 30 below 30 km altitude (near channel)

3.2 Intercomparison

3.2.1 Comparison with other instruments

The intercomparison is hindered by the fact that there are no co-located measurements with sufficient precision. The only available data in the middle atmosphere are satellite based measurements.

Figure 2 shows lidar temperature profiles as well as temperature profiles retrieved from the MLS instrument onboard Aura. Generally, similar structures are visible in both data sets and better agreement is observed below the stratopause. Above, the vertical structures are smeared out in MLS data as the vertical resolution becomes lower at higher altitudes, and there is also

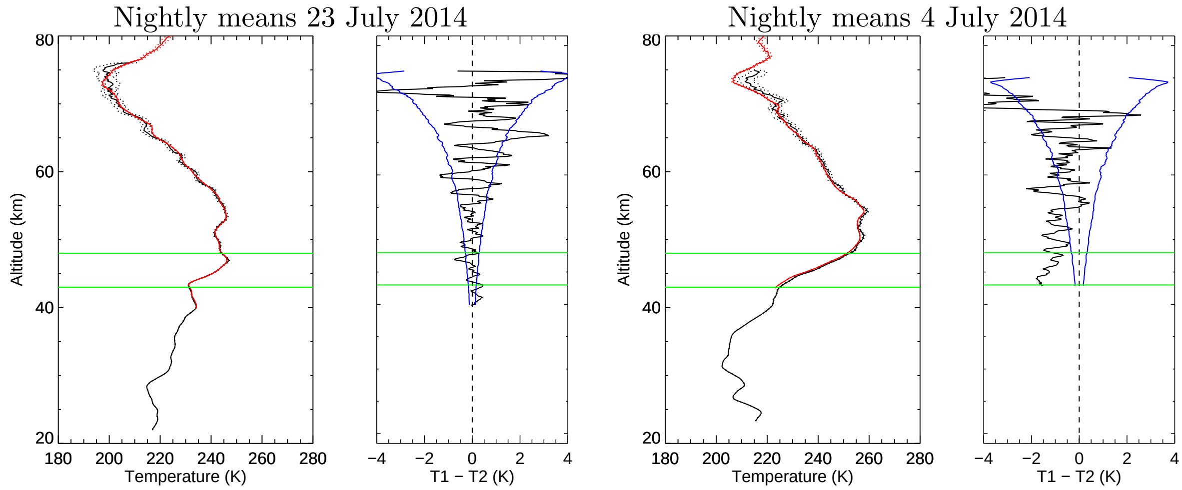

Figure 3: Comparison between temperature profiles retrieved from the far channel (red) and near channel (black). The left panel shows a typical "good" case whereas the right panel is representative for a typical "bad" case. Dotted lines mark uncertainty estimates provided by the retrieval. The blue lines show the combined standard error. For producing the combined temperature profiles, a weighted average is computed in the altitude region marked by green lines.

a cold bias in the upper mesosphere. Also horizontal sampling is certainly an issue as MLS profiles can be up to 1000 km away from the location of the lidar.

3.2.2 Comparison between different detector channels

Temperatures profiles are retrieved from the far channel (avalanche diode detector) and near channel (photo multiplier tube detector) independently. This allows for a rough estimation of temperature errors caused by instrumental (detector) effects such as detector dead-time, afterpulsing and signal induced noise. Instrumental effects may vary over time. For example, detector dead-time effects strongly depend on the maximum photon count rate which in turn is affected by atmospheric transmission (clouds).

Figure 3 shows two examples which are representative for the two extreme cases: good agreement between far and near channel temperature profiles (negligible uncorrected instrumental effects) and a rather bad case with significant instrument related deviance. Uncorrected instrumental effects result typically in temperature errors of much less than 1 K at 45 km altitude. The absolute maximum error is approximately 2 K. Note that these errors are not included in the uncertainty estimates provided with the temperature data set.

4 Data Format

Temperature data is stored in netCDF files, one file for each night. Multiple versions with different horizontal and vertical resolution exist. Filenames are structured as follows: Lauder_Rayleigh_Lidar_YYYYMMDD_TtZz.nc, where YYYY is the year, MM the month, DD the day, and t and z are the temporal (minutes) and vertical resolution (meters). An example is *Lauder_Rayleigh_Lidar_20140701_T10Z900.nc*. This file contains temperature profiles acquired on 1 July 2014 with 10 minutes temporal and 900 meters vertical resolution. Standard data products are listed in Table 1.

The data stored in the netCDF files make use of the following dimensions:

Filename	Temporal resolution	Vertical resolution
T1440V900	nightly mean	900 m
T1440V2900	nightly mean	2900 m
T120V900	120 min	900 m
T120V2900	120 min	2900 m
T60V900	60 min	900 m
T60V2900	60 min	2900 m
T30V900	30 min	900 m
T30V2900	30 min	2900 m
T10V900	10 min	900 m
T10V2900	10 min	2900 m

Table 1: Standard data products

- **time**: This dimension holds the number of time bins
- **altitude**: This dimension holds the number of vertical levels (currently fixed at 1600)
- **channels**: This dimension holds the number of detector channels. It is always 1 for merged profiles discussed in this document.

The netCDF files contain the following variables:

- **station_latitude** (double): Station location, degrees north
- **station_longitude** (double): Station location, degrees east
- **station_height** (unsigned integer): Station height in meters
- **time_offset** (unsigned integer): Seconds since 1970-01-01 00:00:00.00. This value needs to be added to the time variable.
- **altitude_offset** (unsigned integer): Offset of the **altitude** vector in meters. This value is normally zero.
- **wavelength** (double): Wavelength of the lidar (detector channel) in nanometer
- **time** (unsigned integer; vector of dimension times): This vector contains time information for each profile. The unit is milliseconds since midnight UTC of the current day. The variable **time_offset** needs to be added to the vector in order to form a valid UNIX time-stamp or Julian day. The time of element i of vector **time** can be calculated as follows:

$$\begin{aligned} \text{UNIX time} &= \text{time\_offset} + \text{time}[i]/1000 \\ \text{Julian day} &= \text{time\_offset}/86400 + \text{time}[i]/(1000 \cdot 86400) + 2440587.5 \end{aligned}$$

- **altitude** (unsigned integer; vector of dimension altitude): Altitudes for temperature profiles. The unit is meters.
- **integration_start_time** (unsigned integer; vector of dimension time): Start of the integration period in milliseconds since midnight UTC. The variable **time_offset** needs to be added to the vector in order to form a valid UNIX time-stamp.

- **integration_end_time** (unsigned integer; vector of dimension time): End of the integration period in milliseconds since midnight UTC. The variable **time_offset** needs to be added to the vector in order to form a valid UNIX time-stamp.
- **temperature** (float; matrix of dimensions time,altitude): Retrieved lidar temperatures in Kelvin
- **temperature_err** (float; matrix of dimensions time,altitude): Uncertainty estimates for retrieved lidar temperatures. The unit is Kelvin.

Each variable hast at least two attributes: **long_name** and **units**.

Note: **temperature** and **temperature_err** matrices are filled with the value 0 where no data is available.

Important global attributes:

- **version**: Dataset version, currently 27.
- **date_created**: Time-stamp string showing when the data set was created.
- **sim_runs**: Number of Monte Carlo simulations carried out for each temperature profile

Sample output produced by the program ncdump for the file Lauder_Rayleigh_Lidar_20140819_T10Z900.nc:

```
netcdf Lauder_Rayleigh_Lidar_20140826_T10Z900 {
dimensions:
        time = 145 ;
        altitude = 1600 ;
        value = 1 ;
        channels = 1 ;
variables:
        double station_latitude(value) ;
                station_latitude:units = "degrees north" ;
                station_latitude:long_name = "station latitude" ;
        double station_longitude(value) ;
                station_longitude:units = "degrees east" ;
                station_longitude:long_name = "station longitude" ;
        uint station_height(value) ;
                station_height:units = "meter" ;
                station_height:long_name = "height above WGS84" ;
        uint time_offset(value) ;
                time_offset:units = "seconds since 1970-01-01 00:00:00.00 00:00" ;
                time_offset:long_name = "time reference" ;
        uint altitude_offset(value) ;
                altitude_offset:units = "meter" ;
                altitude_offset:long_name = "station height above WGS84" ;
        double wavelength(channels) ;
                wavelength:units = "nanometer" ;
                wavelength:long_name = "wavelength" ;
        uint time(time) ;
                time:units = "milliseconds since time_offset" ;
                time:long_name = "time tag of current profile" ;
```

```
                time:resolution = 600000U ;
        uint altitude(altitude) ;
                altitude:units = "meter" ;
                altitude:resolution = 900U ;
                altitude:long_name = "height above WGS84" ;
        uint integration_start_time(time) ;
                integration_start_time:units = "milliseconds since time_offset" ;
                integration_start_time:long_name = "start of integration" ;
        uint integration_end_time(time) ;
                integration_end_time:units = "milliseconds since time_offset" ;
                integration_end_time:long_name = "end of integration" ;
        float temperature(time, altitude) ;
                temperature:units = "Kelvin" ;
                temperature:long_name = "temperature" ;
                temperature:ancillary_variables = "temperature_err" ;
        float temperature_err(time, altitude) ;
                temperature_err:units = "Kelvin" ;
                temperature_err:long_name = "temperature standard error" ;

// global attributes:
                :title = "Lidar temperature profiles" ;
                :summary = "This file contains temperature profiles retrieved
from lidar backscatter profiles assuming hydrostatic equilibrium" ;
                :comment = "" ;
                :institution = "German Aerospace Center (DLR), Institute of
Atmospheric Physics, Muenchener Strasse 20, 82234 Wessling, Germany" ;
                :creator_name = "Bernd Kaifler" ;
                :creator_url = "http://www.dlr.de/pa" ;
                :creator_email = "bernd.kaifler@dlr.de" ;
                :cmdline = "./CFConvert -db /export/data/malidar/raw/Deepwave.cfdb
-c -o /export/data/malidar/count/Deepwave_new/ -campaign DEEPWAVE -a \n./tanalysis
-tinit /export/data/Saber/SABER_Lauder_v0.nc -o /export/data/malidar/tana/Deepwave/v26/
-nsmooth 9 -simruns 500 -threads 11 -debug_HITRet 0 -debug_rebin 0
/export/data/malidar/count/Deepwave/20140826-1841.nc \n./tanalysis -combine
-i /export/data/malidar/tana/Deepwave/v26
-o /export/data/malidar/tana/Deepwave/v26combined-qc/ -a " ;
                :version = "26" ;
                :date_created = "Jan 25 16:19:56 2016" ;
                :campaign_name = "DEEPWAVE" ;
                :station_name = "Lauder" ;
                :instrument_name = "TELMA" ;
                :history = "Jan 25 16:19:56 2016, file created with MELISA 762." ;
                :date_data_start = "Aug 26 06:42:35 2014" ;
                :snr_limit = 30. ;
                :background_limit = 600. ;
                :signal_limit = 50000. ;
                :sim_runs = 500U ;
data:

 station_latitude = -45.05 ;
```

```
station_longitude = 169.68 ;

station_height = 370 ;

time_offset = 1409011200 ;

altitude_offset = 0 ;

wavelength = 532 ;

time = 24155000, 24455000, 24755000, 25055000, 25355000, 25655000, 25955000,
   26255000, 26555000, 26855000, 27155000, 27455000, 27755000, 28055000,
      ...
```

5 Data Remarks

5.1 Some general rules regarding quality of the data

The precision of Rayleigh lidar temperatures increases with increasing air density (better signal-to-noise ratio), which means profiles are generally better towards the bottom. However, the DLR Rayleigh lidar uses two independent detector channels. The far channel ranges in altitude from the top of the profile to about 43 km. Data below comes from the near channel. Best quality data can therefore be expected in the range 45 to about 70 km altitude. Above, precision and accuracy of temperature profiles are increasingly influenced by photon noise and errors in the seed temperature. At the bottom of the profiles, at altitudes below 32 km, retrieved temperatures may be influenced by stratospheric aerosols. Data in this region is included in the data set because it is still possible to extract gravity wave signatures. However, absolute temperatures may be offset due to the presence of aerosols (which the temperature retrieval does not take into account). Also, if the aerosol layer changes rapidly, this may also show up as wave-like signatures in temperature. Be warned! Please contact the PI if you want to use data below 32 km.

5.2 Issues with the lidar before 25 June 2014

Due to an error in the data acquisition software, the timestamp of the last photon for each laser pulse was lost. This skewed measured backscatter profiles at altitudes above approximately 60 km, causing artificial non-linear temperature gradients in this region. Worse, as the background level changes over time, so does the mean altitude where the second last photon (the last *recorded* photon) was detected. As a result, the artificial gradient also changes over time, defying any attempts to correct for this problem. Still, the temperature profiles can be analyzed for gravity waves when looking at *single* vertical profiles. Data are available on request, but are not included in the standard data sets.

5.3 Issues with the lidar 22 October through 3 November 2014

The original photon counter electronics died after the measurement on 16 October 2014. In the following days we reprogrammed the FPGA, which was previously used to control the lidar, to include photon counters. These *new* counters were slower to read out and introduced noise in the photon count data. However, using these counters, lidar observations could be resumed with the far channel on 22 and 23 October and both channels afterwards. An example of one

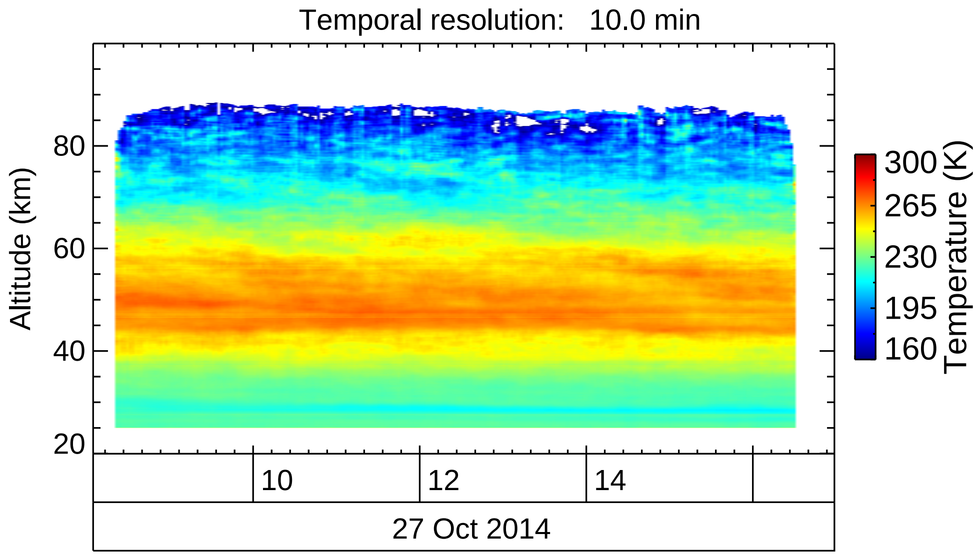

Figure 4: Lidar temperature data acquired on 27 October 2014. Note faint horizontal lines at 47 and 57 km which are caused by noise.

of the noise measurements is shown in Figure 4. Note that the noise is only visible in high vertical resolution data and does not show up in 2900 m vertical resolution profiles.

5.4 Software

An IDL routine for reading netCDF data files is provided.

6 References

Publications using the DLR lidar data:

Ehard, B., B. Kaifler, N. Kaifler, and M. Rapp (2015), Evaluation of methods for gravity wave extraction from middle-atmospheric lidar temperature measurements, *Atmospheric Measurement Techniques*, *8*(11), 4645–4655, doi:10.5194/amt-8-4645-2015.

Kaifler, B., N. Kaifler, B. Ehard, A. Drnbrack, M. Rapp, and D. C. Fritts (2015), Influences of source conditions on mountain wave penetration into the stratosphere and mesosphere, *Geophysical Research Letters*, *42*(21), 9488–9494, doi:10.1002/2015GL066465, 2015GL066465.