# CL51 ceilometer and Halo Doppler lidar calibration using optically thick stratocumulus measurements and ceil\_halo\_calibrate.m

## Introduction

At Chilbolton Observatory we operate 2 low-power, infra-red lidar systems: a Vaisala CL51 ceilometer and a Halo Doppler lidar. Although they are sold as producing calibrated attenuated aerosol backscattering coefficient values, it is essential to check this calibration (as experience shows it to be inaccurate at the ~50% level) and to monitor it regularly for any trends which might indicate an instrument problem.

There are 2 possible calibration methods which can be applied to as-recorded atmospheric data. If there is an appreciable signal from Rayleigh scattering of the laser beam from atmospheric molecules, the lidar can be calibrated using the magnitude of this signal. However, this method is very difficult to apply to IR systems as the molecular scattering signal is very small. It is better suited to visible and UV systems. Fortunately, there is another valid option for IR systems which we use with our IR instruments.

When a lidar system is observing optically thick stratocumulus clouds, the measured signal is dependent on only a few factors [1]. Using this method, the calibration factor of a lidar can be derived. In this code we automatically select cases where we can be largely confident that the cloud was optically thick, non-drizzling (another important factor as it affects the droplet sizes within the cloud) and with low attenuation below it (to not significantly attenuate the cloud signal), then derive the calibration factor based on the signal from that cloud. The data are viewed as a probability density function (PDF) of the number of times each calibration factor in a range of values arises. These PDFs are viewed monthly so as to gain a sufficient number of values to provide meaningful statistics. They are displayed on a plot and saved as a text file which can be read in any plotting program.

## Code process

The user inputs the following parameters:

* Start day (usually the first day of a month)
* End day (usually the last day of a month, must be in the same year as the start day)
* A factor set to 1 or 2, which indicates the minimum peak value of backscattering coefficient to include in the statistics. 1 indicates a minimum peak of 1 x 10-4 m-1sr-1; 2 indicates 2 x 10-4 m-1sr-1. These values refer to the uncalibrated value of backscattering coefficient, as read from the raw data for that instrument. Both values should be used, in order to check that the calibration factor that is derived is not significantly affected by the code of peak height.

The code scrolls through each day in the time period (defined by start day, end day) in turn. For each day it reads the netCDF files that are written at:

CL51 /home/eoconnor/uncalibrated/cl51/*yyyy*/*yyyymmdd*\_chilbolton\_cl51.nc

Halo /data/lidar-doppler-halo/netcdf/calibrated/halo-doppler-lidar-118-co/*yyyy*/

*yyyymmdd*\_chilbolton\_halo-doppler-lidar-118-stare-co.nc

The latter is a Halo file which has been cleaned and processed using the FMI Halo toolbox, a Matlab code which is run daily on the Matlab machine. It is possible to edit the code to change to using the uncalibrated file in /data/lidar-doppler-halo/netcdf/uncalibrated/halo-doppler-lidar-118-co/*yyyy*/.

The code could be edited to read the netCDF files produced using Python, stored in /data/amof-netCDF/ncas-ceilometer-5 and …/ncas-lidar-dop-3.

[1] states that the backscattering coefficient integrated through an optically thick stratocumulus cloud, *B*, is given by

where ** is a factor which accounts for the apparent increase in backscatter and reduction in attenuation due to multiple scattering and *S* is the lidar ratio of the cloud droplets. Figure 5 of [1] shows a plot of ** *S* as a function of height for the Chilbolton Observatory CT75K ceilometer (the predecessor of the CL51) and a value of *S* is given as 18.8 ± 0.8 sr, calculated using Mie scattering theory.

In this code the value of *S* is calculated from a fit to figure 5 for the CL51 and fixed at a value of 20 sr for the Halo2 lidar (taken from a private communication with Ewan O’Connor – the fixed value results from the narrower receiver field of view for the Halo2 lidar compared to the ceilometer).

The method to identify suitable lidar profiles and calculate the calibration factor uses the following steps:

1. Set a range of heights from which clouds will be considered for calibration. This has been chosen as 0.5 – 2.0 km, with the aim of avoiding both very low clouds which are more likely to be drizzling and higher clouds which are less likely to be optically thick.
2. Set a minimum cloud peak height, *h*, for a profile to be included (to help ensure only optically thick clouds are included). This is the value of 1 x 10-4 m-1sr-1 or 2 x 10-4 m-1sr-1 chosen at the command line.
3. Scan the data hour by hour, only including hours where at least 95% of the profiles achieve conditions 2 and 3. This again helps to ensure that only optically thick clouds, which have a high level of sky coverage, are included.
4. Set a maximum elastic backscattering signal below the cloud of 1 x 10-5 m-1sr-1, (up to a few bins below the peak) to exclude cases with drizzle or high boundary layer aerosol loading. To test the sensitivity of the calibration factor to this, change the value in the code.
5. Calculate the appropriate value of ** *S* for the height of the cloud peak.
6. Calculate the uncalibrated integrated cloud peak *Bu* using the points around the peak where the signal is above 5% of the peak height – this defines the peak.
7. Calculate the calibration factor, *k*, defined by

and plot as a probability density function (PDF). For the purposes of applying a calibration to the lidars, PDFs are grouped both by month and by year.

## Matlab commands

The code is currently best run on the Matlab machine as it has a newer version of Matlab than other Chilbolton machines. To run Matlab on this machine you’ll need to log in as a superuser. You also need to enter some commands to enable X11 forwarding to work if you’re running this from a Windows machine using Exceed to provide a graphical display from the PuTTY (or similar) window that you’re using to access Chilbolton machines.

Until now I’ve ssh’ed to ‘matlab’ machine from ‘wilma’. Since the failure of wilma in late January 2022 it has been better to do this from django, in order to make X11 forwarding work. You may, as a result of the xauth list $DISPLAY command, get a long list of cookie addresses – I found that choosing the one that starts ‘matlab/unix:11’ works.

For example:

jla@matlab:~$ xauth list $DISPLAY

This gives a reply something like

matlab/unix:11 MIT-MAGIC-COOKIE-1 3018efd2989cef517ce395f85ef6d565

There may be a long list of replies – if so, pick out the one that starts matlab/unix:11.

Copy this line by highlighting it on the screen, then use the “sudo” command to log in as matlab rather than yourself.

jla@matlab:~$ sudo su matlab

Enter password…

Paste that address into the following command “xauth add…”

matlab@matlab:/home/jla$ xauth add matlab/unix:11 MIT-MAGIC-COOKIE-1 3018efd2989cef517ce395f85ef6d565

The long cookie “address” changes each time you log in so you need to do it each time.

matlab@matlab:/home/jla$ matlab -nodesktop

Matlab will now open. If you get errors regarding X11 forwarding, the above process hasn’t worked.

>>cd /home/jla/matlab/svn\_lidar\_cal/

>> ceil\_halo\_calibrate(*yyyymmdd\_start, yyyymmdd\_end, n*)

where *yyyymmdd\_start, yyyymmdd\_end* are the start, end dates of the period to process (in the same year) and *n* is set to 1 or 2 to indicate the minimum stratocumulus peak height to include in the processing (indicating 1 x 10-4 m-1sr-1 or 2 x 10-4 m-1sr-1 respectively).

## Outputs from the code

Output files from the code are written to /data/range/ceil\_halo\_cal.

### A plot of the PDF of the calibration factor for each instrument

The plot file name format is ceil\_halo\_cal\_*yyyymmdd\_start\_yyyymmdd\_end\_cloud‑threshold\_bl‑threshold*.png

Where *yyyymmdd\_start* and *yyyymmdd\_end* are as defined earlier, *cloud\_threshold* is the peak height threshold defined at the command line and *bl-threshold* is the maximum boundary layer signal below the cloud, currently defined in the code as 1 x 10-5 m-1sr-1.

An example of a yearly plot for 2021 is shown below:

Timeline

Description automatically generated

The plot shows the trend in the calibration factor each month through the period. The different intensities of colour each month indicate the relative numbers of profiles with suitable stratocumulus cloud that were detected during the month.

### A text file showing the data included in the calibration factor plot

The text file name format is ceil\_halo\_pdf\_*yyyymmdd\_start\_yyyymmdd\_end\_cloud‑threshold\_bl‑threshold*.txt.

The format is tab-delimited ASCII, aimed at being simple to read into Excel. There are no header lines.

It contains the data which were used to produce the plot in the following format:

The first column shows the number of the month, from 1 to 12.

There are then a further 200 columns which represent the PDF, first for the CL51 then for Halo2. Each column shows the number of counts within a calibration factor bin of width 0.05, in the range 0 to 5. Hence the first column shows calibration factors in the range 0.0 to 0.05 for the CL51, then 0.05 to 0.1, 0.1 to 0.15 etc. up to 4.95-5.0. The same sequence is then repeated for the Halo2 lidar.

An example of an excerpt from a file after importing into Excel is shown below. The first 4 months, indicated by the first column, are shown. The second column shown is the number of CL51 calibration factor values in the range 0.0 to 0.05 and the last that is visible in this truncated example is for the range 0.95 to 1.0. In the full file this continues up to 4.95 to 5.0 then is repeated for the Halo2 lidar.



This file enables the user to examine the PDF each month and also to sum over several months or the whole year. These data files are recommended for determining the calibration factor to use each month when producing calibrated netCDF files.

### A statistics text file showing details of the profiles from each day

The text file name format is ceil\_halo\_stats\_*yyyymmdd\_start\_yyyymmdd\_end\_cloud‑threshold\_bl‑threshold*.txt.

The format is tab-delimited ASCII, aimed at being simple to read into Excel. There are no header lines.

It contains details of the number of profiles each day that meet the different selection criteria for inclusion in the calibration factor calculation, as described in steps 1-3 above.

The first column shows the date in the format *yyyymmdd*. Columns 2-6 give the number of profiles meeting each criterion for the CL51 ceilometer and columns 7-11 give the same quantities for the Halo2 lidar. Once a profile has failed a particular step, it is not made available for the next one, so the number in each column progressively decreases.

|  |  |
| --- | --- |
| **Column number** | **Selection criterion** |
| 2, 7 | Number of profiles recorded that day |
| 3, 8 | Number of profiles where the maximum signal is at a height in the range 0.5 – 2.0 km |
| 4, 9 | Number of profiles where the maximum signal is greater than the threshold set on the command line. Profiles are only included if at least 95% of profiles in the hour (i.e. 00-01UT, 01-02UT etc.) are above the threshold, to help ensure that the cloud cover is not intermittent. |
| 5, 10 | Number of profiles where the maximum signal is greater than the threshold set on the command line. Profiles are only included if at least 95% of profiles in the hour (i.e. 00-01UT, 01-02UT etc.) are above the threshold, to help ensure that the cloud cover is not intermittent. |
| 6, 11 | Number of profiles where there is no signal below the width of the peak with a signal greater than the threshold set for the boundary layer ( |

An excerpt of a statistics file is shown below, showing all columns for the first 10 days of 2021.



[1] Ewan J. O’Connor, Anthony J. Illingworth and Robin J. Hogan, A Technique for Autocalibration of Cloud Lidar, Journal of Atmospheric and Oceanic Technology, 21(5), 777-786, 2004. <https://doi.org/10.1175/1520-0426(2004)021%3C0777:ATFAOC%3E2.0.CO;2>

## Calibration results

The table below shows the lower limit of the bin with the maximum number of counts in the PDF for each year, for the CL51 and the Halo lidar, using a peak threshold of 1 x 10-4 and 2 x 10-4 m-1sr-1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Year** | **Bin at peak of PDF** | | | |
|  | **CL51 peak = 1 x 10-4 m-1sr-1** | **Halo peak = 1 x 10-4 m-1sr-1** | **CL51 peak =** **2 x 10-4 m-1sr-1** | **Halo peak = 2 x 10-4 m-1sr-1** |
| 2015 | 1.55 – 1.60 |  | 1.50 – 1.55 |  |
| 2016 | 1.60 – 1.65 | 0.45 – 0.50 | 1.55 – 1.60 | 0.45 - 0.50 |
| 2017 | 1.60 – 1.65 | 0.45 – 0.50 | 1.60 – 1.65 | 0.45 – 0.50 |
| 2018 | 1.60 – 1.65 | 0.50 – 0.55 | 1.60 – 1.65 | 0.50 – 0.55 |
| 2019 | 1.60 – 1.65 | 0.50 – 0.55 | 1.60 – 1.65 | 0.50 – 0.55 |
| 2020 | 1.65 – 1.70 | 0.65 – 0.70 | 1.60 – 1.65 | 0.65 – 0.70 |
| 2021 | 1.60 – 1.65 | 0.60 – 0.65 | 1.60 – 1.65 | 0.60 – 0.65 |

In most cases there is agreement between the most common calibration factor bin for the 2 different cloud peak heights. Where there is a difference, that calculated using a minimum peak height of 2 x 10-4 m-1sr-1 is used. The calibration factor is taken as the midpoint of the bin, e.g. 1.625 for the 1.60 – 1.65 bin.

The chosen calibration factor is entered from the command line as described in cao\_lidars\_netcdf\_production.docx and for processing via a cronjob, the scripts py-cl51-today.sh, py-cl51-yesterday.sh and py-cleaned-halo-yesterday.sh must be edited.