Attenuated atmospheric backscatter profiles measured by the CO₂ Sounder lidar in the 2017 ASCENDS/ABoVE airborne campaign

Xiaoli Sun¹, Paul T. Kolbeck², James B. Abshire¹, Stephan R. Kawa¹, and Jianping Mao¹,²

¹NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA
²University of Maryland, College Park, Maryland 20742, USA

Correspondence: Xiaoli Sun (xiaoli.sun-1@nasa.gov)

Received: 13 February 2022 – Discussion started: 17 March 2022
Revised: 22 July 2022 – Accepted: 25 July 2022 – Published: 26 August 2022

Abstract. A series of attenuated atmospheric backscatter profiles were measured at 1572 nm by the CO₂ Sounder lidar during the eight flights of the 2017 ASCENDS/ABoVE (Active Sensing of CO₂ Emission over Nights, Days, and Seasons, Arctic Boreal Vulnerability Experiment) airborne campaign. In addition to measuring the column average CO₂ mixing ratio from the laser signals reflected by the ground, the CO₂ Sounder lidar also recorded the height-resolved attenuated atmospheric backscatter profiles beneath the aircraft. We have recently processed these vertical profiles with a 15 m vertical spacing and 1 s integration time along the flight path (∼200 m) for all the 2017 flights and have posted the results at NASA Distributed Active Archive Center (DAAC) for Biogeochemical Dynamics https://doi.org/10.3334/ORNLDAAC/2051 (Sun et al., 2022). This paper describes the measurement principles, the data processing technique, and the signal to noise ratios.

1 Introduction

NASA Goddard Space Flight Center (GSFC) developed the pulsed multi-wavelength CO₂ Sounder lidar for measuring column average CO₂ mixing ratio (XCO₂) near 1572 nm (Abshire et al., 2018) as an airborne demonstrator for NASA’s planned Active Sensing of CO₂ Emission over Nights, Days, and Seasons (ASCENDS) space mission (Kawa et al., 2018). The airborne CO₂ Sounder lidar has operated successfully during several campaigns, including in the 2017 Arctic Boreal Vulnerability Experiment (ABoVE) airborne campaign onboard the NASA DC-8 aircraft.

The CO₂ Sounder lidar uses a pulsed laser and rapidly step-tunes the laser wavelength across the 1572.335 nm CO₂ absorption line. The lidar receiver uses a single photon sensitive HgCdTe avalanche photodiode (APD) (Sun et al., 2017) and an analog-to-digital convertor (ADC) that records the entire atmospheric backscatter profile, as well as the surface reflected signal. The surface returns are used to retrieve the XCO₂ of the atmosphere column traveled by the laser pulse (Sun et al., 2021). The atmospheric backscatter signals were not directly used for XCO₂ retrieval, except for cloud identifications.

Recently, we have processed the airborne atmospheric backscatter measurements as a new data product. It consists of the atmospheric backscatter profiles from the airplane flight altitude to the surface with a 15 m vertical sampling interval for all eight flights of the 2017 campaign. The results have been posted on the repository for NASA Distributed Active Archive Center (DAAC) for Biogeochemical Dynamics https://doi.org/10.3334/ORNLDAAC/2051 (Sun et al., 2022). Although the backscatter profiles are not used directly in the XCO₂ retrieval, they provide important context for interpreting the retrieved XCO₂ measurements. They can be used for retrieving partial column XCO₂ to cloud tops (Mao et al., 2018). They also enable identifying and profiling smoke plumes from wildfires to help assess their impact on XCO₂ (Mao et al., 2021).

This paper describes the data processing of the airborne CO₂ Sounder lidar’s atmospheric backscatter profiles from the eight flights of the 2017 airborne campaign. We first give a brief description of the lidar and its data structure. We then...
describe the details of the Level-0 and Level-1 data processing and the signal to noise ratio (SNR) of the profiles.

2 A brief description of the CO\textsubscript{2} Sounder lidar instrument

2.1 The airborne CO\textsubscript{2} Sounder lidar

The airborne CO\textsubscript{2} Sounder lidar transmits laser pulses toward nadir from the aircraft and detects and records the laser signals backscattered from the atmosphere and the surface. The wavelength of the single frequency laser is rapidly step-tuned across the CO\textsubscript{2} absorption line centered at 1572.335 nm. The transmitted laser pulse energies are also measured and used to normalize the received signal to correct for variations in laser power at the different wavelengths. The lidar receiver digitizes the pulse waveforms from atmospheric backscatter as well as the surface reflection. Figure 1 shows a simplified instrument block diagram. More details about the airborne CO\textsubscript{2} Sounder lidar can be found in Abshire et al. (2018).

The laser emits 1 \textmu s wide pulses with a nearly rectangular pulse shape continuously at 10 kHz. The laser wavelength is step-tuned across the CO\textsubscript{2} absorption line for 30 pulses, followed by two pulses during the period of wavelength rewind. The wavelengths of the 30 laser pulses during a wavelength scan are listed in Table 1. The laser wavelengths used by the lidar are offset-locked to the CO\textsubscript{2} absorption line in a gas cell at 40 mbar pressure and 296 K temperature. The absorption line center is at 1572.335 nm according to HITRAN 2012 (Numata et al., 2011, 2012). The offset locking frequency is changed between pulses to step the laser wavelength across the CO\textsubscript{2} absorption line.

Figure 2 shows the laser wavelengths plotted across a CO\textsubscript{2} absorption line shape that was computed from a modeled atmosphere. Note that the distribution of laser wavelengths is slightly offset from the line center of the absorption line due to the difference in pressure between the CO\textsubscript{2} in the reference cell and that of the atmosphere being modeled and measured. The laser pulse emission times, the wavelength scan, and the data acquisition are all synchronized to the Coordinated Universal Time (UTC) by the on-board computer. The lidar receiver employs a photon-sensitive HgCdTe APD detector (Sun et al., 2017). The detector has a linear analog response and has sufficient sensitivity and linear dynamic range to record both the time-resolved atmospheric backscatter profiles and the surface reflected signals. The output of lidar detector is digitized continuously with 16 bit resolution at a 100 MHz sampling rate.
set. The lidar controller suspends the laser wavelength scan after all the nine groups of data (a total of 921.6 ms) are collected. The remaining time before the next whole UTC second is used to generate the timestamp and combine the transmitted laser pulse waveforms into the same file. The pulse waveforms from the 31st and 32nd laser pulses during the wavelength rewind are discarded. The transmitted pulse waveforms are appended to the end of the 30th received waveform. Figure 3 shows a plot of a sample data file in the 16 bit signed integer format used by the ADCs.

Figure 4 shows a plot of the first 30 individual pulse waveforms from the first group taken on 8 August 2017 at 23:34:00 (UTC). The first group of pulse waveforms with relatively low amplitude are from the scatter from the aircraft’s nadir window. These are used as the time reference in the laser pulse time-of-flight measurements. The second group of waveforms about 50 µs after the window returns is from thin clouds, and the last set of waveforms are from the ground.

The laser is modulated with rectangular pulses but the amplitude of the actual laser pulse gradually decreases over the 1 µs pulse period. This is caused by the gradual depletion of energy stored in the fiber laser amplifier during the pulse. The detector has a slight ringing after the relatively strong ground return, which is omitted in the subsequent signal processing. There is also a small baseline offset from the detector in addition to the ~1.1 V DC offset added before the ADC. The total baseline offset is estimated by averaging the waveform segment before the transmitter pulses and removed in data processing. The 30 pulse waveforms reflected by the window have nearly the same pulse amplitude. The cloud and ground return pulse waveforms show increasing variation in the detected pulse amplitude caused by atmospheric CO\textsubscript{2} absorption as the laser wavelengths step through the CO\textsubscript{2} absorption line.

The XCO\textsubscript{2} is retrieved from the received laser pulse energies reflected by the surface. At each wavelength, the received pulse energy is calculated by first subtracting the DC offset, then integrating the pulse waveforms reflected by the ground surface and normalizing them by transmitted pulse energies. The relative atmospheric transmissions for each of the laser wavelengths are calculated as the normalized received pulse energies divided by the square of the range measured by the lidar. The surface reflectance is assumed constant during the wavelength scan. Changes in the received pulse energies at different wavelengths are attributed to the absorption by atmospheric CO\textsubscript{2}. The XCO\textsubscript{2} value is retrieved via an algorithm using a linear least-squares fit of the modeled CO\textsubscript{2} line shape to the lidar sampled absorption line shape. Details about the signal processing for the XCO\textsubscript{2} retrieval can be found in Sun et al. (2021).

The signals used to calculate the atmospheric backscatter profile are contained in the same data files and are obtained by processing the signal waveforms before the ground returns. Some preliminary backscatter profiles in raw ADC units (integers) were reported earlier for this campaign (Allan et al., 2018). We have since then calibrated the backscatter profiles and released results on NASA DAAC for Biogeochemical Dynamics https://doi.org/10.3334/ORNLDAAC/2051 (Sun et al., 2022). Details of the data processing and the estimated SNR of the profiles are presented in the remainder of this paper.

### Table 1. List of laser wavelengths used in the 2017 airborne campaign for the CO\textsubscript{2} Sounder lidar. All wavelengths are given as the difference of the actual wavelength from the center of the 1572.335 nm CO\textsubscript{2} absorption line in the reference cell.

<table>
<thead>
<tr>
<th>Laser pulse number</th>
<th>Frequency and wavelength difference from the line center of the CO\textsubscript{2} cell (GHz)</th>
<th>Picometer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.25</td>
<td>−101.013</td>
</tr>
<tr>
<td>2</td>
<td>9.50</td>
<td>−78.388</td>
</tr>
<tr>
<td>3</td>
<td>8.25</td>
<td>−68.031</td>
</tr>
<tr>
<td>4</td>
<td>6.25</td>
<td>−51.539</td>
</tr>
<tr>
<td>5</td>
<td>4.25</td>
<td>−35.047</td>
</tr>
<tr>
<td>6</td>
<td>3.25</td>
<td>−26.801</td>
</tr>
<tr>
<td>7</td>
<td>2.50</td>
<td>−20.616</td>
</tr>
<tr>
<td>8</td>
<td>2.00</td>
<td>−16.493</td>
</tr>
<tr>
<td>9</td>
<td>1.75</td>
<td>−14.431</td>
</tr>
<tr>
<td>10</td>
<td>1.50</td>
<td>−12.370</td>
</tr>
<tr>
<td>11</td>
<td>1.25</td>
<td>−10.308</td>
</tr>
<tr>
<td>12</td>
<td>1.00</td>
<td>−8.246</td>
</tr>
<tr>
<td>13</td>
<td>0.75</td>
<td>−6.185</td>
</tr>
<tr>
<td>14</td>
<td>0.50</td>
<td>−4.123</td>
</tr>
<tr>
<td>15</td>
<td>0.25</td>
<td>−2.062</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Laser pulse number</th>
<th>Frequency and wavelength difference from the line center of the CO\textsubscript{2} cell (GHz)</th>
<th>Picometer</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>−0.25</td>
<td>2.062</td>
</tr>
<tr>
<td>17</td>
<td>−0.50</td>
<td>4.123</td>
</tr>
<tr>
<td>18</td>
<td>−0.75</td>
<td>6.185</td>
</tr>
<tr>
<td>19</td>
<td>−1.00</td>
<td>8.247</td>
</tr>
<tr>
<td>20</td>
<td>−1.25</td>
<td>10.308</td>
</tr>
<tr>
<td>21</td>
<td>−1.50</td>
<td>12.370</td>
</tr>
<tr>
<td>22</td>
<td>−1.75</td>
<td>14.432</td>
</tr>
<tr>
<td>23</td>
<td>−2.00</td>
<td>16.493</td>
</tr>
<tr>
<td>24</td>
<td>−2.50</td>
<td>20.617</td>
</tr>
<tr>
<td>25</td>
<td>−3.25</td>
<td>26.802</td>
</tr>
<tr>
<td>26</td>
<td>−4.25</td>
<td>35.048</td>
</tr>
<tr>
<td>27</td>
<td>−6.25</td>
<td>51.542</td>
</tr>
<tr>
<td>28</td>
<td>−8.25</td>
<td>68.037</td>
</tr>
<tr>
<td>29</td>
<td>−9.50</td>
<td>78.346</td>
</tr>
<tr>
<td>30</td>
<td>−12.25</td>
<td>101.026</td>
</tr>
</tbody>
</table>

https://doi.org/10.5194/essd-14-3821-2022
Figure 3. Plot of the recorded data file structure for the CO\textsubscript{2} Sounder lidar. The data were taken on 8 August 2017 at 23:30:00 (UTC) over the mountains as the aircraft flew toward the Edwards Airforce Base in California at an altitude of 11.2 km. There are nine groups of 30 pulse waveforms, one for each laser wavelength in the scan. Each is averaged over 32 scans. The corresponding waveforms for the 30 transmitted laser pulses, which are sampled at the same rate for 1000 points each, are appended to the end of the 30th received pulse waveform.

Figure 4. Overlay of the 30 received pulse waveforms recorded by the airborne lidar receiver on 8 August 2017 at 23:34:00 UTC.
3 Data processing of the atmospheric backscatter profiles from the CO₂ Sounder lidar

3.1 Level-0 data processing

The Level-0 data processing converts the raw ADC output in 16 bit integers to the atmospheric backscattered signal in the physical units of the detector output. It also removes all the known artifacts in the data, such as baseline DC offsets. To improve the SNR we averaged the backscatter profiles measured for laser wavelengths 2, 3, 4, 27, 28, 29, and 30 (seven in total). We did not include the first wavelength of the scans because occasionally the laser operation was not fully settled after the wavelength rewind. As shown in Fig. 2, these wavelengths undergo only small CO₂ absorption, so this average may be considered as a backscatter measurement at “off-line” wavelengths.

For each received pulse waveform, the DC offset from the detector is first subtracted by calculating the average detector signal for the time interval that occurs immediately before the return from aircraft window. Since the pre-window-return region is long after the ground return from the previous laser pulse but before the next one, it is composed primarily of detected solar background, the dark noise from the detector, and the DC offset of the lidar receiver. One component of the DC offset is a constant offset (−1.1 V) added to the detector output before the ADC. The other component is the inherent DC offset from the detector, which can drift slowly over time. It can also change after the detector is overloaded by excessively strong signals, such as returns from clouds in close range or returns from specular reflection at the ground.

The transmitted pulse energies are calculated from the transmitted pulse waveforms by first subtracting the DC offset and then integrating the signal over the pulse period. They are then used to normalize the corresponding received waveform. The average transmitted pulse energy over the entire flight is also recorded in the Level-0 data file. A known time delay within the lidar electrical and optical system (corresponding to a range of 26.4 m) is removed next. Following this, received pulse waveforms in the 1 s data file are averaged together to improve the SNR. Finally, the averaged pulse waveform is converted into the units of the detector output (V) by dividing the waveform data by the scale factor of the ADC.

Screening is then performed to eliminate any data with the following abnormalities: the transmitted pulse waveforms are missing in the data file, the ground return pulse is saturated with pulse amplitude above 1.1 V after removing the DC offset, or data are taken while the detector is still recovering from saturation where the estimated detector DC offset falls below 0 V or above 0.5 V after removing the known 1.1 V offset (nominal DC offset from the detector is 0.1 to 0.2 V). The mean and standard deviation of the waveform within the pre-window region for each averaged waveform is then recorded in the Level-0 data files in case further screening is needed. Lidar data that contain cloud returns within 3 km of the aircraft window are also flagged because within this distance, the laser beam does not completely overlap the receiver field of view (Allan et al., 2018). The data within this range needs to be further calibrated to account for the overlap factor.

The aircraft’s navigation data are then merged with the lidar data. The aircraft flight data are obtained from archived airborne campaign housekeeping data gathered during the flight from the on-board GPS receiver and other aircraft instruments. The following parameters are extracted from the archived DC-8 housekeeping data for each second: the UTC time, latitude, longitude, altitude, pitch and roll angles, and the DC-8’s radar-measured range to the surface in nadir direction.

The effect of off-nadir pointing is corrected next. The lidar is mounted to the aircraft’s deck, and the laser beam is pointed down perpendicular to the deck. The aircraft flies slightly pitched up during cruise. It also performs pitch and roll maneuvers from time to time. When the laser beam is pointed away from nadir, the laser pulse travels a slant path (a longer distance) before reaching the ground. This effect, if not corrected, would result in a longer length profile and a lower ground elevation when plotted with other profiles taken along nadir direction. To correct this effect, the range bin size (15 m) is first multiplied by the cosine of the combined off-nadir pointing angle. This has the effect of compressing the profile and correcting the elevation of the ground return. The resultant profile is then re-sampled by interpolating at 15 m intervals in the nadir direction. The attenuated backscatter coefficients in the corrected range bins should also be corrected for the additional attenuation by the atmosphere above the layer. Under uniform layer structure, atmospheric transmission from the aircraft to the given altitude equals $T(\theta_{\text{nadir}}) = [T(0)]^{1/\cos(\theta_{\text{nadir}})}$, with $\theta_{\text{nadir}}$ the laser beam pointing angle from nadir. The aircraft pitch angle was mostly within 5° during cruise. The roll angle was near zero, except during sharp turns or spiral-downs. The total decrease in the two-way transmission is < 5% under this condition. We did not correct this effect in the current data release since it is a relatively small effect for this dataset. The pitch and roll angles are included in the released data; the user may apply this correction if needed.

The surface elevation is calculated as the difference between the aircraft altitude from the on-board GPS receiver and the range measured by the lidar after correcting for the effect of off-nadir pointing angle. Missing points in the surface elevation, such as those obscured by clouds, are interpolated linearly between the elevations that border the missing region. The aircraft’s latitude, longitude, altitude, pitch, and roll angle are also linearly interpolated during these conditions, if the gaps are small (seconds), and the change in measured values before and after the gap is small. All data points in the recorded waveforms above the aircraft altitude and below the ground return are removed. In the cases where
there is no discernible ground return, points more than 200 m below the estimated surface elevation are removed.

Finally, the signal waveforms are vertically smoothed via a boxcar averaging window with the integration time equal to the 1 μs laser pulse width. Since the vertical resolution of the atmospheric backscatter measurements is primarily limited by the laser pulse width, the boxcar averaging has little effect on the vertical resolution of the backscatter signal, but it reduces higher frequency noise generated by the detector and the electronics.

3.2 Level-1 data processing

The Level-1 data processing converts the waveforms from the Level-0 data processing to the attenuated atmospheric backscatter profile in terms of cross section per unit volume per steradian (m⁻¹ sr⁻¹), also known as the attenuated backscatter coefficient.

The optical signal power collected by the lidar can be written as

\[ y(t) = E_{\text{lx}} \int_0^{\tau_1} x(\tau) h(t - \tau) \, d\tau, \]

where \( E_{\text{lx}} \) is the transmitted laser pulse energy in joules, \( x(\tau) \) is the normalized laser pulse shape, i.e., \( \int_0^{\tau_1} x(\tau) \, d\tau \equiv 1 \), \( \tau_1 \) is the laser pulse width, and \( h(t) \) is the impulse response of the atmosphere column, i.e., the normalized receiver response to an infinitely short laser pulse. Here, we assume the that lidar detector has a much faster time response than the laser pulse width, and the detector’s output pulse shape is the same as that of the received optical signal.

The impulse response of the atmospheric column is related to the volume atmospheric backscatter coefficient profile by the lidar equation (Measures, 1984; Reagan et al., 1989)

\[ h(t) = \frac{c}{2} \beta[R(t)] \frac{T_2^2[R(t)]}{R^2(t)} \frac{\eta A_{\text{tel}}}{C_1}, \]

Here, \( c \) is the speed of light, \( \beta[R] \) is the backscatter coefficient at range \( R \) in units of m⁻¹ sr⁻¹, \( T_2[R(t)] \) is the one-way atmospheric transmission from the lidar to range \( R \), \( \eta \) is the lidar receiver optical transmission, \( A_{\text{tel}} \) is the collection area of the receiver telescope, \( R(t) = c t / 2 \) is the lidar range at time \( t \) after the laser pulse emission, and \( C_1 \) is the lidar detector responsivity in \( \text{V/W} \)

Substituting Eq. (2) into (1), the lidar signal from the detector can be written as

\[ s(t) = C_1 \left( \frac{c}{2} \eta A_{\text{tel}} E_{\text{lx}} \int_0^{\tau_1} x(\tau) \right) \frac{\beta[R(t - \tau)] T_2^2[R(t - \tau)]}{R^2(t - \tau)} \, d\tau. \]

If we assume that the atmospheric backscatter coefficient and the atmospheric transmission are approximately constant over the laser pulse interval (1 μs, or 150 m in range), the lidar signal from the detector can be written as

\[ s(t) \approx C_1 \left( \frac{c}{2} \eta A_{\text{tel}} E_{\text{lx}} \right) \frac{\beta[R(t)] T_2^2[R(t)]}{R^2(t)} \]

\[ = C_2 \frac{1}{R^2(t)} \langle \beta[R(t)] T_2^2[R(t)] \rangle, \]

where \( C_2 = C_1 \eta A_{\text{tel}} E_{\text{lx}} \) is a constant that depends only on the parameters of the lidar. Therefore, the attenuated atmospheric backscatter coefficient as a function of range can be calculated from the lidar measurement by

\[ \langle \beta[R(t)] T_2^2[R(t)] \rangle = \left[ R^2(t) s(t) \right] \frac{1}{C_2}. \]

The lidar signal term \( s(t) \) in the above equation can be expressed in terms of lidar range by substituting \( t = 2R/c \). The attenuated backscatter coefficient can also be expressed as a function of altitude by subtracting the lidar range from the flight altitude of the aircraft.

For the CO₂ Sounder lidar configuration used in the 2017 airborne campaign, the instrument constants are:

\[ C_1 = \left( \eta_{\text{det}} G_{\text{APD}} \frac{1}{E_{\text{ph}}} \frac{e c}{h} \right) Z_{\text{TIA}} G_{\text{amp}} \text{ cable}, \quad \text{V/W}^{-1}, \]

\[ C_2 = C_1 \eta A_{\text{tel}} E_{\text{lx}} \frac{c}{2} = C_1 \frac{\pi \phi_{\text{tel}}^2}{4} (1 - L_{\text{obs}}) E_{\text{lx}} \frac{c}{2}, \quad \text{V/m}^3. \]

Here, \( \eta_{\text{det}} \) is the detector quantum efficiency, \( G_{\text{APD}} \) is the APD gain, \( E_{\text{ph}} = h c / \lambda_{\text{laser}} \) is the laser photon energy, \( h \) is Planck’s constant, \( \lambda_{\text{laser}} \) is the laser wavelength, \( e c \) is the electron charge, \( Z_{\text{TIA}} \) is the gain of the preamplifier (a transimpedance amplifier) in \( \text{V/W}^{-1} \), \( G_{\text{amp}} \) is the post amplifier voltage gain, \( \eta_{\text{cable}} \) is the electrical cable transmission, \( \phi_{\text{tel}} \) is the receiver telescope diameter, and \( L_{\text{obs}} \) is the fractional loss of the telescope’s collection area caused by its center obscuration.

Table 2 lists the values of these parameters for the 2017 CO₂ Sounder lidar. The resultant detector responsivity is \( C_1 = 6.39 \times 10^8 \text{V/W} \) for nominal APD gain, and the instrument constant is \( C_2 = 5.13 \times 10^{10} \text{V m}^3 \).

The received signal level from the surface varies with lidar range. To prevent detector saturation, the gain of the lidar detector was adjusted manually during the flight to keep the received signal within the detector linear dynamic range. The detector gain was changed in steps of 2 by adjusting the bias voltage of the HgCdTe APD detector. In the data processing, the detector’s gain value was determined from the pulse amplitude of the window returns. After filtering out anomalous window returns, which generally resulted from the aircraft flying through clouds, the pulse energies and the centroid of the window returns are calculated and binned for each 1 s data file for each flight. The window return pulse energies are divided into groups separated by a factor of 2. The average value from the bins with the APD biased at 10 V is defined as
Table 2. CO₂ Sounder Lidar parameters as used in the 2017 airborne campaign.

<table>
<thead>
<tr>
<th>Instrument parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laser pulse energy, $E_{tx}$</td>
<td>25 µJ</td>
</tr>
<tr>
<td>Laser pulse width, $\tau_L$</td>
<td>1.0 µs</td>
</tr>
<tr>
<td>Laser pulse rate</td>
<td>10 kHz</td>
</tr>
<tr>
<td>Laser wavelengths, $\lambda_{laser}$</td>
<td>1572.2 to 1572.5 nm</td>
</tr>
<tr>
<td>Telescope diameter, $\phi_{tel}$</td>
<td>0.20 m</td>
</tr>
<tr>
<td>Telescope center obscuration, $L_{obs}$</td>
<td>16 %</td>
</tr>
<tr>
<td>Receiver optical transmission, $\eta_r$</td>
<td>81.3 %</td>
</tr>
<tr>
<td>Receiver field of view, $\theta_{FOV}$</td>
<td>500 µrad</td>
</tr>
<tr>
<td>Receiver optical filter width, $\Delta \lambda_f$</td>
<td>1.4 nm</td>
</tr>
<tr>
<td>Receiver integration time</td>
<td>1 s</td>
</tr>
<tr>
<td>Detector quantum efficiency, $\eta_{det}$</td>
<td>69.3 %, including the fill factor</td>
</tr>
<tr>
<td>Detector avalanche gain, $G_{APD}$</td>
<td>190 (at 10 V APD bias)</td>
</tr>
<tr>
<td>Detector excess noise factor, $F_{ex}$</td>
<td>1.05</td>
</tr>
<tr>
<td>Receiver noise equivalent power, NEP</td>
<td>1.7 fW Hz$^{-1/2}$</td>
</tr>
<tr>
<td>Transimpedance amplifier gain, $Z_{TIA}$</td>
<td>0.5 MHz</td>
</tr>
<tr>
<td>Post amplifier voltage gain, $G_{amp}$</td>
<td>320 kV A$^{-1}$</td>
</tr>
<tr>
<td>Cable transmission, $\eta_{cable}$</td>
<td>12.6</td>
</tr>
<tr>
<td>Overall receiver responsivity</td>
<td>$6.39 \times 10^8$ V W$^{-1}$ at 10 V APD bias, decreases by a factor of two for every 1 V reduction in APD bias</td>
</tr>
</tbody>
</table>

For each second, the attenuated atmosphere backscatter coefficient is obtained by multiplying the signal waveforms from Level-1 data processing by the square of the lidar range and then dividing it by the lidar scaling factor $C_2$ according to Eq. (5). Figure 5 shows the atmosphere backscatter profile for the flight on 8 August 2017 along the flight path shown in Fig. 6. Figures 7 and 8 show expanded views of the same data. During this flight, aircraft flew over mountains, ocean, desert, etc. (Fig. 6) and performed seven spiral-down maneuvers to compare the lidar $X_{CO₂}$ measurements with those from the in situ gas analyzer on board the aircraft. The surface type and reflectance affected the solar background noise in the atmospheric backscatter measurement. For example, when the aircraft flew over ocean surface in between Spiral-Downs #3 and #4, the scattered sunlight was much lower than that for land surfaces, and the noise in the received data became noticeably lower, as shown in Fig. 5. Figure 9 shows a line plot of the vertical atmospheric backscatter profile measured at time 23:34:00 UTC.

The raw lidar signal is digitized at 10 ns (1.5 m) intervals by the ADC to enable meter-level calculations of the lidar range to the surface. However, the vertical resolution of the computed atmospheric backscatter profiles is much wider due to the 1 µs (150 m) laser pulse width. To reduce the data volume, the backscatter profile data are re-sampled with 100 ns (15 m) bin width. This still oversamples the backscatter profile but preserves certain temporal features in the data, such as the ground returns and cloud boundaries. The horizontal resolution along the aircraft ground track is the distance traveled by the aircraft in 1 s, or about 200 m. These results from most flights showed the values of attenuated backscatter coefficients within the boundary layer are comparable to those reported by Spinhirne et al. (1997), which were measured at a nearby (1540 nm) laser wavelength. Similar atmospheric backscatter profiles have been measured with a ground-based 2 µm lidar that used the same type of HgCdTe APD detector (Refaat et al., 2020).

These backscatter profiles also contain the broadened lidar returns from the ground expressed in terms of attenuated backscatter coefficient. With the 1.0 µs wide laser pulse and the 1.0 µs wide boxcar smoothing, the smoothed return signal is usually triangular with a width of 10 range bins at the half maximum points and 20 range bins at the base. Any changes in surface elevation during the receiver integration time can cause additional broadening of the return pulse shape to 15 to 20 range bins (1.5 to 2 µs), as shown in Fig. 9. These ground return signals at the end of the backscatter profiles can be
used to calculate the attenuated surface reflectance by summing the backscatter values over all the range bins containing the ground return and then multiplying the results by \( \pi \) and the bin width (15 m). Figure 10 shows the attenuated surface reflectance as the aircraft approached the Edward Air Force Base on 8 August 2017. Over a mountainous area for the time period from 23:00:00 to 23:40:00 UTC, the attenuated surface reflectance varied between about 4% when there were clouds and 18% when it was cloud free. The average attenuated surface reflectance from 00:01:00 to 00:11:00 UTC over desert near the end of the flight was 22%. If we estimate the average desert reflectance to be 45% (Kuze et al., 2011), the one-way atmospheric transmission would be 70%, which is consistent with the sky condition shown on the aerial photograph (Fig. 11). When the surface reflectance can be estimated by some other means, the lidar profiles may be used to calculate the extinction to backscatter ratio of thin clouds and aerosols in the path.

### 4 Signal to noise ratio calculations

To derive the SNR of the attenuated backscatter profiles we first express the average detected signal from each laser pulse given by Eq. (4) in terms of the rate of detected signal photons as a function of the lidar range, as

\[
\langle \dot{n}_{\text{sig}}(R) \rangle \approx \frac{c}{2} \frac{E_{\text{tx}}}{E_{\text{ph}}} A_{\text{tel}} \eta_{\text{det}} \eta_{\text{r}} \left( \frac{T_{\text{a}}^2(R)}{R^2} \right) \langle \beta(R) \rangle,
\]

where \( R \) is the lidar range.

For each laser pulse, the standard deviation of the rate of the detected signal photons can be calculated as (Gagliardi and Karp, 1995)

\[
\sigma_{\text{noise}}(R) = \sqrt{F_{\text{ex}} \left( \langle \dot{n}_{\text{sig}}(R) \rangle + \langle \dot{n}_b \rangle \right) + \left( \frac{\text{NEP} \lambda_{\text{laser}} c}{h} \right)^2 B_n},
\]

where \( F_{\text{ex}} \) is the detector excess noise factor from the randomness of the APD gain, \( \langle \dot{n}_b \rangle \) is the average detected background photon rate, NEP is the detector noise equivalent power due to the detector dark noise and the preamplifier noise, and \( B_n \) is the noise bandwidth of the lidar receiver used for the atmospheric backscatter calculations. The noise bandwidth is equal to \( 1/2t_{\text{box}} \) with \( t_{\text{box}} \) the width of the boxcar averaging window in the signal processing.
Figure 7. Same as Fig. 5 but expanded for the last segment from time 23:00:00 UTC to the end of the flight.

Figure 8. Same as Fig. 7 but expanded for the 5 min between 23:30:00 and 23:35:00 UTC.

Figure 9. Plot of the attenuated backscatter profile and ground return measured on 8 August 2017 at 23:34:00 UTC with a 15 m vertical sampling interval. The results are from the same data shown in Fig. 4 after Level 1 processing.

Figure 10. Product of the surface reflectance and two-way atmospheric transmission calculated by integrating the ground return of the backscatter profile as the aircraft flew toward the Edwards Air Force Base during 8 August 2017.
For the CO$_2$ Sounder lidar, the lidar range $R$ in Eqs. (9) and (10) consists of a discrete set of lidar ranges for a series of range bins. Although the attenuated atmospheric backscatter profiles are given in the 15 m range bin size, the actual profiles are averaged over 150 m vertical layers due to the effect of the laser pulse width and the boxcar smoothing in the data processing. Therefore, the noise standard deviation given in Eq. (10) is for the 150 m vertical integration interval. This noise may be further reduced by averaging data from adjacent 150 m vertical atmospheric layers.

When the lidar views a sunlit diffuse scattering surface through a clear sky the average detected background photon rate is given by

$$
\langle \hat{n}_b \rangle = I_s \Delta \lambda_f \pi \left( \frac{\theta_{FOV}}{2} \right)^2 \rho \cdot A_{tel} \eta_r \eta_{det},
$$

(10)

where $I_s$ is the solar irradiance at the 1572 nm laser wavelength, $\Delta \lambda_f$ is the receiver optical filter bandwidth, $\theta_{FOV}$ is the receiver field of view diameter, and $\rho$ is the diffuse reflectance of the surface. These and other instrument parameter values are listed in Table 2.

For a single laser pulse, the SNR of the attenuated backscatter coefficient measurement is the same as the SNR of the rate of detected signal photons, which is the ratio of Eqs. (9) to (10). Therefore, the SNR of a 150 m vertical bin at range $R$ when averaged over the number of laser pulse measurements within the 1 s receiver integration time is given by

$$
\text{SNR}_{av}(R) = \sqrt{\frac{N_{ave} \langle \hat{n}_{sig}(R) \rangle}{\sigma_{noise}(R)}},
$$

(11)

with $N_{ave}$ the total number of laser pulses measured within the receiver integration time. For the CO$_2$ Sounder lidar with a 1 s receiver integration time, $N_{ave} = 32 \times 9 \times 7 = 2016$.

The SNR is proportional to the attenuated backscatter coefficient and inversely proportional to the square of the lidar range. Hence, for measurements from an airborne lidar, the SNR is a strong function of the range to the atmospheric layers being measured. Figure 12 shows the calculated SNR of the attenuated backscatter coefficient vs. the lidar range for $10^{-6}$, $10^{-5}$, and $10^{-4}$ m$^{-1}$ sr$^{-1}$ backscatter coefficients at local noon and night. For example, for a 12 km aircraft altitude above ground surface and a 1 s receiver integration time, the averaged SNR is about 65 for clouds with a backscatter coefficient of $10^{-5}$ m$^{-1}$ sr$^{-1}$ that are 7 km below the aircraft (5 km above surface) during daytime at 45$^\circ$ sun angle. Therefore, these attenuated backscatter profiles can be used to reliably identify mid-altitude clouds. The signals from aerosols at lower altitudes are much weaker. For example, at 10 km lidar range in the daytime, the averaged SNR for a 1 s integration time is about 2.0 for aerosols with a backscatter coefficient of $10^{-6}$ m$^{-1}$ sr$^{-1}$.

As a comparison, we calculated the average attenuated backscatter coefficient profile and the corresponding SNR from 100 consecutive backscattering profiles starting from 8 August 2017, at 23:55:00 UTC (a small cloud-free segment shown in Fig. 7). The mean and the ratio of the mean to standard deviation are plotted in Fig. 13. The ratio of the mean to standard deviation gives an estimate of the SNR. It is about 2.4 for attenuated backscattering coefficients of $10^{-6}$ m$^{-1}$ sr$^{-1}$ at 4.8 km altitude (7.8 km lidar range). The extrapolated SNR at 10 km lidar range is 1.5. The SNR estimated from the measurement data is expected to be lower since it also includes the effects of small atmosphere variation during the measurement period. The SNR can be improved by averaging over more measurements along the flight path and thicker vertical layer.

5 Data availability

The atmospheric backscatter profiles by the CO$_2$ Sounder from the 2017 airborne lidar measurements are available from the NASA Oak Ridge National Laboratory (ORNL) Distributed Active Archive...
Figure 12. Estimated SNR of the attenuated backscatter coefficients obtained from the CO\textsubscript{2} Sounder lidar in the 2017 airborne campaign for 1 s integration times and 150 m vertical resolution. The SNR of the backscatter profiles may be further improved by averaging over a longer time intervals and by using thicker atmospheric layers.

Figure 13. Average attenuated backscatter coefficients and the ratio of the mean to standard deviation calculated from 100 consecutive backscattering profiles starting on 8 August 2017 at 23:55:00 UTC.

6 Code availability

The software codes used for the data processing of the atmospheric backscatter profiles described in this paper are posted on the same data depository website as the data, https://doi.org/10.3334/ORNLDAAC/2051 (Sun et al., 2022).

7 Conclusions

In addition to measuring the column average CO\textsubscript{2} mixing ratio (XCO\textsubscript{2}), the NASA GSFC CO\textsubscript{2} Sounder lidar measures the attenuated atmospheric backscatter profiles in the laser beam’s path. We have recently processed the atmospheric backscatter profiles from the 2017 ASCENDS/ABoVE (Active Sensing of CO\textsubscript{2} Emission over Nights, Days, and Seasons /Arctic Boreal Vulnerability Experiment) airborne campaign and produced a new dataset of range-resolved attenuated atmospheric backscatter coefficients at 1572 nm laser wavelength measured for each of the eight flights. The analysis shows that the signal to noise ratios of the CO\textsubscript{2} Sounder lidar backscatter profiles are sufficient to identify clouds and estimate the height of aerosol layers. This new dataset also provides additional information to help interpret the retrieved XCO\textsubscript{2}. These same types of measurements may be obtained in the future from a space-based lidar that uses the same measurement technique and scaled for a similar power-aperture product.
Appendix A

The following is the flowchart of the data processing software, which was developed using Matlab. The software consists of a main program “backscatterProcessing.m” and a set of functions indicated in blue text. The software code is available at the same website, where the atmospheric backscatter profiles are archived (https://doi.org/10.3334/ORNLDAAC/2051, Sun et al., 2022).

**Figure A1.** Flowchart of the software used to process the data described in this paper.
**Author contributions.** XS led the data processing and the writing of the manuscript; PTK developed the algorithm and the software to process the data; JBA led the CO$_2$ Sounder lidar team, the instrument development, the 2017 airborne campaign, and its data analysis; SRK led the data archiving to the NASA Airborne Science Data website; and JM processed the 2017 airborne data for XCO$_2$ retrieval; all co-authors participated the writing of the manuscript.

**Competing interests.** The contact author has declared that none of the authors has any competing interests.

**Disclaimer.** Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Acknowledgements.** We are grateful to Haris Riris, Bill Hasselbrack, Jeffrey Chen, and Kenji Numata at NASA Goddard for the development of the CO$_2$ Sounder lidar, supporting the 2017 airborne campaign and for collecting the lidar measurements. We are particularly grateful to the late Graham Allan for his work on the CO$_2$ Sounder lidar, participating in its airborne campaigns and his early work in analyzing the atmospheric backscatter profile data.

**Financial support.** This research has been supported by the NASA Earth Sciences Technology Office (ESTO), the NASA ASCENDS Mission pre-formulation program and the 2017 NASA Arctic Boreal Vulnerability Experiment campaign.

**Review statement.** This paper was edited by Nellie Elguindi and reviewed by three anonymous referees.

**References**


https://doi.org/10.5194/essd-14-3821-2022