



# Methodology for deriving the telescope focus function and its uncertainty for a heterodyne pulsed Doppler lidar

Pyry Pentikäinen<sup>1</sup>, Ewan James O'Connor<sup>2,3</sup>, Antti Juhani Manninen<sup>2</sup>, and Pablo Ortiz-Amezcu<sup>4,5</sup>

<sup>1</sup>Institute for Atmospheric and Earth System Research/Physics, Faculty of Science, University of Helsinki, Helsinki, Finland

<sup>2</sup>Finnish Meteorological Institute, Helsinki, Finland

<sup>3</sup>Department of Meteorology, University of Reading, Reading, UK

<sup>4</sup>Andalusian Institute for Earth System Research (IISTA-CEAMA), 18006 Granada, Spain

<sup>5</sup>Department of Applied Physics, University of Granada, 18071 Granada, Spain

**Correspondence:** Pyry Pentikäinen (pyry.pentikainen@helsinki.fi)

Received: 21 December 2019 – Discussion started: 9 January 2020

Revised: 9 April 2020 – Accepted: 27 April 2020 – Published: 29 May 2020

**Abstract.** Doppler lidars provide two measured parameters, radial velocity and signal-to-noise ratio, from which winds and turbulent properties are routinely derived. Attenuated backscatter, which gives quantitative information on aerosols, clouds, and precipitation in the atmosphere, can be used in conjunction with the winds and turbulent properties to create a sophisticated classification of the state of the atmospheric boundary layer. Calculating attenuated backscatter from the signal-to-noise ratio requires accurate knowledge of the telescope focus function, which is usually unavailable. Inaccurate assumptions of the telescope focus function can significantly deform attenuated backscatter profiles, even if the instrument is focused at infinity. Here, we present a methodology for deriving the telescope focus function using a co-located ceilometer for pulsed heterodyne Doppler lidars. The method was tested with Halo Photonics StreamLine and StreamLine XR Doppler lidars but should also be applicable to other pulsed heterodyne Doppler lidar systems. The method derives two parameters of the telescope focus function, the effective beam diameter and the effective focal length of the telescope. Additionally, the method provides uncertainty estimates for the retrieved attenuated backscatter profile arising from uncertainties in deriving the telescope function, together with standard measurement uncertainties from the signal-to-noise ratio. The method is best suited for locations where the absolute difference in aerosol extinction at the ceilometer and Doppler lidar wavelengths is small.

## 1 Introduction

Coherent Doppler lidar systems are capable of providing accurate radial Doppler velocities at high temporal and spatial resolution and have been employed across a wide range of scientific and operational fields. Meteorological applications include the retrieval of turbulent properties to determine the strength, location, and source of mixing in the atmospheric boundary layer and, with many systems having scanning capability, the retrieval of winds. Information on the targets responsible for the radial Doppler velocities measured by the Doppler lidar (e.g. aerosol, cloud, precipitation) would greatly aid the interpretation of both the velocities and the products derived from them, but this requires quantitative use of the signal power received by the instrument.

The performance of a Doppler lidar depends on the signal-to-noise ratio, SNR, of the system, as SNR determines the radial velocity uncertainty (Rye and Hardesty, 1993; Pearson et al., 2009). The outgoing laser beam can be focused to improve the SNR at ranges close to the focal length (Pearson et al., 2002), and this is often used to improve the Doppler lidar velocity data quality and data availability, particularly in the atmospheric boundary layer. The optimal choice of focus will depend on the atmospheric conditions at the deployment location (Hirsikko et al., 2014).

Knowledge of how the choice of instrument parameters, such as the effective focal length of the telescope, impacts the SNR profile is necessary in order to obtain profiles of the attenuated backscatter coefficient (Zhao et al., 1990). A com-

prehensive overview of the theoretical considerations in determining the performance of coherent Doppler lidar systems was given by Frehlich and Kavaya (1991), who provided analytical expressions for deriving the expected signal measured by the coherent detector for a given target for a range of instrument configurations, including analytical expressions for the telescope focus function (also termed coherent responsivity). Most analytical expressions assume ideal Gaussian beams, which may not always be appropriate (Hill, 2018); hence experimental approaches have also been used to determine the impact of beam aberrations (Hu et al., 2013).

The profile of the attenuated backscatter coefficient has the potential to be used in real time by weather forecasters (Illingworth et al., 2019), as it can be used in the same manner as for ceilometers. This includes the detection of liquid, supercooled-liquid, mixed-phase, and ice clouds (Hogan et al., 2003; Van Tricht et al., 2014; Tonttila et al., 2015), aerosol layer and mixing-height determination (Flentje et al., 2010; Kotthaus and Grimmond, 2018), and the retrieval of precipitation parameters (Lolli et al., 2018).

In addition to providing velocity estimates for wind and turbulence, the inclusion of the profile of the attenuated backscatter coefficient is advantageous for Doppler lidar boundary layer classification schemes (Tucker et al., 2009; Harvey et al., 2013; Manninen et al., 2018) by enhancing the discrimination between aerosol, cloud, and precipitation, and it can be used for tracking elevated aerosol plumes (Hannon et al., 1999). The combination of attenuated backscatter profiles from coherent Doppler lidars with other profiling instruments permits additional retrievals; for example, together with a ceilometer (Westbrook et al., 2010b), or with a cloud radar (Träumner et al., 2010), it can yield drizzle drop size and precipitation rate. There is also an advantage to obtaining attenuated backscatter and Doppler velocity measurements in the same measurement volume, since this will simplify the calculation of cloud or aerosol mass fluxes (Engelmann et al., 2008).

Therefore, an accurate profile of the attenuated backscatter coefficient requires confidence in the parameters used to generate the telescope focus function. The parameters may not be known a priori or may differ from what is assumed, and incorrect values can result in artefacts and very large biases in the attenuated backscatter coefficient. We present a methodology for deriving the parameters of the telescope focus function experimentally from co-located Doppler lidar and ceilometer observations, together with the uncertainties in the function parameters. The ceilometer, for which the overlap function is known, provides our reference attenuated backscatter profiles. This methodology is relevant for coherent Doppler lidars designed for meteorological applications with maximum ranges suitable for observing the full extent of the boundary layer and beyond. Note that a calibration constant may still need to be determined and applied after implementing the calculated telescope focus function to

retrieve the profile of the attenuated backscatter coefficient (Westbrook et al., 2010a; Chouza et al., 2015).

The theoretical description of the telescope focus function is outlined in Sect. 2. In Sect. 3, we introduce the instruments and the methodology for deriving the parameters of the telescope focus function experimentally. An iterative least-squares regression using weighted mean square error (MSE) is used to find the best solution for the telescope focus function, where the weights represent the measurement uncertainties in both instruments. The use of long time periods (1 year or more) also provides an estimate of the uncertainties in the parameters for the telescope focus function, which can then be propagated through to uncertainties in the retrieved attenuated backscatter coefficients. The methodology is applied to different instruments in multiple locations in Sect. 4, and the validation of the method is presented in Sect. 5.

## 2 Theory

### 2.1 Telescope focus function

Following Frehlich and Kavaya (1991), the coherent Doppler lidar equation can be expressed as

$$\text{SNR}(R) = \frac{\eta c E}{2h\nu B} \frac{A_e(R)}{R^2} \beta'(R), \quad (1)$$

where SNR is the signal-to-noise ratio, varying as a function of range,  $R$ , from the instrument;  $\beta'$  is the attenuated backscatter coefficient;  $\eta$  is the detector quantum efficiency;  $c$  is the speed of light;  $E$  is the beam energy;  $h$  is Planck's constant;  $\nu$  is the optical frequency;  $B$  is the receiver bandwidth; and  $A_e$  is the effective receiver area.

For a monostatic system emitting a circular Gaussian beam, using a circular aperture, and having matched filters, the effective receiver area is given by Frehlich and Kavaya (1991) and Henderson et al. (2005):

$$A_e(R) = \frac{\pi D^2}{4 \left( 1 + \left( \frac{\pi D^2}{4\lambda R} \right)^2 \left( 1 - \frac{R}{f} \right)^2 + \left( \frac{D}{2\rho_0} \right)^2 \right)}, \quad (2)$$

where  $D$  is the  $1/e^2$  effective diameter of a Gaussian beam,  $\lambda$  is the laser wavelength,  $f$  is the effective focal length of the telescope for the transmitter and receiver, and  $\rho_0$  is a turbulent parameter, also termed transverse field coherence length.

Collecting the range-dependent terms, we obtain a unitless telescope focus function:

$$T_f(R) = \frac{A_e(R)}{R^2}, \quad (3)$$

which is also termed the coherent responsivity (Frehlich and Kavaya, 1991).

The profile of the attenuated backscatter coefficient is then obtained by rearranging Eq. (1):

$$\beta'(R) = \frac{2h\nu B \text{SNR}(R)}{\eta c E T_f(R)}. \quad (4)$$

Figure 1a shows how  $T_f(R)$  depends on the telescope focal length,  $f$ , and Fig. 1b how  $T_f(R)$  depends on  $D$ . Both figures show that the apparent focus – i.e. range to the  $T_f(R)$  maximum – is always closer than  $f$  and that decreasing  $D$  shortens the apparent focus. This makes estimation of the parameters by eye in  $T_f(R)$  prone to errors, since the apparent focus cannot be translated into  $f$  without knowledge of  $D$ .

Figure 1c shows that even if the telescope is focused at infinity, knowledge of the  $D$  is essential to derive attenuated backscatter coefficient profiles. While the gradient of  $T_f(R)$  may be independent of  $D$  at the near and far ranges, the relative magnitude is not, and the potential variation is high in the range of the profile that is commonly of most interest.

## 2.2 Uncertainty in the attenuated backscatter coefficient

Assuming that the parameters  $T_f(R)$  and SNR are independent, and have uncertainties that can be described as Gaussian, the relative random uncertainty in the attenuated backscatter coefficient is

$$\sigma_{\beta'} = \sqrt{\sigma_S^2 + \sigma_{T_f}^2}, \quad (5)$$

where  $\sigma_S$  is the relative uncertainty in the Doppler lidar SNR, and  $\sigma_{T_f}$  is the relative uncertainty in  $T_f(R)$ . An expression for deriving  $\sigma_S$  is given by Manninen et al. (2018), and we describe our method for obtaining  $\sigma_{T_f}$  in Sect. 4.2.

## 3 Application to data

There are three range-dependent unknowns in Eq. (2):  $f$ ,  $D$ , and  $\rho_0$ . We first assume that we can neglect  $\rho_0$  and then describe a method for estimating  $f$  and  $D$ , together with their uncertainties, which can then be propagated to obtain the uncertainty in the attenuated backscatter coefficient. The impact of the  $\rho_0$  parameter is discussed in Sect. 4.3.

### 3.1 Instruments

We used measurements taken from the U.S. Department of Energy Atmospheric Radiation Measurement (ARM, Mather et al., 2016) observatories. We selected five sites with co-located ceilometer and Doppler lidar instruments: Southern Great Plains, US (SGP); tropical west Pacific, Darwin, Australia (Darwin); Barrow, Alaska, US (North Slope of Alaska, NSA); Graciosa, Azores (Graciosa); and Ascension Island, Atlantic, UK (Ascension).

The Doppler lidars operated by ARM comprise both Halo Photonics StreamLine and StreamLine XR versions. These

**Table 1.** Halo Photonics StreamLine and StreamLine XR heterodyne Doppler lidar specifications. Values in parentheses refer to the specification of the Doppler lidar during the first period in Darwin.

Wavelength	1.5 $\mu\text{m}$
Pulse repetition rate	15 kHz
Nyquist velocity	19.8 $\text{m s}^{-1}$
Sampling frequency	50 MHz
Points per range gate	10 (16)
Range resolution	30 m (48 m)
Pulse duration	0.2 $\mu\text{s}$
Divergence	33 $\mu\text{rad}$
Antenna	monostatic fibre-optic coupled

**Table 2.** Vaisala CL31 ceilometer specifications.

Wavelength	910 nm
Pulse repetition rate	5.57 kHz
Range resolution	30 m
Lens diameter	14.5 cm
Divergence	0.75 mrad

are commercially available heterodyne pulsed systems capable of full hemispheric scanning and operated at a temporal resolution of 1–2 s (see Table 1). The focus for the StreamLine version can be set by the operator, whereas the StreamLine XR has the focus set by the manufacturer; however ARM has had some instruments upgraded from their original specification.

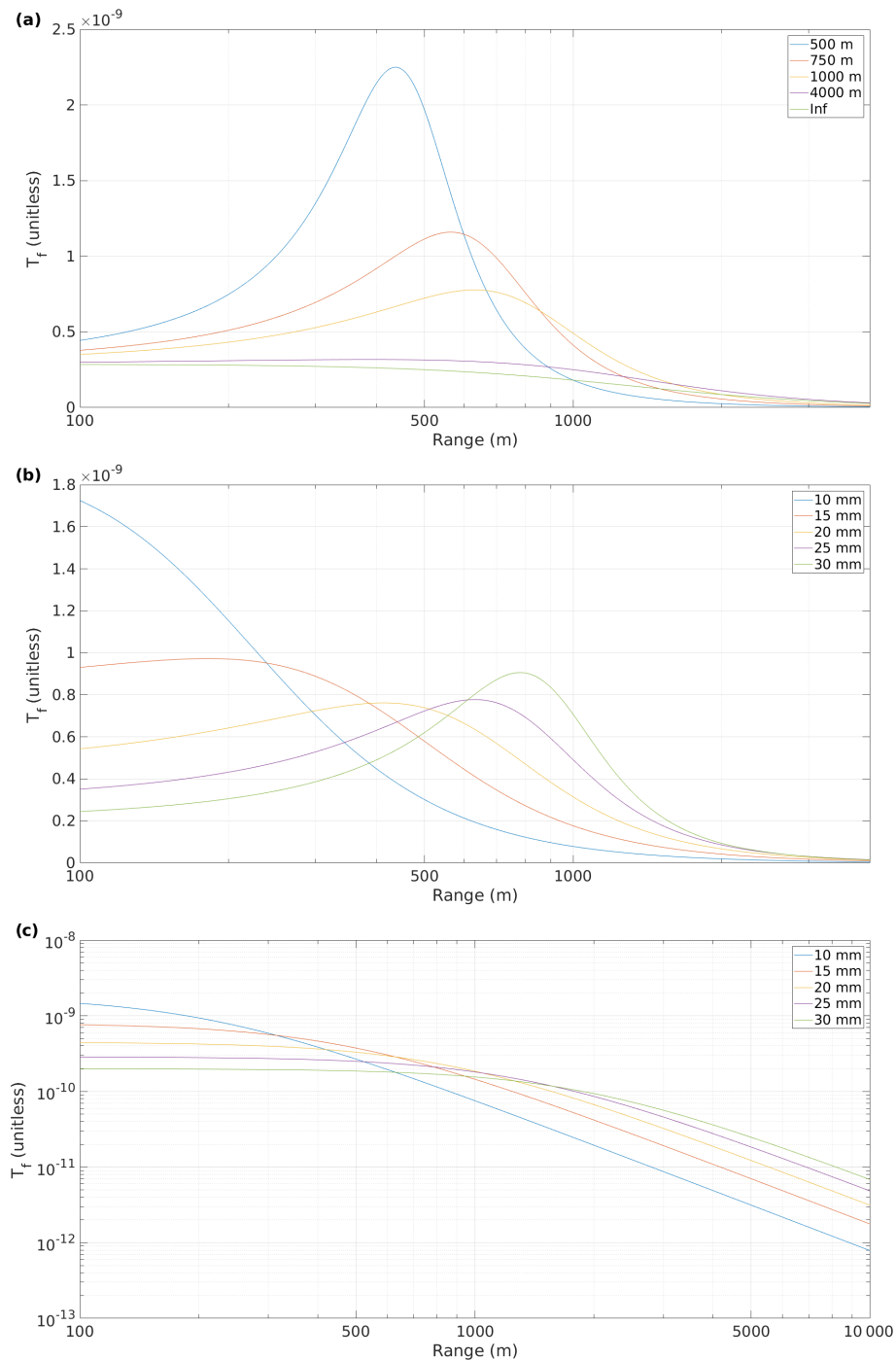
The ceilometer at all sites was a Vaisala CL31 ceilometer, which has a coaxial design and full overlap before 100 m and a temporal resolution of 30 s (more specifications given in Table 2).

## 3.2 Methodology

### 3.2.1 Telescope focus function parameter estimation

The methodology for deriving the parameters of the telescope focus function compares profiles from a co-located Doppler lidar and ceilometer using an iterative least-squares regression to find the best solution. The method follows the process diagram given in Fig. 2.

Before input, the Doppler lidar SNR data had a background correction applied to reduce bias (Manninen et al., 2016). Both ceilometer and Doppler lidar data were averaged to a common 30 min, 30 m vertical resolution grid. If the Doppler lidar vertical resolution was larger than 30 m (as was the case for one period from Darwin), linear interpolation was used to match resolutions. After averaging, data below a minimum threshold of  $-22.2$  dB (Manninen et al., 2018) were discarded. The threshold is based on the expected noise floor for the instruments considered here (Halo Stream-

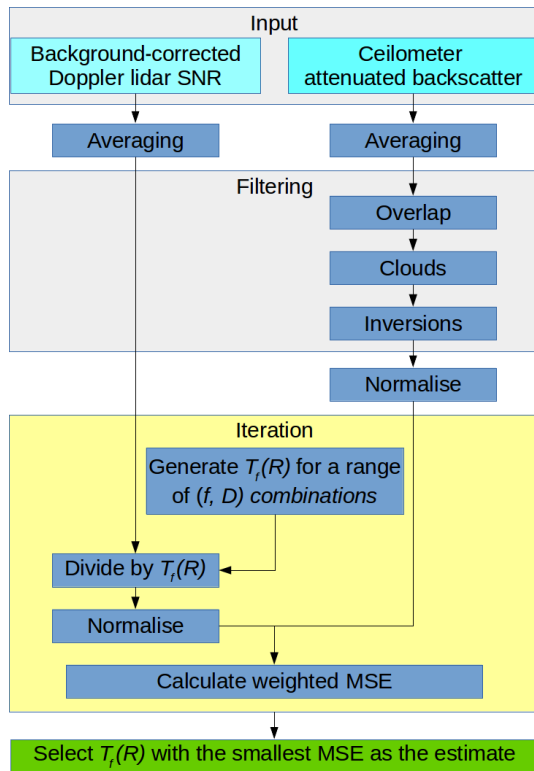


**Figure 1.** Telescope focus functions for (a) varying  $f$  with  $D = 70$  mm, (b) varying  $D$  with  $f = 1000$  m, and (c) varying  $D$  with  $f$  being infinity.

Line and StreamLine XR) and should probably be modified for different instruments.

The data were then filtered to select only those portions of the profiles that are considered reliable for comparison. Ceilometer data below 195 m were discarded to ensure that only data with full overlap were used.

Due to the wavelength difference between the Doppler lidar and the ceilometer, it cannot be assumed that the atmospheric backscattering properties are the same at both wavelengths. However, we are only interested in the profile shape, not the absolute values, so profiles from the Doppler lidar and the ceilometer can be compared as long as they contain



**Figure 2.** Process diagram of the telescope focus function parameter estimation.

only one type of scatterer, and one which can be assumed to be distributed homogeneously throughout the portion of the profile used for comparison. Hence, the portion of a profile selected for comparison should contain only one aerosol layer, no clouds, and no precipitating hydrometeors.

We removed clouds by identifying the range gate 150 m below the cloud base detected by the ceilometer and excluding all data beyond this. Elevated aerosol layers and precipitating hydrometeors were filtered out by identifying layers using a convolution of the ceilometer profile with a haar-wavelet to detect changes in the gradient. The base of the second layer was identified where the gradient was increasing over two range gates, and all data above this were discarded. This process also eliminates noisy profiles with low SNR.

The  $T_f(R)$  parameter estimation is performed on a profile-by-profile basis for each profile where the filtering process leaves eight successive range gates present. From Eq. (4), dividing the Doppler lidar SNR profile with the appropriate  $T_f(R)$  will generate a Doppler lidar attenuated backscatter profile whose shape should match that of the ceilometer attenuated backscatter profile.

We use a brute-force approach to iterate through a range of reasonable  $f$  and  $D$  values, generating a corresponding  $T_f(R)$  and Doppler lidar attenuated backscatter profile for each combination of values. The ceilometer profile and re-

sulting Doppler lidar attenuated backscatter profiles are normalised so that the integral value of the unfiltered portion is unity. We then use a least-squares regression using weighted MSE to find the best solution (smallest MSE), where the weights represent the measurement uncertainties in both instruments.

Collecting results over many profiles results in a bivariate distribution; the peak of this distribution is chosen as the best estimate of  $f$  and  $D$ , and hence the best estimate of  $T_f(R)$ , using Eq. (3).

### 3.2.2 Outlier removal

Occasionally, data of poor quality pass the filtering step in Fig. 2. The most common issues are noisy ceilometer data and a bias in the Doppler lidar SNR profiles. If not screened, these occasional profiles result in significantly altered  $T_f(R)$  estimation. Any noise in the ceilometer data is magnified by the profile length often being relatively short, and hence large uncertainty in even a single range gate can skew the regression. Doppler lidar SNR bias will impact the normalisation process, changing the  $T_f(R)$  selected by the method due to the now incorrect profile shape. Due to the non-linearity of the  $T_f(R)$  parameter estimation process, these issues result in regression solutions wildly inconsistent with the estimates based on good data. These outliers, which do not fall within the normal uncertainty observed in good data, are then removed from the bivariate distribution of solutions before calculating the uncertainty estimates.

We used the median absolute deviation, MAD (Huber and Ronchetti, 2009; Leys et al., 2013), to distinguish outliers in the bivariate distribution of estimated  $f$  and  $D$ . MAD can be calculated using

$$MAD = b \text{ med}\{|x_i - \text{med}\{x_i\}|\}, \quad (6)$$

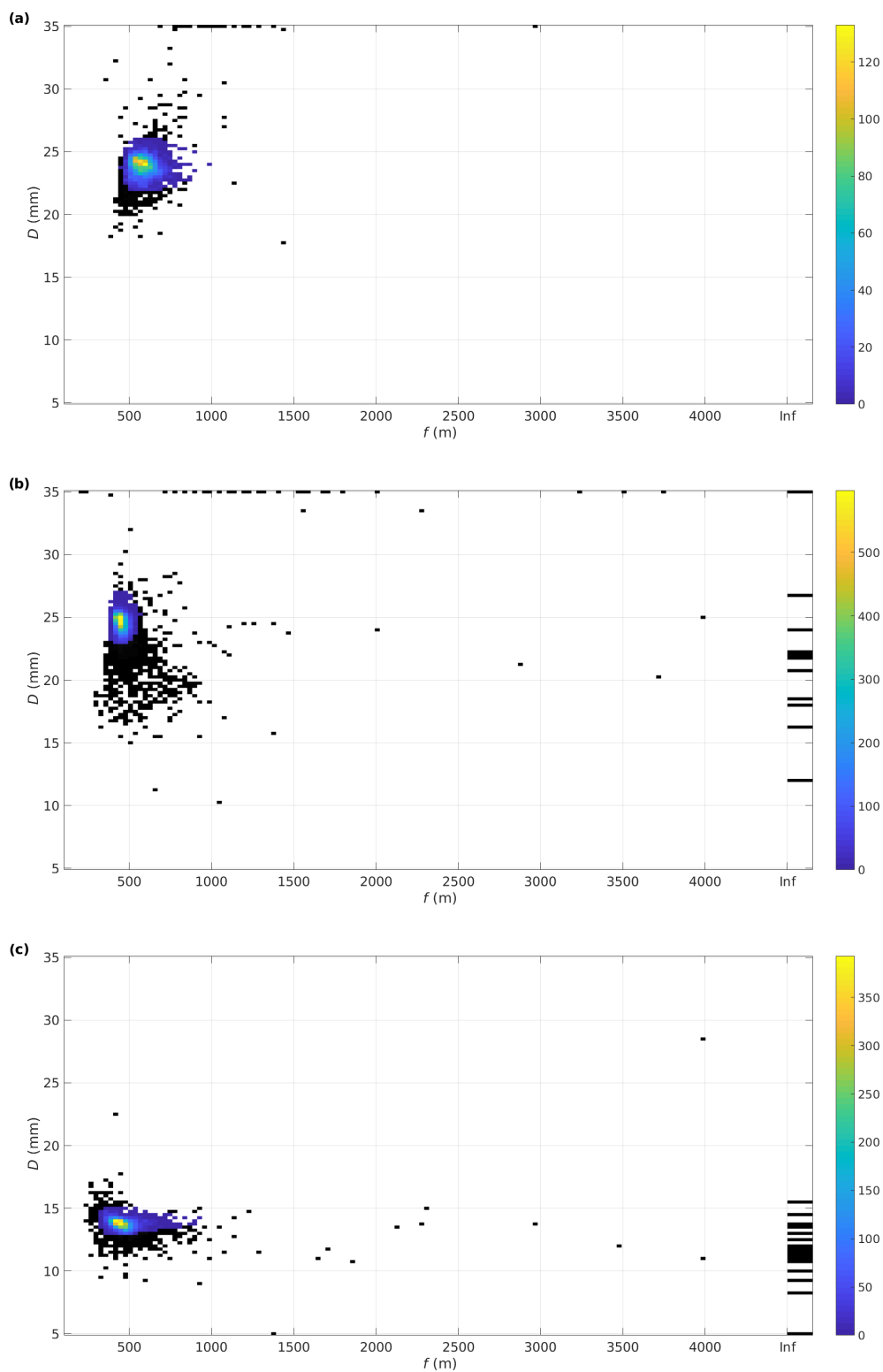
where  $b = 1.4826$  when the distribution excluding the outliers is normal. However, the distribution of  $f$  and  $D$  may not meet this criterion due to the non-linearity of  $T_f(R)$  and the computational  $T_f(R)$  estimation process. We expect the distributions of  $D$  and  $f^{-2}$  to be close to normal and will use  $f^{-2}$  rather than  $f$  to determine outliers. Additionally, the peak of the bivariate distribution may not always coincide with the medians of the univariate  $D$  and  $f^{-2}$  distributions, and, hence, we use a modified form of Eq. (6),

$$MAD = b \text{ med}\{|x_i - \text{peak}\{x_i, y_i\}|\}. \quad (7)$$

We selected three MADs as the threshold for flagging outliers:

$$\sqrt{\left(\frac{f_i^{-2} - \text{med}\{f_i^{-2}\}}{MAD_{f^{-2}}}\right)^2 + \left(\frac{D_i - \text{med}\{D_i\}}{MAD_D}\right)^2} \geq 3. \quad (8)$$

Assuming the distribution excluding the outliers to be normal, three MADs correspond to 3 standard deviations of the underlying distribution. In cases where  $f$  is at infinity, all estimates with a finite  $f$  will be flagged as outliers.



**Figure 3.** Distributions of  $T_f(R)$  parameter estimates from (a) Darwin from 21 June 2011 to 22 July 2012, (b) SGP from 1 January 2015 to 2 May 2016, and (c) SGP from 3 May 2016 to 5 June 2017. Outliers filtered using  $\text{MAD} \geq 3$  are marked in black.

**Table 3.** Best estimates of  $f$  and  $D$  together with their uncertainties for Doppler lidars at five ARM sites. Total estimates refers to the number of ceilometer and Doppler lidar profiles suitable for comparison. Good estimates refers to the number of estimates remaining after outlier filtering with MAD.

Location	Period	$f$	$D$	Available profiles	Total estimates	Good estimates
Ascension	20160906–20170930	$550 \pm 34$ m	$25.3 \pm 0.5$ mm	18 586	62	56
Darwin	20110621–20120722	$590 \pm 62$ m	$24.0 \pm 0.7$ mm	18 988	3684	3528
Darwin	20120921–20140626	$545 \pm 53$ m	$25.0 \pm 0.8$ mm	30 836	5046	4878
Graciosa	20150124–20161114	$625 \pm 80$ m	$23.5 \pm 0.7$ mm	31 124	3737	3161
NSA	20140730–20171231	Infinity	$11.8 \pm 1.5$ mm	56 832	1132	589
SGP	20150101–20160502	$440 \pm 29$ m	$25.0 \pm 0.7$ mm	22 916	9198	8426
SGP	20160503–20170605	$425 \pm 74$ m	$14.0 \pm 0.4$ mm	14 212	5814	5337

## 4 Results

### 4.1 Parameter estimation

We applied the  $T_f(R)$  estimation method to Doppler lidars at five ARM atmospheric observatories. Figure 3a shows the distribution of  $f$  and  $D$  calculated for the Doppler lidar operating at Darwin in northern Australia between 21 June 2011 and 22 July 2012. This Doppler lidar is a StreamLine and the distribution of  $f$  is positively skewed, as explained in Sect. 3.2.2. The distribution displays a slightly wider peak than expected for a normal distribution.

Figure 3b shows the distribution of  $f$  and  $D$  for the StreamLine Doppler lidar operating at SGP from 1 January 2015 to 2 May 2016. The distribution close to the peak is really tight, while the outliers have substantial spread. Many of the poor estimates responsible for the outlier spread occur during January and February in both years, while for the rest of the period the estimates are remarkably consistent. On 3 May 2016 the Doppler lidar at SGP was changed to a StreamLine XR, and Fig. 3c shows the distribution of  $f$  and  $D$  from 3 May 2016 to 5 June 2017. The change in instrument version, from StreamLine to StreamLine XR, is clearly seen in the change in  $D$ , whereas the best estimate for  $f$  did not change. However, inspecting the data by eye would suggest a significantly shorter apparent focus, and the  $T_f(R)$  calculated using the best estimates for  $f$  and  $D$  also exhibits a significantly shorter apparent focus. Consequently, the StreamLine XR in SGP has been noted to suffer from poor SNR at the boundary layer top.

The bivariate distributions of  $f$  and  $D$  show notable variations in how tight they are around the peak, likely a result of differences in data quality between the instruments. The best estimates of  $f$  and  $D$  and their uncertainties for all sites are presented in Table 3. The Doppler lidar measurements at Darwin were split into two periods, as there was a 2-month break in the measurements between these two periods. We performed the  $T_f(R)$  parameter estimation separately for both periods. The best estimates from these periods differ from each other, which is expected as some adjustments were made to the instrument. The telescope focal

length for this instrument is directly adjustable by any operator, while the beam diameter is set by the manufacturer and is not modifiable by the operator. We note that the  $D$  estimates are the same for these two periods within the margin of error calculated.

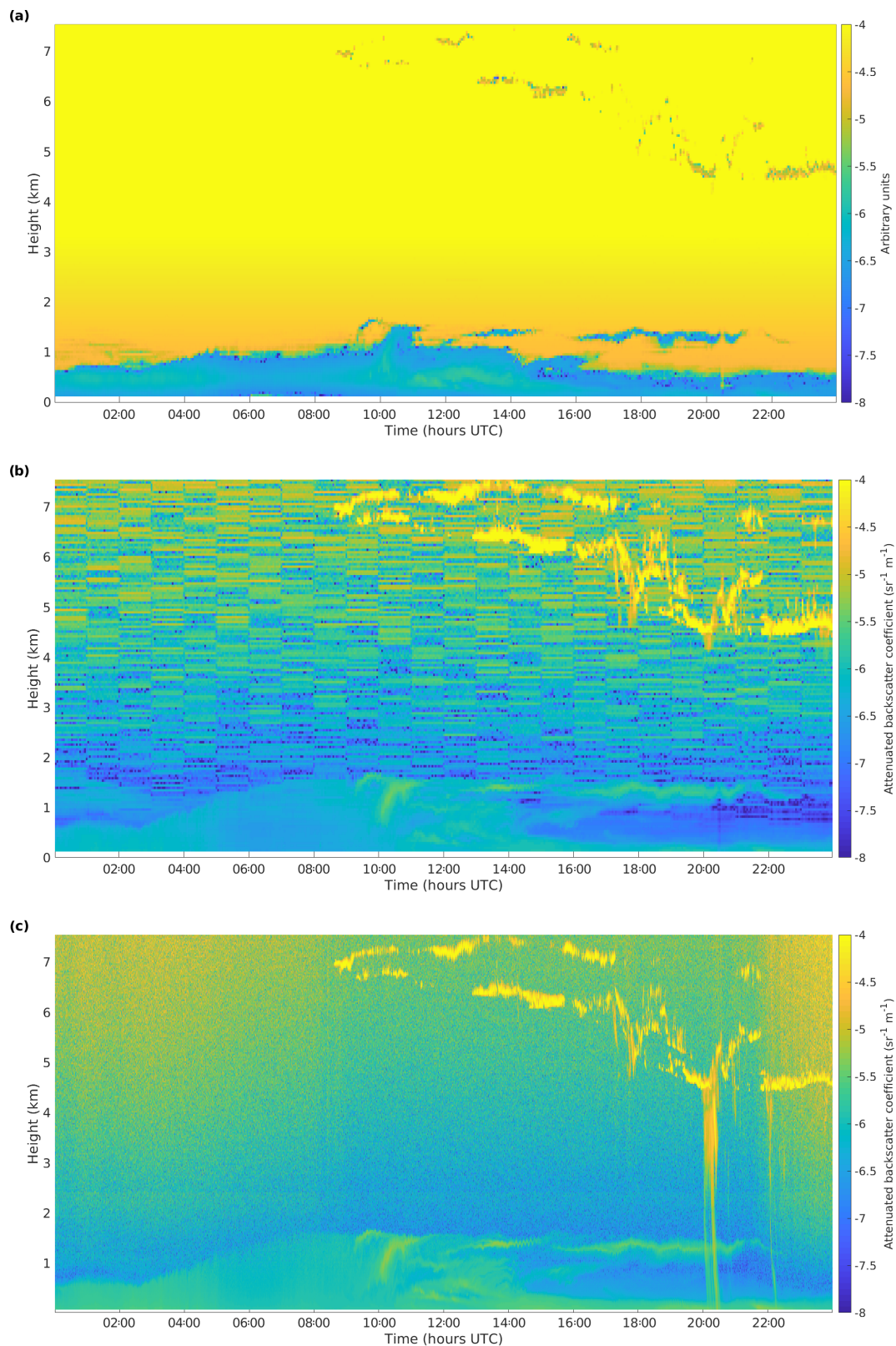
For the sites and instruments selected here, only the Doppler lidar at NSA had  $f$  set to infinity. In fact, all StreamLine lidars have  $D$  in the vicinity of 25 mm, whereas  $D$  for the StreamLine XR versions is about half this. Nevertheless, the variation between instruments of the same version is not negligible and should be taken into account when calculating  $T_f(R)$  and then attenuated backscatter.

The final step to obtain attenuated backscatter profiles is to apply a calibration constant, which can be achieved using the liquid cloud calibration method (Westbrook et al., 2010a; O'Connor et al., 2004).

The parameters  $f$  and  $D$  calculated for period 1 in Darwin have been used to derive  $T_f(R)$  and the results applied in Fig. 4. This shows the utility of the method, able to provide reliable Doppler lidar attenuated backscatter profiles in Fig. 4b that show no overcorrection below 1 km and display similar in-cloud values to the ceilometer in Fig. 4c. It is expected that the aerosol attenuated backscatter coefficients will differ due to the different scattering properties of aerosol at the different wavelengths; the scattering properties of cloud droplets remain similar at the two wavelengths (O'Connor et al., 2004; Westbrook et al., 2010a).

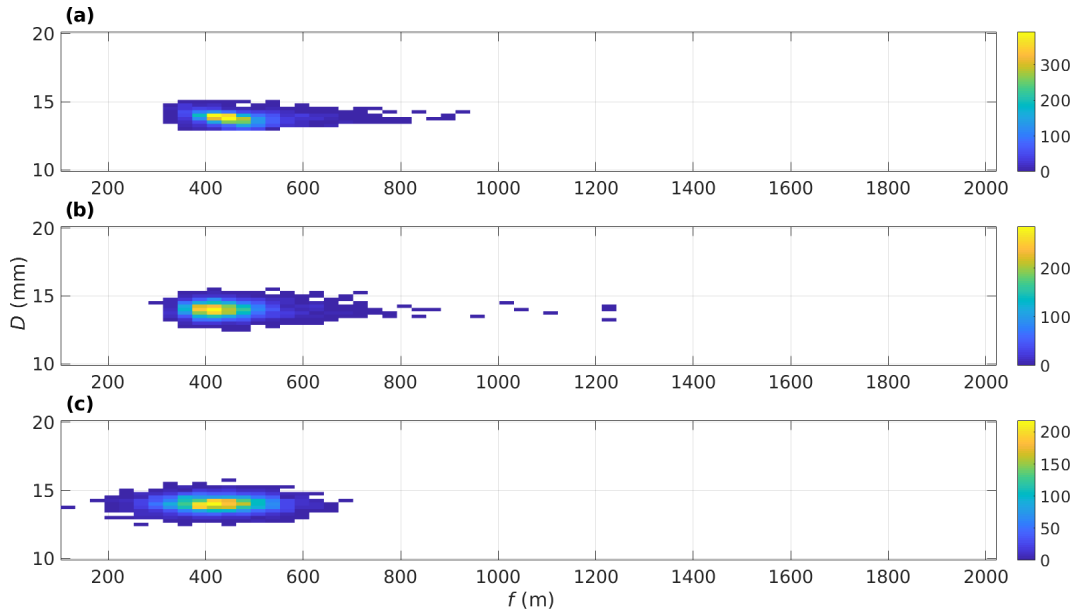
### 4.2 Uncertainty

A computational method was used to calculate the uncertainty in the estimated  $T_f(R)$  as it is a non-linear function of  $f$  and  $D$ . We used Monte Carlo simulation (MCS) (Morgan and Henrion, 1990), where a distribution of input values is fed into a model, here the effective receiver area Eq. (2), and the uncertainty is obtained from the distribution of the output. The input values can be created either from observed statistics or by bootstrapping, i.e. resampling the data. We created three different sets of input values for our MCS:



**Figure 4.** (a) Doppler lidar attenuated backscatter coefficient assuming a generic  $T_f(R)$ , (b) corrected Doppler lidar attenuated backscatter coefficient, and (c) ceilometer attenuated backscatter coefficient, from Darwin on 28 May 2012.





**Figure 5.** Distributions of the Monte Carlo simulation (MCS) input values used for calculating  $\sigma_{T_f}(R)$ . Values are obtained from (a) resampling, (b) assuming  $N(f^{-2}, D)$ , and (c) assuming  $N(f, D)$ . All distributions contain 5337 samples.

**Table 4.**  $\sigma_{T_f}(R)$  uncertainty envelopes generated using MCS with three different sets of input values.

Location	Period	Resampling	$N(f^{-2}, D)$	$N(f, D)$
Ascension	20160906–20170930	0.15	0.14	0.16
Darwin	20110621–20120722	0.21	0.23	0.25
Darwin	20120921–20140626	0.23	0.21	0.28
Graciosa	20150124–20161114	0.23	0.19	0.30
NSA	20140730–20171231	0.32	0.32	0.30
SGP	20150101–20160502	0.20	0.18	0.22
SGP	20160503–20170605	0.13	0.13	0.23

1. Resampling the individual estimates of  $f$  and  $D$  provided directly by the  $T_f(R)$  estimation method (i.e. those displayed in Fig. 3) after excluding outliers.
2. Generating the values from the statistics presented in Table 3, assuming that  $D$  and  $f^{-2}$  are normally distributed and independent,  $N(f^{-2}, D)$ .
3. Generating the values from statistics presented in Table 3, assuming  $D$  and  $f$  to be normally distributed and independent,  $N(f, D)$ .

For each set of input values, the relative uncertainty in  $T_f(R)$  is calculated as

$$\sigma_{T_f}(R) = \frac{\sigma_{T_\phi}(R)}{T_f(R)}, \tag{9}$$

where  $\sigma_{T_\phi}(R)$  is expressed in terms of the mean-squared deviation of the MCS-simulated telescope focus function,

$T_\phi(R)$ , from the best estimate of  $T_f(R)$ ,

$$\sigma_{T_\phi}(R) = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (T_{\phi_i}(R) - T_f(R))^2}, \tag{10}$$

to avoid underestimating the uncertainty resulting from the asymmetry in  $T_\phi(R)$ . This also allows us to estimate the impact of refractive turbulence on the uncertainty estimate.

Examples of the three input parameter distributions are presented in Fig. 5. Resampling (Fig. 5a) is the most accurate method as it does not require assumptions about the parameter distributions and their independence. We recommend resampling as the primary method for uncertainty calculation. Using the  $N(f^{-2}, D)$  distribution (Fig. 5b) produces a set of input values that appear to be a reasonable approximation, except that the distribution is not as tight around the peak. Using the  $N(f, D)$  distribution (Fig. 5c) produces a set of input values that tend to overemphasise shorter values of  $f$  and underemphasise higher values. We also note that the central bin of the resampled distribution contains 50 % more samples

than the central bin of the statistically generated distributions do. We presume that this is a consequence of the variation in SNR not being necessarily normally distributed.

Figure 6a displays  $\sigma_{T_f}(R)$  for Darwin showing the range dependence of the uncertainty, with much larger uncertainties for ranges close to either side of the focus ( $f = 590$  m). The profile of uncertainties obtained with each set of MCS input values exhibit a similar shape, with  $N(f^{-2}, D)$  being closer to resampling than  $N(f, D)$  in the near field.

Figure 6b displays  $\sigma_{T_f}(R)$  for the Doppler lidar at NSA, which has  $f$  set to infinity; therefore  $\sigma_{T_f}(R)$  is only dependent on the uncertainty in  $D$ . Note the reduced uncertainties around 200–400 m, which are expected when examining Fig. 1c.

The largest value of  $\sigma_{T_f}(R)$  provides the uncertainty envelope value for each site, which is presented in Table 4. Resampling provides values ranging from 0.12 for the updated instrument at SGP to 0.32 at NSA. MCS values created using  $N(f^{-2}, D)$  provided similar values, whereas MCS using  $N(f, D)$  often provided much larger uncertainties.

### 4.3 Impact of refractive turbulence

So far we have neglected the potential impact of turbulence on  $T_f(R)$  arising from the refractive turbulent parameter,  $\rho_0$ , in Eq. (2). An expression for  $\rho_0$  is given in Frehlich and Kavaya (1991),

$$\rho_0(R) = \left[ Hk^2 \int_0^R C_n^2(z) (1 - z/R)^{5/3} dz \right]^{-3/5}, \quad (11)$$

where  $H = 2.914383$ ,  $k = 2\pi/\lambda$ , and  $C_n^2(z)$  is the refractive turbulence at range  $z$ . We chose three profiles with constant  $C_n^2(z)$  and a realistic vertical profile based on the most turbulent case presented by Roadcap and Tracy (2009). Figure 7 shows the impact that these different profiles have on  $T_f(R)$  and the resulting resampling calculation of  $\sigma_{T_f}(R)$  for two Doppler lidar instruments with different  $T_f(R)$ . Values of  $C_n^2$  up to  $10^{-14} \text{ m}^{-2/3}$  have negligible impact on  $T_f(R)$ , and even the realistic profile only showed a slight increase in  $\sigma_{T_f}(R)$  for the instrument with a focus set closer than infinity. This suggests that the impact of turbulence can be safely neglected for low values of  $C_n^2$ , and for most applications, it can also be neglected when operating in the vertical. Hence, turbulence has no significant impact on the methodology described here for deriving the parameters  $f$  and  $D$  and their uncertainties from vertical profiles but can be included for completeness.

However, it is clear that the turbulent impact should not be ignored when measuring at low elevation angles close to the horizon, where a profile similar to  $C_n^2 = 10^{-13} \text{ m}^{-2/3}$  may be possible. Figure 7 shows that such a profile has a major impact on  $T_f(R)$ , especially in the far range. In these cases, the parameters  $f$  and  $D$  obtained from vertical mea-

surements are still applicable, but  $\rho_0(R)$  must also be calculated or estimated in order to derive the profile of attenuated backscatter,  $\beta'(R)$ .

## 5 Validation

The liquid cloud calibration method (O'Connor et al., 2004; Westbrook et al., 2010a) is used to determine a calibration constant by integrating attenuated backscatter profiles containing fully attenuating liquid clouds, which have a well-constrained apparent lidar ratio,  $\eta S$ , where  $\eta$  is a multiple scattering factor and  $S$  is the lidar ratio. In the absence of multiple scattering,  $\eta S$  can be assumed to be independent of the height of the cloud.

This calibration method can be used to evaluate the estimated  $T_f(R)$  for Doppler lidar by checking whether the attenuated backscatter profiles obtained for the Doppler lidar after applying  $T_f(R)$  indeed provide similar  $\eta S$  values for liquid clouds at different heights.

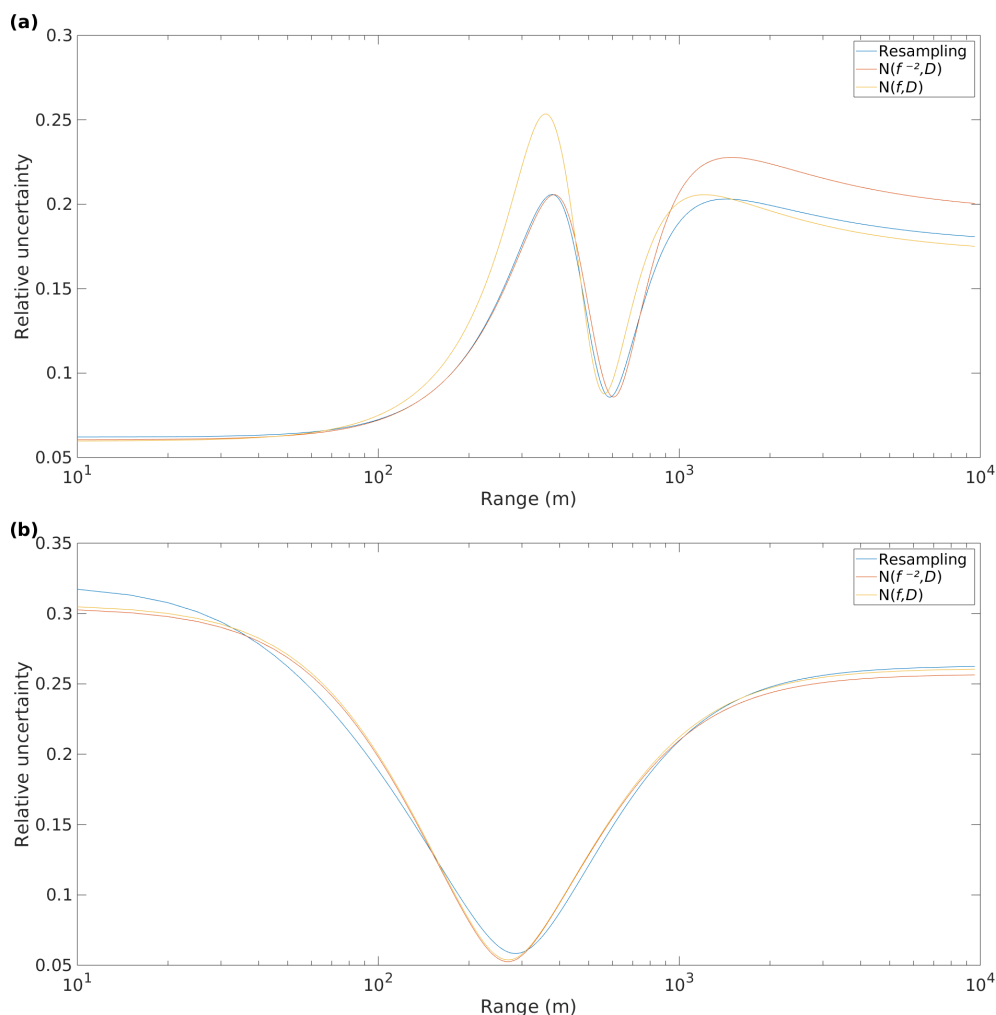
Figure 8 shows examples of Doppler lidar attenuated backscatter profiles after calibration and the derived apparent lidar ratio at two sites, Darwin and NSA. These sites have different values of  $f$ : Darwin has  $f = 590$  m and NSA has  $f$  set to infinity. For both cases, liquid clouds are present throughout the day, with altitudes varying from 2 to 6 km. When fully attenuating liquid clouds are present, the apparent lidar ratio is close to the expected value of 20 sr, regardless of the height of the cloud, thus confirming that the method of estimating  $T_f(R)$  is valid.

### Limitations

Table 3 shows that the proportion of data that can be used for the  $T_f(R)$  parameter estimation varies significantly from site to site. Over a third of the available profiles from SGP are used, whereas only 0.3% pass the filtering for Ascension. The lack of suitable profiles at Ascension is explained by the almost constant low cloud cover at this site, with very few profiles having a sufficient number of successive range gates.

Data quality is also a limiting factor, so at sites with very low aerosol optical depth, AOD, such as NSA, the Doppler lidar SNR decreases so rapidly that again there are few profiles having a sufficient number of successive range gates. Low AOD also impacts the performance of the ceilometer, with 48% of the estimates at NSA discarded as outliers even after the initial filtering was performed. While the outlier removal can separate the good and the poor estimates, the largest uncertainty in  $D$  was at NSA. We attempted to perform the  $T_f(R)$  parameter estimation on a Doppler lidar from an ARM campaign in Cape Cod but could not obtain reliable estimates due to the low SNR of the ceilometer data.

The  $T_f(R)$  parameter estimation method is suitable only in situations where there is minimal difference in atmospheric extinction within the aerosol layer between the two instru-

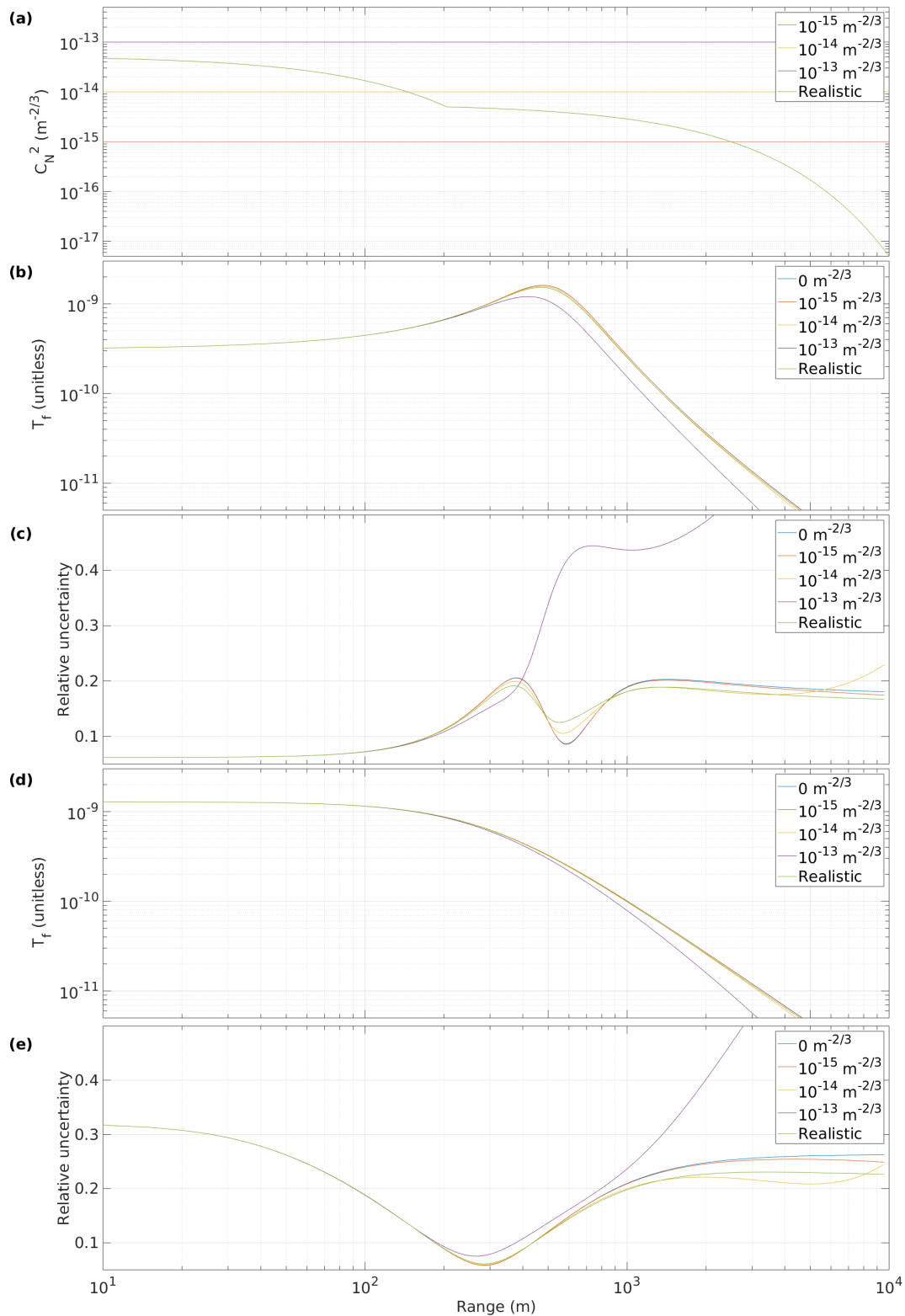


**Figure 6.** Relative telescope focus function uncertainties,  $\sigma_{T_f}(R)$ , generated using MCS with three different sets of input values for (a) Darwin from 21 June 2011 to 22 July 2012 and (b) NSA from 30 July 2014 to 31 December 2017.

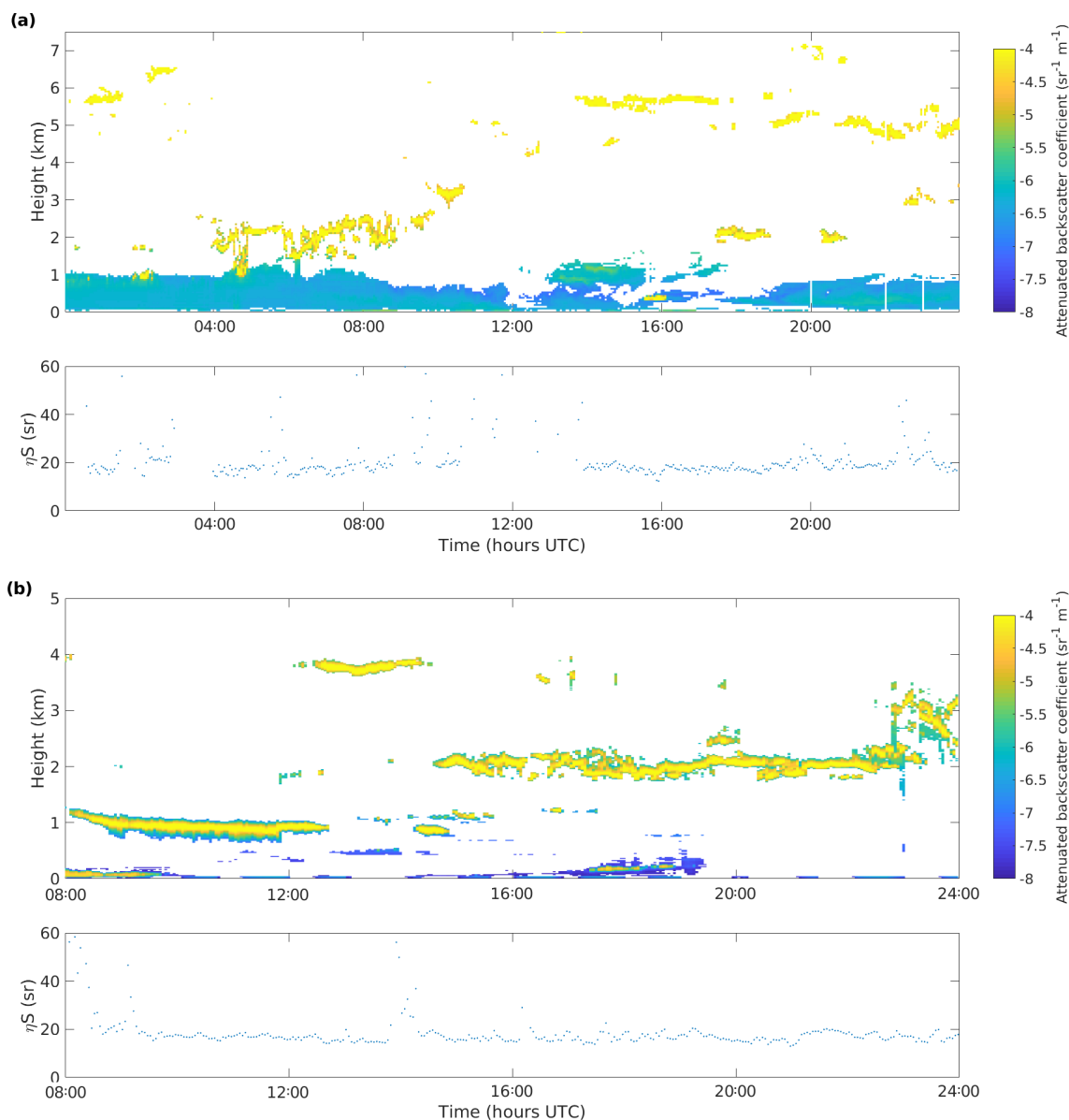
ment wavelengths of 910 and 1500 nm. Using AERONET (Aerosol Robotic Network) AOD measurements co-located at the ARM atmospheric observatories, the median difference in AOD at 870 and 1640 nm varied between 0.016 and 0.027, which should correspond closely to what might be expected for the difference between the ceilometer and Doppler lidar. Very occasional periods of notable AOD differences were observed at some sites, but including these profiles in time series extending beyond a year will have negligible impact on the  $T_f(R)$  parameter best estimate. However, there were breaks in the AOD measurements, and some periods experiencing a significant differential extinction may have gone unnoticed. An additional filter using AERONET AOD measurements to remove profiles experiencing significant differential extinction could be included in Fig. 2 for those sites where this may be an issue.

## 6 Conclusions

We have developed a method for deriving the telescope focus function and its uncertainty for pulsed heterodyne Doppler lidars, and we applied the method to Halo Photonics Stream-Line and XR Doppler lidars. The method compares profiles of the Doppler lidar SNR to profiles of the attenuated backscatter coefficient from a co-located ceilometer, producing estimates for two parameters of the  $T_f(R)$ : the effective focal length for the telescope,  $f$ ; and the  $1/e^2$  effective diameter of a Gaussian beam,  $D$ . This method was developed because it also provides uncertainties in  $f$ ,  $D$ , and  $T_f(R)$ , which are necessary for quantitative use of the subsequently derived attenuated backscatter profiles. The method can be used to check the manufacturer specifications for these parameters, calculate them if not known, and also check their stability over time. The method does not necessarily require a permanently co-located ceilometer as the estimates of  $f$  and  $D$  can be made from a short time series with a co-located



**Figure 7.** Impact of turbulent parameter,  $\rho_0$ , on the telescope focus function,  $T_f(R)$ , and relative uncertainties,  $\sigma_{T_f}(R)$ , for different  $C_n^2$  profiles. **(a)** Selected profiles of  $C_n^2$  with range; **(b)**  $T_f(R)$  and **(c)**  $\sigma_{T_f}(R)$  for Darwin from 21 June 2011 to 22 July 2012; **(d)**  $T_f(R)$  and **(e)**  $\sigma_{T_f}(R)$  for NSA from 30 July 2014 to 31 December 2017.



**Figure 8.** Doppler lidar attenuated backscatter coefficient and apparent lidar ratio,  $\eta S$ , from (a) Darwin on 8 May 2012 and (b) NSA on 20 August 2014.

ceilometer; the Doppler lidar can then be moved to another location without a co-located ceilometer using the  $T_f(R)$  estimated previously, assuming that there has been no modification to the instrument parameters.

The method was applied to data from Doppler lidars with different configurations deployed at five ARM sites. Relative uncertainties in  $f$  for these instruments ranged from 6 % to 17 %, with the median uncertainty being 10 %; the relative uncertainty in  $D$  ranged from 2 % to 12 %, with a median of 3 %. The uncertainty in  $T_f(R)$  was calculated using Monte Carlo simulation, using three methods to prepare the input values. We recommend the direct resampling method, but reasonable results were obtained using statistically de-

rived input values assuming a normal distribution. The envelope of relative uncertainties in  $T_f(R)$  ranged from 13 % to 32 % and depended on both the instrument configuration and the instrument location. We also show that, even for a Doppler lidar with the focus set at infinity, uncertainty remains in estimating  $T_f(R)$  arising from the uncertainty in  $D$ . The method was validated by calculating the apparent lidar ratio of fully attenuating liquid clouds from  $T_f(R)$  corrected profiles of Doppler lidar attenuated backscatter.

The impact of turbulence on  $T_f(R)$  was also investigated and was found to have no significant impact on the methodology described here for deriving the parameters  $f$  and  $D$  and their uncertainties from vertical profiles. However, the tur-

bulent impact should not be ignored when measuring at low elevation angles close to the horizon, as it can modify  $T_f(R)$  considerably, especially in the far range. In these cases, the parameters  $f$  and  $D$  obtained from vertical measurements are still applicable, but the turbulent contribution to  $T_f(R)$  should be included when deriving the attenuated backscatter coefficient.

The  $T_f(R)$  estimation method is suitable only for conditions where the differential extinction at the two wavelengths of the Doppler lidar and the ceilometer is small, which can be identified, for example, using AOD from co-located AERONET observations.

*Data availability.* The ground-based data used in this article are generated by the Atmospheric Radiation Measurement (ARM) user facility and are made available from the ARM Data Discovery website as follows: ceilometer data (CEIL, 2002) from <https://doi.org/10.5439/1181954> and Doppler lidar data (DLFPT, 2010) from <https://doi.org/10.5439/10251850>.

*Author contributions.* EJO supervised and conceptualised the project. PP created the methodology and performed the analysis. AJM provided the processed input data. POA and AJM contributed to the development of the methodology. PP and EJO wrote the paper.

*Competing interests.* The authors declare that they have no conflict of interest.

*Special issue statement.* This article is part of the special issue “Tropospheric profiling (ISTP11) (AMT/ACP inter-journal SI)”. It is a result of the 11th edition of the International Symposium on Tropospheric Profiling (ISTP), Toulouse, France, 20–24 May 2019.

*Acknowledgements.* The Doppler lidar and ceilometer data used in this study were obtained from the Atmospheric Radiation Measurement (ARM) user facility, managed by the Office of Biological and Environmental Research for the U.S. Department of Energy Office of Science.

*Financial support.* This research has been supported by the U.S. Department of Energy’s Atmospheric System Research (ASR), an Office of Science, Office of Biological and Environmental Research (BER) programme (grant no. DE-SC0017338).

Open access funding provided by Helsinki University Library.

*Review statement.* This paper was edited by Pauline Martinet and reviewed by two anonymous referees.

## References

- CEIL: Atmospheric Radiation Measurement (ARM) user facility, updated hourly. Ceilometer (CEIL). 2011-06-21 to 2017-12-21, ARM Mobile Facility (ASI) Ascension Island, South Atlantic Ocean; AMF1 (M1), Eastern North Atlantic (ENA) Graciosa Island, Azores, Portugal (C1), North Slope Alaska (NSA) Central Facility, Barrow AK (C1), Southern Great Plains (SGP) Central Facility, Lamont, OK (C1), Tropical Western Pacific (TWP) Central Facility, Darwin, Australia (C3), compiled by: Morris, V., Flynn, C., and Ermold, B., ARM Data Center, <https://doi.org/10.5439/1181954>, 2002.
- Chouza, F., Reitebuch, O., Groß, S., Rahm, S., Freudenthaler, V., Toledano, C., and Weinzierl, B.: Retrieval of aerosol backscatter and extinction from airborne coherent Doppler wind lidar measurements, *Atmos. Meas. Tech.*, 8, 2909–2926, <https://doi.org/10.5194/amt-8-2909-2015>, 2015.
- DLFPT: Atmospheric Radiation Measurement (ARM) user facility, updated hourly. Doppler lidar fixed-pointing (DLFPT). 2011-06-21 to 2017-12-21, ARM Mobile Facility (ASI) Ascension Island, South Atlantic Ocean; AMF1 (M1), Eastern North Atlantic (ENA) Graciosa Island, Azores, Portugal (C1), North Slope Alaska (NSA) Central Facility, Barrow AK (C1), Southern Great Plains (SGP) Central Facility, Lamont, OK (C1), Tropical Western Pacific (TWP) Central Facility, Darwin, Australia (C3), compiled by: Newsom, R. and Krishnamurthy, R., ARM Data Center, <https://doi.org/10.5439/1025185>, 2010.
- Engelmann, R., Wandinger, U., Ansmann, A., Müller, D., Žeromskis, E., Althausen, D., and Wehner, B.: Lidar observations of the vertical aerosol flux in the planetary boundary layer, *J. Atmos. Ocean. Tech.*, 25, 1296–1306, <https://doi.org/10.1175/2007JTECHA967.1>, 2008.
- Flentje, H., Heese, B., Reichardt, J., and Thomas, W.: Aerosol profiling using the ceilometer network of the German Meteorological Service, *Atmos. Meas. Tech. Discuss.*, 3, 3643–3673, <https://doi.org/10.5194/amtd-3-3643-2010>, 2010.
- Frehlich, R. G. and Kavaya, M. J.: Coherent laser radar performance for general atmospheric refractive turbulence, *Appl. Optics*, 30, 5325–5352, <https://doi.org/10.1364/AO.30.005325>, 1991.
- Hannon, S. M., Thomson, J. A. L., and Smith, D. D.: Plume detection and tracking using Doppler lidar aerosol and wind data, in: *Application of Lidar to Current Atmospheric Topics III*, edited by: Sedlacek III, A. J. and Fischer, K. W., 3757, 28–39, International Society for Optics and Photonics, SPIE, <https://doi.org/10.1117/12.366435>, 1999.
- Harvey, N. J., Hogan, R. J., and Dacre, H. F.: A method to diagnose boundary-layer type using Doppler lidar, *Q. J. Roy. Meteor. Soc.*, 139, 1681–1693, <https://doi.org/10.1002/qj.2068>, 2013.
- Henderson, S. W., Gatt, P., Rees, D., and Huffaker, R. M.: Wind Lidar, in: *Laser Remote Sensing*, edited by: Fujii, T. and Fukuchi, T., 469–722, CRC Taylor and Francis, 2005.
- Hill, C.: Coherent focused lidars for Doppler sensing of aerosols and wind, *Remote Sens.*, 10, 466, <https://doi.org/10.3390/rs10030466>, 2018.
- Hirsikko, A., O’Connor, E. J., Komppula, M., Korhonen, K., Pfüller, A., Giannakaki, E., Wood, C. R., Bauer-Pfundstein, M., Poikonen, A., Karppinen, T., Lonka, H., Kurri, M., Heinonen, J., Moisseev, D., Asmi, E., Aaltonen, V., Nordbo, A., Rodriguez, E., Lihavainen, H., Laaksonen, A., Lehtinen, K. E. J., Laurila, T., Petäjä, T., Kulmala, M., and Viisanen, Y.: Observing wind,

- aerosol particles, cloud and precipitation: Finland's new ground-based remote-sensing network, *Atmos. Meas. Tech.*, 7, 1351–1375, <https://doi.org/10.5194/amt-7-1351-2014>, 2014.
- Hogan, R. J., Illingworth, A. J., O'Connor, E. J., and Polares Baptista, J. P. V.: Characteristics of mixed-phase clouds. II: A climatology from ground-based lidar, *Q. J. Roy. Meteor. Soc.*, 129, 2117–2134, <https://doi.org/10.1256/qj.01.209>, 2003.
- Hu, Q., Rodrigo, P. J., Iversen, T. F. Q., and Pedersen, C.: Investigation of spherical aberration effects on coherent lidar performance, *Opt. Express*, 21, 25670–25676, <https://doi.org/10.1364/OE.21.025670>, 2013.
- Huber, P. J. and Ronchetti, E.: *Robust Statistics*, chap. 5, 105–123, John Wiley & Sons, Ltd, <https://doi.org/10.1002/9780470434697.ch5>, 2009.
- Illingworth, A. J., Cimini, D., Haeffele, A., Haefelin, M., Hervo, M., Kotthaus, S., Löhnert, U., Martinet, P., Mattis, I., O'Connor, E. J., and Potthast, R.: How can existing ground-based profiling instruments improve European weather forecasts?, *B. Am. Meteorol. Soc.*, 100, 605–619, <https://doi.org/10.1175/BAMS-D-17-0231.1>, 2019.
- Kotthaus, S. and Grimmond, C. S. B.: Atmospheric boundary-layer characteristics from ceilometer measurements. Part I: A new method to track mixed layer height and classify clouds, *Q. J. Roy. Meteor. Soc.*, 144, 1525–1538, <https://doi.org/10.1002/qj.3299>, 2018.
- Leys, C., Ley, C., Klein, O., Bernard, P., and Licata, L.: Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median, *J. Exp. Soc. Psychol.*, 49, 764–766, <https://doi.org/10.1016/j.jesp.2013.03.013>, 2013.
- Lolli, S., D'Adderio, L. P., Campbell, J. R., Sicard, M., Welton, E. J., Binci, A., Rea, A., Tokay, A., Comerón, A., Barragan, R., Baldasano, J. M., Gonzalez, S., Bech, J., Afflitto, N., Lewis, J., and Madonna, F.: Vertically resolved precipitation intensity retrieved through a synergy between the ground-based NASA MPLNET lidar network measurements, surface disdrometer datasets and an analytical model solution, *Remote Sens.*, 10, 1102, <https://doi.org/10.3390/rs10071102>, 2018.
- Manninen, A. J., O'Connor, E. J., Vakkari, V., and Petäjä, T.: A generalised background correction algorithm for a Halo Doppler lidar and its application to data from Finland, *Atmos. Meas. Tech.*, 9, 817–827, <https://doi.org/10.5194/amt-9-817-2016>, 2016.
- Manninen, A. J., Marke, T., Tuononen, M., and O'Connor, E. J.: Atmospheric boundary layer classification with Doppler lidar, *J. Geophys. Res.-Atmos.*, 123, 8172–8189, <https://doi.org/10.1029/2017JD028169>, 2018.
- Mather, J. H., Turner, D. D., and Ackerman, T. P.: Scientific maturation of the ARM Program, *Meteorol. Monogr.*, 57, 4.1–4.19, <https://doi.org/10.1175/AMSMONOGRAPHS-D-15-0053.1>, 2016.
- Morgan, M. G. and Henrion, M.: *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*, 172–219, Cambridge University Press, <https://doi.org/10.1017/CBO9780511840609.009>, 1990.
- O'Connor, E. J., Illingworth, A. J., and Hogan, R. J.: A technique for autocalibration of cloud lidar, *J. Atmos. Ocean. Tech.*, 21, 777–786, [https://doi.org/10.1175/1520-0426\(2004\)021<0777:ATFAOC>2.0.CO;2](https://doi.org/10.1175/1520-0426(2004)021<0777:ATFAOC>2.0.CO;2), 2004.
- Pearson, G., Davies, F., and Collier, C.: An analysis of the performance of the UFAM pulsed Doppler lidar for observing the boundary layer, *J. Atmos. Ocean. Tech.*, 26, 240–250, 2009.
- Pearson, G. N., Roberts, P. J., Eacock, J. R., and Harris, M.: Analysis of the performance of a coherent pulsed fiber lidar for aerosol backscatter applications, *Appl. Optics*, 41, 6442–6450, <https://doi.org/10.1364/AO.41.006442>, 2002.
- Roadcap, J. R. and Tracy, P.: A preliminary comparison of daylit and night  $C_n^2$  profiles measured by thermosonde, *Radio Sci.*, 44, <https://doi.org/10.1029/2008RS003921>, 2009.
- Rye, B. J. and Hardesty, R. M.: Discrete spectral peak estimation in incoherent backscatter heterodyne lidar. I. Spectral accumulation and the Cramer-Rao lower bound, *IEEE T. Geosci. Remote*, 31, 16–27, <https://doi.org/10.1109/36.210440>, 1993.
- Tonttila, J., O'Connor, E. J., Hellsten, A., Hirsikko, A., O'Dowd, C., Järvinen, H., and Räisänen, P.: Turbulent structure and scaling of the inertial subrange in a stratocumulus-topped boundary layer observed by a Doppler lidar, *Atmos. Chem. Phys.*, 15, 5873–5885, <https://doi.org/10.5194/acp-15-5873-2015>, 2015.
- Träumner, K., Handwerker, J., Wieser, A., and Grenzhäuser, J.: A Synergy Approach to Estimate Properties of Rain-drop Size Distributions Using a Doppler Lidar and Cloud Radar, *J. Atmos. Ocean. Tech.*, 27, 1095–1100, <https://doi.org/10.1175/2010JTECHA1377.1>, 2010.
- Tucker, S. C., Senff, C. J., Weickmann, A. M., Brewer, W. A., Banta, R. M., Sandberg, S. P., Law, D. C., and Hardesty, R. M.: Doppler lidar estimation of mixing height using turbulence, shear, and aerosol profiles, *J. Atmos. Ocean. Tech.*, 26, 673–688, <https://doi.org/10.1175/2008JTECHA1157.1>, 2009.
- Van Tricht, K., Gorodetskaya, I. V., Lhermitte, S., Turner, D. D., Schween, J. H., and Van Lipzig, N. P. M.: An improved algorithm for polar cloud-base detection by ceilometer over the ice sheets, *Atmos. Meas. Tech.*, 7, 1153–1167, <https://doi.org/10.5194/amt-7-1153-2014>, 2014.
- Westbrook, C., Illingworth, A., O'Connor, E., and Hogan, R.: Doppler lidar measurements of oriented planar ice crystals falling from supercooled and glaciated layer clouds, *Q. J. Roy. Meteor. Soc.*, 136, 260–276, <https://doi.org/10.1002/qj.528>, 2010a.
- Westbrook, C. D., Hogan, R. J., O'Connor, E. J., and Illingworth, A. J.: Estimating drizzle drop size and precipitation rate using two-colour lidar measurements, *Atmos. Meas. Tech.*, 3, 671–681, <https://doi.org/10.5194/amt-3-671-2010>, 2010b.
- Zhao, Y., Post, M. J., and Hardesty, R. M.: Receiving efficiency of monostatic pulsed coherent lidars. 1: Theory, *Appl. Optics*, 29, 4111–4119, <https://doi.org/10.1364/AO.29.004111>, 1990.