A generalised background correction algorithm for a Halo Doppler lidar and its application to data from Finland

Atmos. Meas. Tech., 9, 817–827, 2016
www.atmos-meas-tech.net/9/817/2016/
doi:10.5194/amt-9-817-2016

Atmospheric Measurement Techniques
Open Access
EGU

A generalised background correction algorithm for a Halo Doppler lidar and its application to data from Finland

Antti J. Manninen[1], **Ewan J. O'Connor**[2,3], **Ville Vakkari**[2], **and Tuukka Petäjä**[1]

[1]Department of Physics, University of Helsinki, Helsinki, Finland
[2]Finnish Meteorological Institute, Helsinki, Finland
[3]Department of Meteorology, University of Reading, Reading, UK

Correspondence to: Antti J. Manninen (antti.j.manninen@helsinki.fi)

Received: 3 September 2015 – Published in Atmos. Meas. Tech. Discuss.: 28 October 2015
Revised: 10 February 2016 – Accepted: 15 February 2016 – Published: 3 March 2016

Abstract. Current commercially available Doppler lidars provide an economical and robust solution for measuring vertical and horizontal wind velocities, together with the ability to provide co- and cross-polarised backscatter profiles. The high temporal resolution of these instruments allows turbulent properties to be obtained from studying the variation in radial velocities. However, the instrument specifications mean that certain characteristics, especially the background noise behaviour, become a limiting factor for the instrument sensitivity in regions where the aerosol load is low. Turbulent calculations require an accurate estimate of the contribution from velocity uncertainty estimates, which are directly related to the signal-to-noise ratio. Any bias in the signal-to-noise ratio will propagate through as a bias in turbulent properties. In this paper we present a method to correct for artefacts in the background noise behaviour of commercially available Doppler lidars and reduce the signal-to-noise ratio threshold used to discriminate between noise, and cloud or aerosol signals. We show that, for Doppler lidars operating continuously at a number of locations in Finland, the data availability can be increased by as much as 50 % after performing this background correction and subsequent reduction in the threshold. The reduction in bias also greatly improves subsequent calculations of turbulent properties in weak signal regimes.

1 Introduction

The greatest uncertainties in understanding the radiative balance of Earth are related to the effects of atmospheric aerosol particles (Boucher et al., 2013). In the atmosphere, aerosols can affect the radiative balance directly by absorbing and scattering solar radiation, and indirectly by changing cloud properties (e.g. Allen and Sherwood, 2010; Bony et al., 2013; Lohmann and Feichter, 2001; Lohmann and Hoose, 2009). In order to reduce the uncertainty in the impact of aerosol particles on Earth's climate, temporally and spatially representative measurements of their physical and chemical properties are vital. Although such measurements have been carried out at a number of ground-based measurement stations (Collaud Coen et al., 2013; Asmi et al., 2013), challenges remain in understanding the transport mechanisms that allow us to relate surface measurements to properties above the surface.

Doppler light detection and ranging (lidar) instruments provide a way of tackling this challenge through continuous measurements of the air motion simultaneously with scattering from aerosol. Turbulent profiles can be derived from these high-resolution measurements of vertical velocity (O'Connor et al., 2010), which can then be used to identify the mixing height, i.e. the height of the layer that is constantly in contact with the ground (e.g. Emeis et al., 2008; Pearson et al., 2010), and thus infer the aerosol transport.

At present, commercially available Doppler lidars, such as the Halo Photonics Stream Line Doppler lidar (Pearson et al., 2009), represent a solution for routinely measuring profiles of radial Doppler velocity and co- and cross-polarised signal-to-noise ratio, from which profiles of horizontal and vertical

winds are derived. These systems are based on fibre-optic technology utilising solid-state lasers in the infrared spectral band, and are capable of continuous operation for months or more (e.g. Harvey et al., 2013). However, the high-pulse repetition, low-pulse energy mode that these systems operate in, which means that they conform to eye-safety requirements, requires a certain amount of aerosol loading in the atmosphere for sufficient sensitivity (e.g. Pearson et al., 2009). In regions where the aerosol load is low, such as in remote continental regions, the resulting lack of suitable scatterers in the atmosphere becomes the limiting factor for their performance.

In this study, we present a novel post-processing algorithm to improve the background noise performance of Halo Doppler lidars. This algorithm is provided in the accompanying Supplement as a MATLAB programme together with a short atmospheric measurement and a ready-to-run example script. We demonstrate that the algorithm can significantly increase the data availability in low-aerosol conditions without decreasing the time or height resolution of the data set. The increase in data availability will enhance the capability for connecting surface aerosol measurements with aerosols that take part in cloud formation. This will, in turn, contribute towards reducing the uncertainties in the effects of aerosols on Earth's climate.

2 Measurements and instrument description

The Halo Photonics Stream Line Scanning Doppler lidar is a 1.5 µm pulsed Doppler lidar with a heterodyne detector that can switch between co- and cross-polar channels (Pearson et al., 2009). Here, we have configured the instrument to cover a range from 90 to 9600 m with 30 m resolution. The accumulation time per ray is also user-configurable and may vary from 1 to 30 s depending on environment and application. The instrument also has a full hemispheric scanning capability, but here we concentrate on the vertically pointing data. The technical specifications of the instrument as configured for standard operation are summarised in Table 1.

The instrument outputs a profile of back-scattered light intensity, in terms of signal-to-noise ratio (SNR), together with a profile of the radial Doppler velocity determined from the Doppler shift of the back-scattered light (Pearson et al., 2009). The attenuated backscatter profile can then be calculated from the SNR profile if the telescope function is known (Hirsikko et al., 2014). The uncertainty for each Doppler velocity measurement, σ_v, is then calculated from the corresponding SNR value. According to Rye and Hardesty (1993) and O'Connor et al. (2010), for a direct detection system the SNR is proportional to σ_v:

$$\sigma_v^2 \sim \frac{1}{\sqrt{\mathrm{SNR}}}, \quad (1)$$

Table 1. Summary of the technical specifications of the Halo lidar.

Wavelength	1.5 µm
Detector	heterodyne
Pulse repetition frequency	15 kHz
Nyquist velocity	20 m s^{-1}
Sampling frequency	50 MHz
Velocity resolution	0.038 m s^{-1}
Points per range bin	10
Range resolution	30 m
Pulse duration	0.2 µs
Lens diameter	8 cm
Lens divergence	33 µrad
Minimum range	90 m
Maximum range	9600 m
Telescope	monostatic optic-fibre coupled

whereas for a heterodyne system the relation is more complex:

$$\sigma_v^2 \sim \frac{1}{\mathrm{SNR}^2}(1+\mathrm{SNR})^2. \quad (2)$$

Thus, any bias or other error in SNR will directly bias the expected uncertainty σ_v. For low SNR conditions, especially, accurate knowledge of σ_v is crucial for calculating the dissipation rate of turbulent kinetic energy (O'Connor et al., 2010) and discriminating between turbulent mixing and instrumental noise (e.g. Vakkari et al., 2015).

The instrument performs a periodical background noise determination, typically once an hour for about 10 s when operating continuously. Since software version 10, the raw signals accumulated during the background determination are stored as text files on the lidar internal PC. However, in spite of this check, a small offset often remains in the instrument background. The manufacturer recommends post-processing the Halo Doppler lidar signal to identify and remove measurements that have an SNR of less than 0.015 (−18.2 dB). In most cases, this threshold is stringent enough to remove all issues arising from any background imperfections. However, this threshold places a severe restriction on data availability in very clean atmospheric environments as it reduces the likelihood of obtaining any useful signals; thus the strong motivation to reduce the threshold and improve the data availability.

In this study we utilise continuous measurements from a number of Halo Photonics Stream Line scanning Doppler lidars with similar configurations (Hirsikko et al., 2014) to illustrate the background correction method described in Sect. 3. The sites and the time periods included in testing the method are described in Table 2.

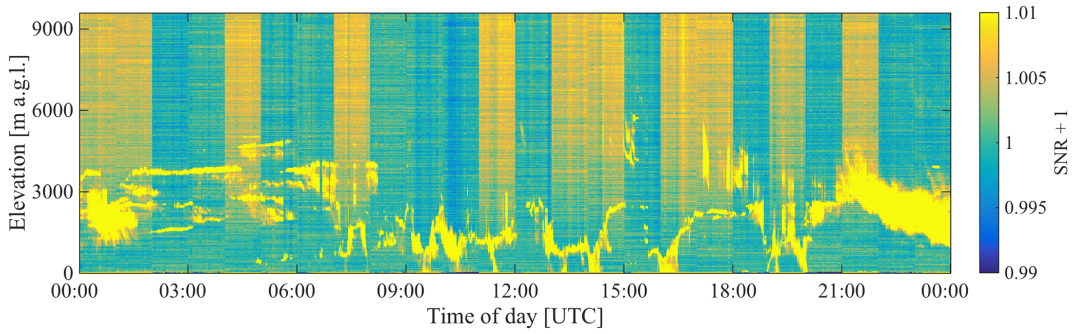

Figure 1. Time series of uncorrected Halo Doppler lidar SNR as a function of elevation above ground level (a.g.l.) measured at Hyytiälä, Finland, on 21 June 2014.

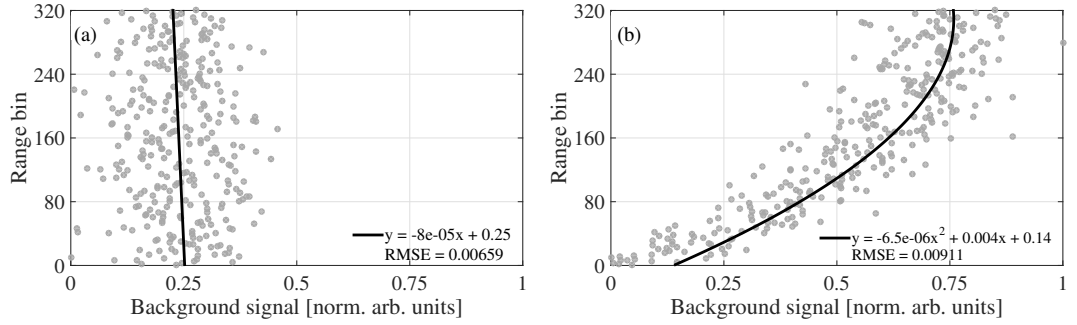

Figure 2. Two examples of background profiles with fitted polynomials emphasising the shape of the background. **(a)** Linear background profile measured on 21 June 2014, 13:00:21 UTC. **(b)** Second-order polynomial background profile measured on 21 June 2014, 12:00:20 UTC.

Table 2. Measurement site locations and data set time periods included in this study.

Site	Data set time period
Helsinki	1 Jan to 31 Mar 2014
Hyytiälä	1 Jan 2013 to 20 Apr 2014
Hyytiälä	21 May to 31 Dec 2014
Kuopio	1 Sep to 30 Nov 2014
Sodankylä	1 Sep to 30 Nov 2014
Uto	1 Sep to 30 Nov 2014

3 Background correction method

3.1 Description of the background artefact in Halo Doppler lidar

The existence of a background offset in the instrument can be seen in the time–height plot of SNR, where the hourly background determination introduces step changes in the SNR profile (Fig. 1). In older software versions (pre-10) the amplitude of these steps can be larger if the background determination is conducted when the instrument is receiving in the opposite polarisation to the previous determination. In addition, the shape of the instrument background can be either a linear function of range (Fig. 2a), or the background can follow a second-order polynomial (Fig. 2b).

3.2 Method for correcting the Halo Doppler lidar background artefact

There are two critical phases in correcting the background artefact in the Halo Doppler lidar SNR. The first one is the detection of any steps in the background, which occur due to periodic background determinations carried out by the instrument. The second one is to determine the shape of the background. The first phase is much more simple with software versions 10 and above, since the timing of each background determination can be retrieved directly from the background determination timestamp. However, there is a large amount of data collected with previous software versions, where background files are not available. In addition, there may be step changes in the background SNR that do not occur after a background determination, but have the same characteristics. Thus, the step detection algorithm may be useful even when the background files are available.

The workflow of the background SNR correction algorithm that has been developed is illustrated in Fig. 3. The algorithm applies the correction to the lidar SNR, and only SNR values are processed during the algorithm steps until the recalculation of the attenuated backscatter coefficients.

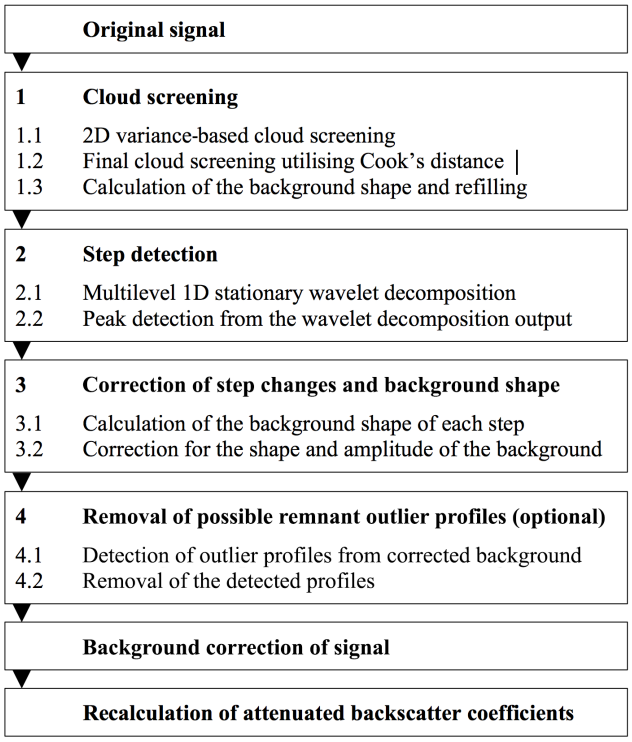

Figure 3. Chart showing the workflow of the Halo lidar background correction algorithm.

The correction algorithm is performed separately for each polarisation.

3.2.1 Cloud screening

In order to isolate any atmospheric measurements from the noise that will be used to characterise the shape and the magnitude of the background, returns originating from clouds and aerosols must first be identified so that they can be removed from subsequent processing. The cloud screening is performed in two consecutive steps: the initial coarse step identifies the regions of high variance in SNR, indicating the presence of cloud and aerosol; the second step identifies and removes additional atmospheric (cloud or aerosol) data points missed by the first step by calculating Cook's distance (Cook, 1977) values from robust bi-square weighted linear regression fits to each profile.

The initial component of the cloud-screening scheme (step 1.1 in Fig. 3) assumes that the variance in the SNR from clouds and aerosols is significantly larger than the variance of pure background noise. The variance is calculated by sliding an m by n (where m is the range bin and n is time) window over each profile, range bin by range bin. The size of the window affects how accurately the clouds and aerosols are detected and also the processing time. In this study, the window was selected to be 33 by 1. A dynamic threshold for the variance is then used to differentiate clouds and aerosols from pure background noise (Fig. 4). Besides assuming that the SNR from clouds and aerosols has higher variance, it is also assumed that the furthest range bins contain mainly background SNR and thus have lower variance. To find the dynamic threshold, the variance calculated from the SNR, which is contained in the upper 20 % of the range bins, is divided into n subsections. A larger number of subsections increases the probability of finding subsections containing mainly background SNR. In the results shown in this study, the value of n was chosen to be 64. The algorithm selects half of the subsections with the lowest median values as reference areas (Fig. 4). Then, an initial low threshold is chosen, which masks areas of SNR with high variance, but also areas of low variance. In this study we used the 50th percentile of the SNR variance contained in all range bins as the initial threshold. The threshold is then increased iteratively until the amount of masked pure background SNR is less than 1 % of the selected reference area. Thus, the dynamic threshold depends on the background noise statistics of the particular instrument in question.

The next step (step 1.2 in Fig. 3) is the fine-resolution cloud screening, which aims to remove any cloud or aerosol signals that were not removed by the variance-based method, and is based on the method discussed in Hoaglin and Welsch (1978). In short, each profile of the SNR is modelled with a robust bi-square weighted linear regression after the variance-based screening. Then, leverage points are calculated in each profile by constructing a hat matrix. Together with residuals of the modelled fits, the leverage points are used to calculate a Cook's distance, which describes an individual data point's influence on the least squares regression analysis. In literature, general threshold values based on the Cook's distance, D_{Cook}, have been suggested: $D_{\mathrm{Cook}} > 1$ (Cook, 1982), or $D_{\mathrm{Cook}} > 4/n$, where n is the number of observations (Bollen and Jackman, 1990). If D_{Cook} for a data point is larger than the threshold value, the point is considered as an outlier, and will be removed. We used the $D_{\mathrm{Cook}} < 4/n$ as a rule to separate the wanted background noise from the remaining data outliers due to aerosols or clouds (Fig. 5).

The combination of the SNR variance and Cook's distance schemes forms a robust mask for separating atmospheric signals (arising from clouds and aerosols) from noise-only signals. This mask is then used in later stages of the workflow. Note that for this particular application the mask may not be a reliable cloud detection mask as it may also contain non-atmospheric signals; it is designed so that the inverse of the mask contains random noise only.

SNR regions removed by the cloud screening have to be infilled for the wavelet decomposition used in the step detection phase (step 1.3 in Fig. 3 and Sect. 3.2.2). This is done by calculating first- and second-degree polynomials to each masked profile to characterise the shape of the cloud-screened. The best fit for each profile is chosen according to the goodness-of-fit indicator root-mean-square error (RMSE)

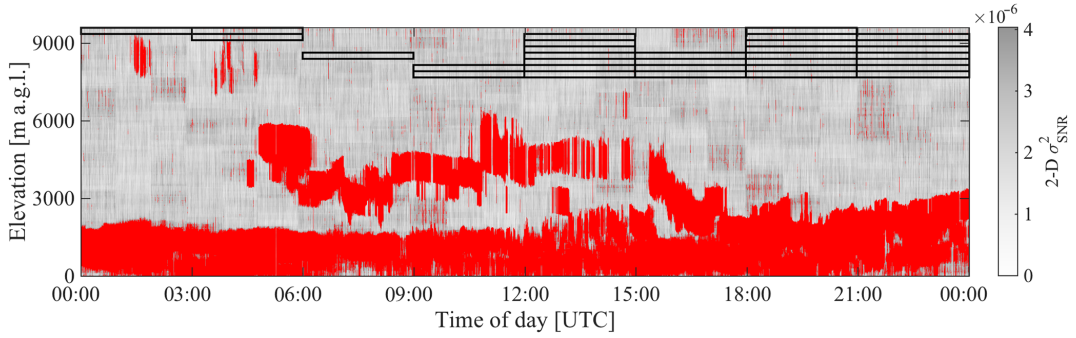

Figure 4. Time–height series of 2-D variance of SNR calculated from a Halo Doppler lidar at Hyytiälä, Finland, on 21 May 2014. The high variance areas (red) are masked using a dynamic threshold, which is found iteratively using the automatically selected reference areas (rectangles).

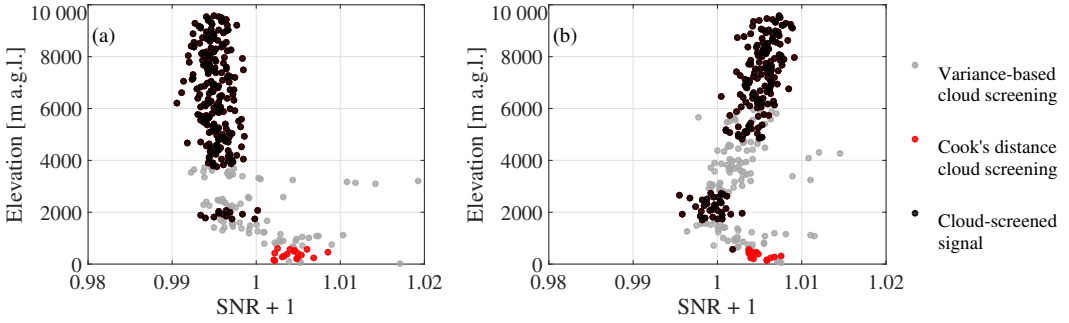

Figure 5. Two example vertical profiles of SNR with different background shapes measured by a Halo Doppler lidar on 21 May 2014 at **(a)** 07:24:24 UTC and **(b)** 08:33:35 UTC in Hyytiälä, Finland, illustrating the results of the two-cloud-screening scheme.

and then used to infill the cloud-screened regions. A robust 2-D interpolation is used for profiles where fitting is not possible due to insufficient data points.

3.2.2 Step detection from the background

The step detection routine finds both the times that background determinations were performed and minor step changes, which can occur between each background determination. A matrix (time by range) of SNR values for 1 day are processed in the time dimension using the multilevel 1-D stationary wavelet decomposition method (Nason and Silverman, 1995) using the orthogonal Haar wavelet (Daubechies, 1992). The chosen wavelet decomposition level, i.e. the number of iterations, affects the robustness of the step detection. For the data set used in this study, level 5 is the lowest level, which enables robust step detection.

The multilevel wavelet decomposition provides two outputs per level: the approximation coefficients, and the detail coefficients for each range bin. The wavelet decomposition is performed for SNR values over the furthest 75 % of the range. The detail coefficients from the highest selected level, here level 5, are summed together over all selected range bins. All of the peaks in the detail coefficients occur at the same time for all range bins, because for an individual profile, all range bins share the same timestamp. The summation of the detail coefficients over the whole range makes the step changes in the background more pronounced so that, for the range-summed detail coefficients (RSDCs), any step changes are represented as peaks (Fig. 6). Thus, the time of any step changes in the background is obvious and is determined using peak detection.

The peak detection uses a peak threshold, for which we selected the 75th percentile of the absolute RSDC. Then, the absolute RSDC time series are processed iteratively. A peak is defined as a local maximum whose difference to a preceding local minimum is higher than the chosen peak threshold.

For higher levels of wavelet decomposition, the step appears smoother and smoother in the approximation coefficients of the previous level. This shifts the RSDC peak positions towards the beginning of the time series. The shift is constant and directly proportional to the half-lengths of a particular wavelet level's high- and low-pass filters, and at level 5 the shift is 15 units on the time axis.

3.2.3 Correction for the step changes and the shape of the background

Within two consecutive step changes, a small temporal drift may occur in the background. To correct for this drift, the me-

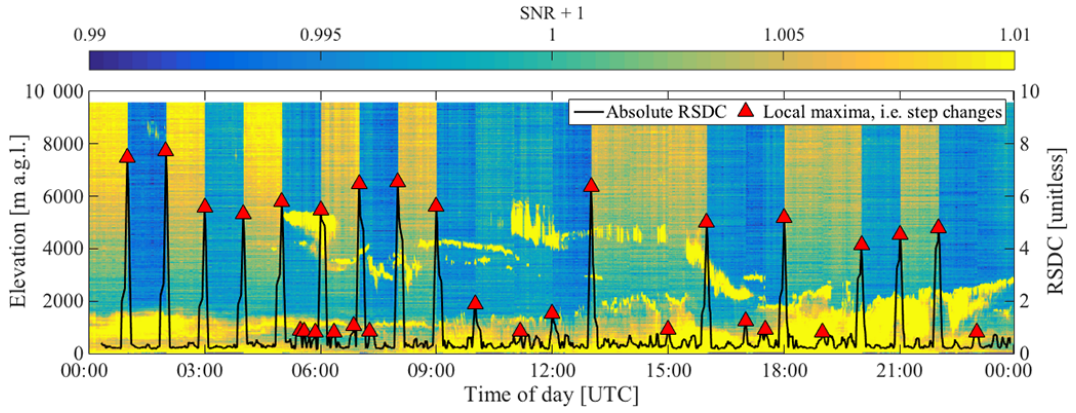

Figure 6. Time–height plot of uncorrected SNR from a Halo Doppler lidar operating at Hyytiälä, Finland, on 21 May 2014, together with the calculated absolute RSDC at wavelet decomposition level 5 (line), and the detected local maxima i.e. step changes (triangles).

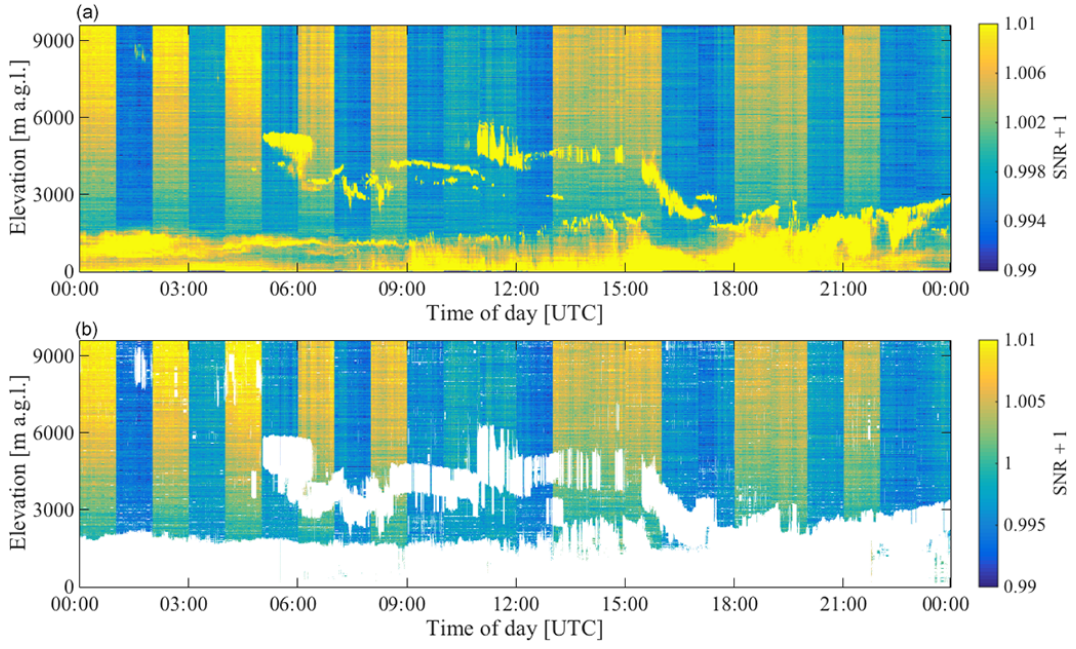

Figure 7. An example illustrating the performance of the cloud-aerosol masking scheme: **(a)** uncorrected SNR from a Halo Doppler lidar at Hyytiälä, Finland, on 21 May 2014, and **(b)** same data after cloud and aerosol have been identified and removed (masked regions in white).

dian cloud-screened SNR is calculated from each profile between two consecutive step changes. Then, the median drift is estimated with robust bi-square weighted linear fits, which are calculated from the medians. The temporal drift is then subtracted from the cloud-screened SNR, and stored for the final correction step.

The shape of the background noise-only SNR is determined by modelling the median profile of cloud-screened, and drift-corrected SNR between two consecutive step changes by robust bi-square weighted first- and second-order polynomials. The fit with the smallest RMSE is chosen to model the median profile. If the RMSE of the best fit is significantly larger than the expected noise level of the instrument, then the background correction between two consecutive step changes in question is rejected.

If the number of SNR pixels in the nearest half of the range for any particular step is very low, e.g. less than 5 %, the shape of the background is modelled with robust bi-square weighted first-order polynomial fit constrained to pass through $\mathrm{SNR} = 0$ at the nearest range bin. Finally, if an acceptable background shape can be determined, the fitted background and temporal drifts are subtracted from the original measured SNR profiles between the respective steps in the background (step 3.2 in Fig. 3).

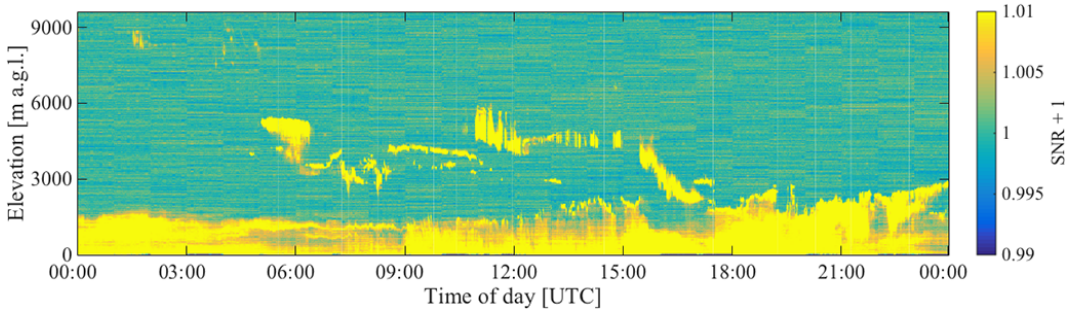

Figure 8. Corrected SNR from a Halo Doppler lidar operating at Hyytiälä on 21 May 2014.

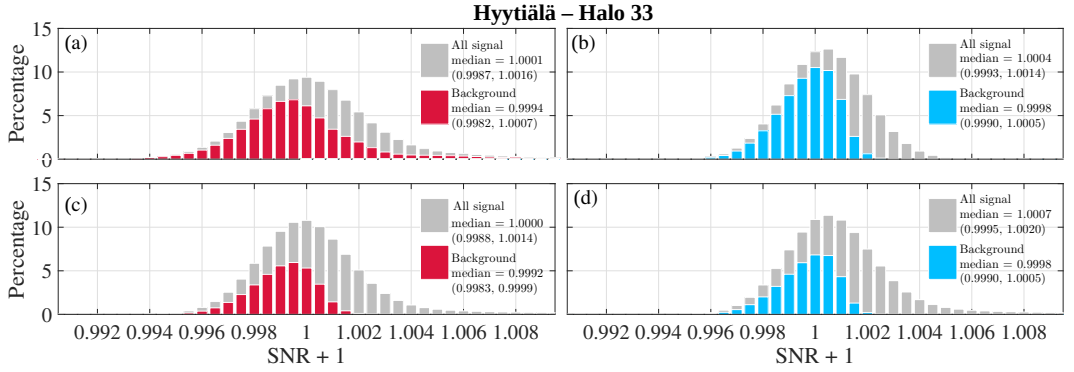

Figure 9. Histograms of **(a, c** grey) uncorrected, and (**b, d** grey) corrected SNR; (**a, c** red) uncorrected and (**b, d** blue) corrected background for the upper 20 % of range bins (**a, b**) and the lowest 20 % of the range bins (**c, d**). All data were measured with a Halo Doppler lidar (instrument ID: 33) between 21 May 2014 and 31 December 2014 at Hyytiälä, Finland.

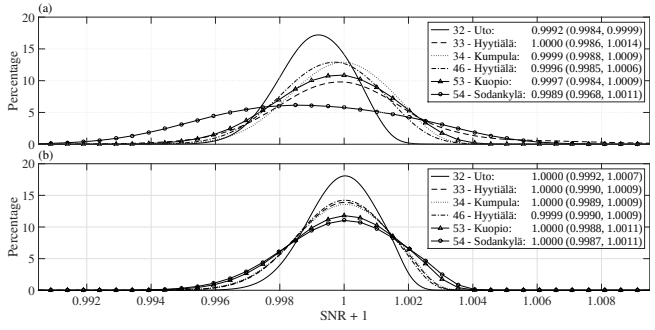

Figure 10. Histograms showing the **(a)** uncorrected and the **(b)** corrected background SNR for the upper 20 % range bins of the Halo Doppler lidar units in different locations. The specific measurement periods for the different locations are given in Table 2.

3.2.4 Removal of the remnant outlier profiles

For instruments that are not operating optimally, the background noise for some profiles may not be very well represented by the averaged approach in Sect. 3.2.3. To identify these profiles, the median background noise SNR of each profile after correction is calculated after reapplying the cloud mask from Sect. 3.2.1. Outlier profiles can then be detected as they exhibit peaks in the time series of the median background through using the same method as in Sect. 3.2.2. The outlier profiles can then be flagged and rejected, or the user may choose to apply the background noise profile shape detection and correction on a profile-by-profile basis, if the cloud-screened data availability permits.

4 Algorithm performance at several locations in Finland

In this section, the performance of the algorithm on this data set is evaluated. The cloud and aerosol screening is evaluated in Sect. 4.1; Sect. 4.2 addresses the accuracy of the step detection; the background step change and shape correction are discussed in Sect. 4.3; and finally, in Sect. 4.4, the effect of the algorithm on Halo Doppler lidar noise statistics and subsequently on the data coverage is presented.

4.1 Evaluating cloud and aerosol screening

Evaluation of cloud-screening methods is difficult for a single instrument if there is no ground truth to compare it with. Here, we were able to compare the Doppler lidar cloud and aerosol mask with observations from a co-located cloud radar and High Spectral Resolution Lidar at Hyytiälä over an 8-

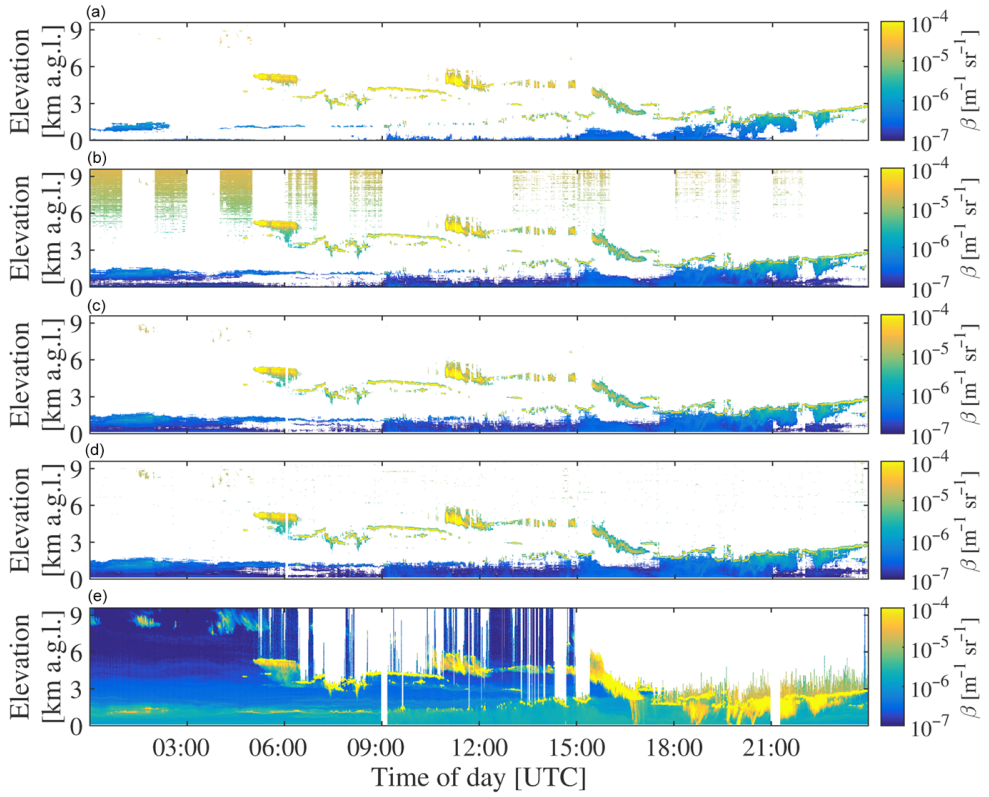

Figure 11. Time–height plots of attenuated backscatter at Hyytiälä, Finland, on 21 May 2014: **(a)** uncorrected attenuated backscatter (denoted β) with SNR limit of 0.0015, **(b)** SNR limit of 0.0075; **(c)** corrected attenuated backscatter with SNR limit of 0.0075, **(d)** SNR limit of 0.0055; **(e)** attenuated backscatter from a co-located High Spectral Resolution Lidar.

month period, with good agreement found. We illustrate the mask performance for an example day, where the background noise varies considerably as a function of range, with marked step changes over time (Fig. 7a). The characteristics of this example day form very challenging conditions for the masking algorithm, but the algorithm is able to identify the cloud and aerosol signals even when the noise-only signals may sometimes appear to have a higher SNR (Fig. 7b). The good performance of the masking algorithm enables the robust performance of the subsequent phases of the algorithm.

4.2 Step change detection accuracy

The accuracy of the step change detection was evaluated by comparing the available background determination timestamps with those detected from the data alone. There is a slight offset by default since the background determination cannot occur at the same time as a measurement. It is also possible for the background determination to occur before or in between either a co- or cross-polarised measurement. This slight time lag was compensated for by selecting a timestamp of the nearest backscatter profile measurement from a co- or cross-polarised channel nearest to the background determination timestamp. We can then compare the detected step times with the selected measurement timestamps.

More than 90 % of the detected steps match the timestamps for the background determinations in the data sets used in this study. The remaining 10 % can be explained by the fact that occasionally a step change is not present after a new background determination and that the step detection algorithm also picks up minor changes in the background that are not due to a new background determination, but some other change.

4.3 Evaluation of the background step change and profile shape correction

The background step change and shape correction is first evaluated visually with the chosen example day, 21 May 2014, then comprehensively by calculating histograms at two different ranges, and finally by calculating histograms from the background for all of the locations given in Table 2. Figure 8 shows SNR from a Halo Doppler lidar after performing the background correction. The white vertical lines note some remnant outlier profiles that have been removed (step 4 in Fig. 3). Figure 9 shows the SNR (unmasked) and the background SNR (masked) for the farthest 20 % (upper row pan-

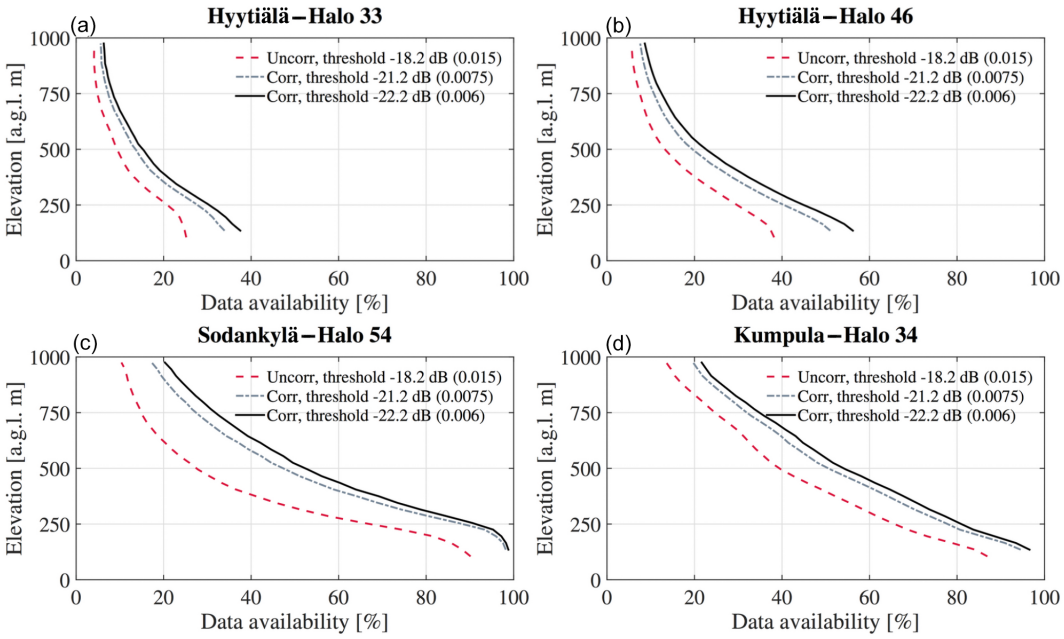

Figure 12. Data availability below 1000 m, where most signals originate from aerosol, for different SNR thresholds: (**a** and **b**, dashed red line) before, and (**a** and **b**, dot dashed gray and black lines) after processing with the background correction algorithm. Instrument ID 33 measured from 21 May to 31 December 2014 at Hyytiälä, Finland; instrument ID 46 between 1 January 2013 and 20 April 2014 at Hyytiälä, Finland; instrument ID 54 between 1 September and 30 November 2014 at Sodankylä, Finland; instrument ID 34 between 1 January and 31 March 2014 at Kumpula in Helsinki, Finland.

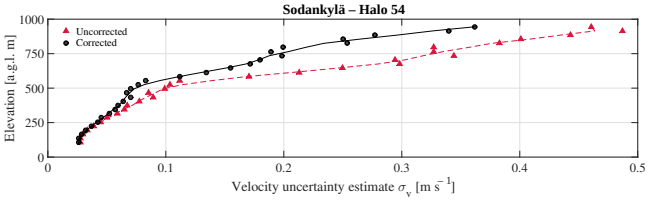

Figure 13. The velocity uncertainty estimate σ_V before and after applying the background correction for Halo Doppler lidar measurements during 1 September–30 November 2014 at Sodankylä, Finland. Each marker represents the calculated σ_V for the range bin in question, and the solid and the dashed lines represent the running medians with a window length of three range bins.

els) and the nearest 20 % (lower row panels) of the range bins from the instrument measured at Hyytiälä, Finland, during 21 May–31 December 2014, and Fig. 10 shows the background SNR for the farthest 20 % of the range bins at different locations measured during the time periods described in Table 2.

The results show that the full background correction algorithm successfully corrects for both the step changes and the shape of the background profile, producing a field of SNR with a homogeneous background. The histograms show that the median of the background SNR after the correction is closer to 1 (Figs. 9b, d, and 10b). The background spread is also reduced and, after the correction, the background SNR has nearly the same median and spread at both near and far ranges (Fig. 9b and d) as well as in the different locations (Fig. 10b).

4.4 Impact of the background correction algorithm on noise statistics and data coverage

Applying the background correction algorithm will significantly reduce the noise threshold level necessary to apply to data from a Halo Doppler lidar, especially for instruments that exhibit a strong background artefact (cf. Fig. 1). The background artefact correction (step correction) allows a substantial reduction in the threshold for automatic acceptance of data and, therefore, significantly increases the data availability in low-signal conditions. This is clearly shown in Fig. 11, where the amount of accepted data is increased dramatically in the lower altitudes after SNR has been processed with the background correction algorithm.

The corrected homogeneous background, as shown in Fig. 8, allows the SNR threshold to be set much lower than what has been suggested in earlier studies. For example, Pearson et al. (2010) suggested a threshold of −17 dB (0.020), the instrument manufacturer, Stream Line Photonics, has suggested a threshold of −18.2 dB (0.015), and Päschke et al. (2015) discussed decreasing the threshold to −20 dB (0.010). After the background correction, the SNR threshold can be set to −21.2 dB (0.0075; Fig. 11c) and tentatively even to −22.6 dB (0.0055; Fig. 11d). After applying this background correction, a SNR threshold of −22.2 dB

(0.006) is suggested as a robust value for processing larger data sets.

The impact of applying this background correction on data availability depends on the location, and more precisely the aerosol loading and number of weak SNR measurements that can be recovered by decreasing the SNR threshold. The SNR threshold can then be lowered after the background correction. In predominantly weak-signal environments, such as Hyytiälä, Finland, the background correction algorithm can lead to as much as a 50 % increase in the data availability, as shown in Fig. 12. The general differences in data availability between the four lidars in Fig. 12 are due to differences in atmospheric aerosol loads rather than instrument performance during the campaigns.

In addition to increasing data availability, the correction also improves the overall bias seen in SNR. For all instruments, the median bias is reduced to about 0.0002, with improvements of a factor of 5 or 10. Since the velocity uncertainty estimate, σ_v, is obtained directly from the SNR value (e.g. Rye and Hardesty, 1993), the reduction in SNR bias immediately impacts σ_v at low SNR, as shown in Fig. 13 where the σ_v is reduced more in the range bins where the SNR is low. Using typical velocity uncertainty estimates derived using the standard instrument specifications (O'Connor et al., 2010), σ_v for an SNR of 0.02 (-20 dB) is about $0.50\,\mathrm{m\,s^{-1}}$; a bias of 0.002 (as found for one instrument) for a similar uncorrected SNR value would lead to corrected σ_v of $0.54\,\mathrm{m\,s^{-1}}$ (SNR $= 0.022$) or $\sigma_v = 0.45\,\mathrm{m\,s^{-1}}$ (SNR $=$ 0.018) after correction depending on the sign of the bias. Since turbulent calculations require an accurate estimate of the contribution from velocity uncertainty estimates, a bias of about 10 % in the velocity uncertainty estimate can dominate any measurable turbulent contribution at low SNR in quiescent atmospheres and severely skew turbulent retrievals.

5 Conclusions

Halo Doppler lidars have been operating continuously at Hyytiälä and other locations in Finland since January 2013. Commercially available Doppler lidars offer a solution for obtaining high-resolution profiles of wind, turbulence, and the presence of cloud and aerosol layers.

However, the low-pulse energy and high-pulse repetition operation can result in sensitivity limitations for these instruments, and thus reduced data availability, when operated in regions with low-aerosol loads such as in boreal forests. Any attempts to average data to obtain signals below the standard operating thresholds can suffer if there are artefacts present in the background noise output by the instrument, since it can be difficult to discriminate between noise, and cloud or aerosol.

We have described a background correction algorithm which successfully corrects a number of artefacts present in the standard data output and enables the use of lower SNR thresholds, which can significantly increase the data availability by as much as 50 % at low altitudes in low-aerosol regimes. In addition, the reduction of any biases in the SNR propagates directly to the velocity uncertainty estimate and hence to reducing biases in turbulent calculations.

The background correction method can potentially be applied to any instrument types that display similar artefacts and can therefore improve data availability by reducing the SNR threshold required to discriminate between good signals and noise.

Additionally, the proposed cloud-screening scheme can be used in combination with other instrumentation to improve cloud detection in general. The main goal of the proposed cloud mask is to screen all of the atmospheric signal so that only the background signal remains. However, it is possible to adapt it to only mask clouds, and thus use it in cloud detection.

The Supplement related to this article is available online at doi:10.5194/amt-9-817-2016-supplement.

Acknowledgements. We gratefully acknowledge the financial support by the Nessling foundation (grant 201500326), Office of Science (BER) of the US Department of Energy, the Academy of Finland Centre of Excellence program (project number 1118615), and the Centre of Excellence (grant no. 272041).

Edited by: U. Wandinger

References

Allen, R. J. and Sherwood, S. C.: Aerosol-cloud semi-direct effect and land-sea temperature contrast in a GCM: AEROSOL EFFECT ON CLOUDS, Geophys. Res. Lett., 37, 1–5, doi:10.1029/2010GL042759, 2010.

Asmi, A., Collaud Coen, M., Ogren, J. A., Andrews, E., Sheridan, P., Jefferson, A., Weingartner, E., Baltensperger, U., Bukowiecki, N., Lihavainen, H., Kivekäs, N., Asmi, E., Aalto, P. P., Kulmala, M., Wiedensohler, A., Birmili, W., Hamed, A., O'Dowd, C., G Jennings, S., Weller, R., Flentje, H., Fjaeraa, A. M., Fiebig, M., Myhre, C. L., Hallar, A. G., Swietlicki, E., Kristensson, A., and Laj, P.: Aerosol decadal trends – Part 2: In-situ aerosol particle number concentrations at GAW and ACTRIS stations, Atmos. Chem. Phys., 13, 895–916, doi:10.5194/acp-13-895-2013, 2013.

Bollen, K. A. and Jackman, R. W.: Regression diagnostics: An expository treatment of outliers and influential cases, in: Modern methods of data analysis, Sage, Newbury Park, CA, 257–291, 1990.

Bony, S., Bellon, G., Klocke, D., Sherwood, S., Fermepin, S., and Denvil, S.: Robust direct effect of carbon dioxide on tropical circulation and regional precipitation, Nat. Geosci., 6, 447–451, doi:10.1038/ngeo1799, 2013.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S., Sherwood, S., Stevens, B., and Zhan, X.: Clouds and Aerosols, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T., Qin, D., Plattner, G., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Collaud Coen, M., Andrews, E., Asmi, A., Baltensperger, U., Bukowiecki, N., Day, D., Fiebig, M., Fjaeraa, A. M., Flentje, H., Hyvärinen, A., Jefferson, A., Jennings, S. G., Kouvarakis, G., Lihavainen, H., Lund Myhre, C., Malm, W. C., Mihapopoulos, N., Molenar, J. V., O'Dowd, C., Ogren, J. A., Schichtel, B. A., Sheridan, P., Virkkula, A., Weingartner, E., Weller, R., and Laj, P.: Aerosol decadal trends – Part 1: In-situ optical measurements at GAW and IMPROVE stations, Atmos. Chem. Phys., 13, 869–894, doi:10.5194/acp-13-869-2013, 2013.

Cook, R. D.: Detection of influential observation in linear regression, Technometrics, 19, 15–18, doi:10.2307/1268249, 1977.

Cook, R. D.: Residuals and influence in regression, Chapman & Hall, New York, NY, 1982.

Daubechies, I.: Ten Lectures On Wavelets, vol. 61 of CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, 8th Edn., 1992.

Emeis, S., Schäfer, K., and Münkel, C.: Surface-based remote sensing of the mixing-layer height – a review, Meteorol. Z., 17, 621–630, doi:10.1127/0941-2948/2008/0312, 2008.

Harvey, N. J., Hogan, R. J., and Dacre, H. F.: A method to diagnose boundary-layer type using Doppler lidar: A Method to Diagnose Boundary-Layer Type, Q. J. Roy. Meteor. Soc., 139, 1681–1693, doi:10.1002/qj.2068, 2013.

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, doi:10.5194/amt-7-1351-2014, 2014.

Hoaglin, D. C. and Welsch, R. E.: The Hat Matrix in Regression and ANOVA, Am. Stat., 32, 17–22, doi:10.2307/2683469, 1978.

Lohmann, U. and Feichter, J.: Can the direct and semi-direct aerosol effect compete with the indirect effect on a global scale?, Geophys. Res. Lett., 28, 159–161, doi:10.1029/2000GL012051, 2001.

Lohmann, U. and Hoose, C.: Sensitivity studies of different aerosol indirect effects in mixed-phase clouds, Atmos. Chem. Phys., 9, 8917–8934, doi:10.5194/acp-9-8917-2009, 2009.

Nason, G. and Silwerman, B.: The stationary wavelet transform and some statistical applications, Lect. Notes Stat., 103, 281–299, 1995.

O'Connor, E. J., Illingworth, A. J., Brooks, I. M., Westbrook, C. D., Hogan, R. J., Davies, F., and Brooks, B. J.: A Method for Estimating the Turbulent Kinetic Energy Dissipation Rate from a Vertically Pointing Doppler Lidar, and Independent Evaluation from Balloon-Borne In Situ Measurements, J. Atmos. Ocean. Tech., 27, 1652–1664, doi:10.1175/2010JTECHA1455.1, 2010.

Päschke, E., Leinweber, R., and Lehmann, V.: An assessment of the performance of a 1.5 µm Doppler lidar for operational vertical wind profiling based on a 1-year trial, Atmos. Meas. Tech., 8, 2251–2266, doi:10.5194/amt-8-2251-2015, 2015.

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, doi:10.1175/2008JTECHA1128.1, 2009.

Pearson, G., Davies, F., and Collier, C.: Remote sensing of the tropical rain forest boundary layer using pulsed Doppler lidar, Atmos. Chem. Phys., 10, 5891–5901, doi:10.5194/acp-10-5891-2010, 2010.

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, 1993.

Vakkari, V., O'Connor, E. J., Nisantzi, A., Mamouri, R. E., and Hadjimitsis, D. G.: Low-level mixing height detection in coastal locations with a scanning Doppler lidar, Atmos. Meas. Tech., 8, 1875–1885, doi:10.5194/amt-8-1875-2015, 2015.