

# Authors' response to the referee comments on *"Ground-based lidar processing and simulator framework for comparing models and observations"*

P. Kuma, A. J. McDonald, O. Morgenstern, R. Querel, I. Silber, and C. J. Flynn

14 October 2020

Dear Editor and Anonymous Referees,

We would like to thank the anonymous referees for their comments. Please find our response to the comments below. The original referee comments are marked in **bold** above our response. The revised manuscript and a latexdiff document highlighting the changes are appended at the end. The page and line numbers (PxxLxx) in our response refer to this most recent latexdiff document. References to sections, figures and tables in our response refer to the revised manuscript.

## **Anonymous Referee #1**

The paper fits to the scope of GMD however, before publication a few clarifications are required, and the structure must be revised (order of figures should be 1,2,3,..., outline of the paper), it seems that the reorganisation has been done very fast after the quick review .

### **General comments:**

1. The requirements on simulating and/or evaluating ALC-signals depend on the application: for the determination of cloud bottom heights (CBHs) they are certainly quite different compared to particle backscatter coefficients. It should be made very clear in the paper, which application is aimed at and how the requirements for this application are fulfilled. An outlook on planned extensions can be given for some relevant applications not yet implemented.

We want to clarify that we do not evaluate cloud base height in our manuscript, but rather the full cloud mask as determined by the cloud detection algorithm based on the attenuated volume backscattering coefficient. This approach provides a more comprehensive comparison between observations and models. It is more dependent on accurately calibrated absolute attenuated volume backscattering coefficient. The ALCF output also contains cloud base height, but it is simply determined from the lowest level of the cloud mask.

Aerosols are not currently implemented in the lidar simulator (this is also true for the original spaceborne lidar simulator) and also not implemented in the detection algorithm. We added a sentence to the abstract "While the correct focus of the framework is on clouds, support for aerosol in the lidar simulator is planned in the future." (P1L20) We also added text to Sect. 4 stating explicitly that aerosols are not implemented in the simulator and the focus is on cloud evaluation (P9L17) and text in Discussion and conclusions (P36L11).

2. Section 2: The lidar simulator should be explained in more detail and whether the focus of ALCF1.0 is on clouds or aerosols or both. In the present state it seems to be clouds , as this is much easier to be treated: The numerical models/reanalyses provide the necessary information and the variability of the lidar ratios is comparable small. Low clouds typically consists of droplets so the consideration of non-spherical ice crystals is not highly important (provided that the focus is on CBHs only). Anyway: it is strongly recommended to include to consideration of optical properties of ice crystals into ALCF2.0 (phase functions are available since decades).

We recognise that implementation of aerosols would be a desirable feature of the simulator. This is, however, a relatively complicated problem due to the relatively large variety of aerosol types implemented in different models

and would deserve its own paper. Also reanalyses generally do not provide aerosol concentrations in their standard output. The current focus of the simulator is on clouds. Aerosols affect the analysis of backscattering from clouds by attenuating the lidar signal before reaching the cloud. We do not expect this attenuation to significantly affect cloud detection unless the measurements are done in a location with high levels of aerosol concentration (P36L11).

As we note above, the current focus is on clouds and we made clarifications in the text (P1L20, P9L17). In the revised version we implemented backscattering from ice (Sect. 4.3, P9L20, P35L7), where the lidar ratio varies much more considerably than in liquid clouds. As we note above, cloud base height is not evaluated in the manuscript but rather the cloud mask.

The knowledge of the phase function of particular ice crystals is not enough for implementation of ice in the simulator because of the unknown distribution of shapes and size of ice crystals. Therefore, we decided to base the ice lidar ratio on global observations from CALIPSO (Garnier et al., 2015).

**The treatment of aerosols is not sufficiently explained: the backscatter signal depends on the aerosol distribution and their optical properties. Where is this input coming from in cases when the model/reanalysis does not provide aerosol information (NWP not necessarily consider aerosols). What about non-spherical particles – in the case of e.g. dust or volcanic ash the application of Mie theory is certainly not justified. Even if aerosol applications are not included in the paper this topic must be discussed (maybe as outline of the next version of ALCF). Chan et al. (2018) have demonstrated that there is an influence of particle shape on the backscatter profiles (and the intercomparison with model results), and that consideration of aerosols and their non-sphericity is possible.**

Chan et al: Evaluation of ECMWF-IFS (version 41R1) operational model forecasts of aerosol transport by using ceilometer network measurements, *Geosci. Model Dev.*, 11, 3807–3831, <https://doi.org/10.5194/gmd-11-3807-2018>, 2018.

As we note above, aerosol is not currently included in the simulator, even though some models such as the UM potentially provide information about aerosol concentration. Our case studies are affected by relatively little aerosol. The atmosphere of New Zealand has little anthropogenic aerosol except for major cities especially in winter. Attenuation by aerosol is unlikely to be a significant factor for studying backscattering from clouds, which generally produce about an order of magnitude or more backscattering.

Aerosol affects the attenuated volume backscattering coefficient by producing additional backscattering on top of the molecular and cloud backscattering, as well by attenuating the signal. This could be partially avoided by filtering out time periods with high concentrations of aerosol identified manually (automatic identification would probably require some kind of a machine learning algorithm). In our datasets aerosol with volume backscattering coefficient of up to about  $2 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  was present, spanning range of up to about 4 km. If we assume lidar ratio of 50 sr, this would produce two-way attenuation of up to about 50%. Therefore, this can be expected as the worst-case limit for accuracy of the molecular backscattering calibration in our dataset. For the purpose of determination of the cloud mask this should be still acceptable, as the cloud volume backscattering coefficient varies over multiple orders of magnitude (from approximately 2 to  $200 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  in our dataset). We clarified this in Discussion and conclusions (P36L11).

The optical properties listed in Chan et al. (2018) would be a good start for implementation of aerosols in the simulator. However, more aerosols type might need to be calculated for models such as the UM, which provides a much larger variety of aerosol species in its output.

**3. Section 4: In the introduction of this section again the description of the treatment of aerosols has been forgotten.**

We we note above, we clarified this in the abstract and in the introduction of Sect. 4 (P1L20, P9L17).

**In section 4.1 (starting with the paragraph above Eq. 2) it should clearly stated whether the authors talk of clouds or aerosols. Aerosol size distributions typically are not described by gamma-distributions, and the**

application of Mie theory is often not adequate.

We replaced "atmospheric constituents" with "cloud droplets" (P11L16).

Moreover, the authors should comment how they deal with the lidar ratio: when CBHs shall be derived the magnitude of the lidar ratio is not very important: the backscatter coefficient is such large that the uncertainty of the lidar ratio is more or less a second order effect. They should also comment on the consequences of the wrong description of ice crystals by the Mie theory: is this tolerable in view of the overall relevance of the lidar ratio in case of CBHs?

As we note above, cloud base height is not evaluated in the manuscript but rather the cloud mask. The variability of lidar ratio of cloud droplets is relatively small (ranging between approximately 14 and 19 sr for effective radius between 10 and 30  $\mu\text{m}$  and lidar wavelength between 532 and 1064 nm (Fig. 2b), i.e. variability of about 30%, this is potentially significant and we think it is justified to parametrise lidar ratio as a function of the effective radius and lidar wavelength in the simulator. Ice crystals have a greater variability of lidar ratio between approximately 20 and 50 sr (Fig. 2c), and therefore correct parametrisation of lidar ratio is even more important.

**Are the derived microphysics of the cloud particles (mainly size distribution) checked against the input used in the models/reanalyses. If it is available it should be consistent (ALCF vs. model).**

We showed that the size distribution assumption does not produce a significant difference in the lidar ratio (Fig. 2b). Therefore, we do not expect that it would produce a significant effect on the attenuated volume backscattering coefficient. In the future the size distribution assumption could be made consistent with each type of model and reanalysis (if this information is available). This kind of information is not generally present in model and reanalysis output.

**In Section 4.3 multiple scattering is briefly described. Of course it belongs to a full description of radiative transfer but in case of CBH-determination I don't believe that this is really relevant (the photons stay close to the optical axis). A clarifying sentence might be added.**

As we note above, cloud base height is not evaluated in the manuscript but rather the cloud mask. Previous studies identified potentially large multiple scattering correction required for a ceilometer with a similar field of view as our instruments (110–830  $\mu\text{rad}$ ) (O'Connor et al., 2004).

**4. Section 7: This section is quite long and should be separated into subsections. If done so 7.1 would deal with Fig. 4, Section 7.2 with Figs. 5-7, and Section 7.3 with Fig. 8 and so on.**

We split Section 7 into subsections "Cass", "Lauder", "Christchurch", "Backscattering on daily scales", "Cloud fraction and cloud albedo", "Molecular backscattering, aerosol backscattering and noise".

**In this context I suggest to delete Section 7.3 (Fig. 8): The arguments of the authors with respect to the cloud albedo are very difficult to understand and thus not convincing, and for me highly speculative: How is the effect of overlapping clouds in different levels considered?**

The effect of overlapping clouds is treated implicitly by the lidar simulator. It simulates cloud obscuration as would happen in reality for the given model cloud mixing ratios and cloud fraction. In that sense, the comparison of LR between observations and models is unbiased, but only if the effect of aerosol can be neglected. We agree that the lack of aerosol in the lidar simulator could cause a systematic bias in the results presented in Fig. 8. Therefore, we decided to remove Fig. 8 and the related discussion (Sect. 7.3 and P11L13).

**How accurate is the estimation of the lidar ratio (ice crystals)?**

The revised version of the simulator has parametrisation of lidar ratio and multiple scattering coefficient of ice crystals based on temperature (Sect. 4.3). Uncertainty still remains due to the fact that we used global empirical distributions derived from CALIPSO and there are more factors than temperature which affect the optical properties of ice crystals.

The calculation of the lidar ratio from the vertically integrated backscatter is not explained (in contrast to the extended description of  $r_{\text{eff}}$  etc. in Section 4.1). How is obscuration of high clouds by low clouds considered, and so on. An ALC can provide a lot of useful information so it is not necessary to invent retrievals of further parameters associated with low confidence.

We added Sect. 4.1 which reviews lidar ratio and its derivation from the vertically integrated attenuated volume backscattering coefficient. This is an established parameter used by previous studies (O'Connor et al., 2004).

In connection with Figs. 5–7 it should be outlined why the backscatter coefficient is plotted and not the cloud mask. In the latter case the model will show clouds that might be obscured by low clouds in case of ALC-measurements. This could give further insight into the fundamental differences between remote sensing and modeling (and the problems associated with the resolution).

In the original Fig. 4–6 the lower limit for the attenuated volume backscattering coefficient was chosen the same as the threshold for cloud detection. Therefore, showing the cloud mask would not provide additional information. In the revised figures the lower limit is  $0.1 \times 10^{-6} \text{ m}^{-1} \text{ sr}^{-1}$  and the cloud mask is plotted as dashed outlines. The input fields of cloud liquid water and ice mixing ratios and cloud fraction are in panels below. It should be clear how the obscuration by low cloud affects higher level clouds in the simulated attenuated volume backscattering coefficient.

In connection with Figs. 9–10 again the aerosol related questions from above show up.

Aerosol affects these figures by producing additional backscattering on top of the molecular and cloud backscattering, as well by attenuating the signal. This could be partially avoided by filtering out time periods with high concentrations of aerosol identified manually (automatic identification would probably require some kind of a machine learning algorithm). As we note above, in our dataset aerosol could produce additional attenuation of up to about 50% and this can be assumed to be the worst-case limit of the molecular backscattering calibration method, which is still good enough for absolute calibration for the purpose of cloud detection.

**Specific/minor comments:**

1. 2/8: aerosol optical depth should be replaced by particle backscatter coefficient: AOD is not a primary output of ALCs as it is an integral quantity, and depends on the lidar ratio.

We replaced this with "particle volume backscattering coefficient" (P2L6).

2. 2/14: thousands is a little bit exaggerated: each network rather consists of a few hundreds at maximum.

We replaced "thousands" with "hundreds" (P2L15).

3. 2/18: model evaluation : here a few references would be fine.

We added references to Hogan et al. (2001); Illingworth et al. (2007); Morcrette et al. (2012); Warren et al. (2018); Lamer et al. (2018) and Hansen et al. (2018) (P2L18).

4. 4/13 ff: check order!

We reformulated this paragraph (P4L17).

5. 5/18 (and throughout the paper): please avoid terms like backscatter : either backscatter coefficient, particle backscatter coefficient, attenuated backscatter or any other clearly defined physical quantity.

Where appropriate, we replaced occurrences of "backscatter" with "attenuated volume backscattering coefficient".

6. 6/1: in the near IR spectrum can be deleted, the same is true for in the visible spectrum in line 17.

We put "near IR" and "green colour in the visible spectrum" in parentheses (P7L1, P7L7, P7L18). We think this is still useful information, even though most readers might be well aware of this.

7. 6/2: the native resolution is indeed 5 m, but the typical output-resolution is 15 m.

We clarified this with "vertical sampling resolution 5 m in the first 150 m and 15 m above" (P7L2).

8. 6/3: uncalibrated : this is a little bit misleading as the signals (to my knowledge) are calibrated against a standard instrument. This has been done to make all instruments within a network comparable. I don't know if this applies to the ceilometers in New Zealand as well.

According to information in (Hopkin et al., 2019, Fig. 8, Fig. 13) the instruments are not all absolutely calibrated to the same reference. The instruments in our study are off-the-shelf instruments. Their calibration was done by the manufacturer. Therefore, determination of the calibration coefficient is necessary to get the absolutely calibrated attenuated volume backscattering coefficient. Without the calibration step the per-instrument default values of the calibration coefficient are used.

9. 6/18: is up to 5 m : what does this mean? Is the resolution coarser (10 m or so) or finer (3 m)?

We replaced this with "The vertical resolution is 5–75 m and sampling rate 1 s." (P7L21).

10. 6/19: The instrument can be housed... : Why is this mentioned? It is true for the other instruments as well.

This is mentioned because there is an option to operate this instrument without an enclosure, which is not an option with the CL31, CL51 and CHM 15k, which are permanently integrated with their enclosure. Another reason is that the enclosure has a scanning head (pointing vertically in our case studies). We clarified this in the text (P7L23, P24L3, P24L12).

11. 7/2 (and later): when giving the number of vertical levels it should be added how many levels are within the range of the ALCs, e.g. CL31. And the typical vertical resolution in this range should be mentioned.

The number of model levels in the instrument range is generally dependent on the geographical location due to orography. We added a table listing the number of levels for each of the three sites and instruments (Table 7) and a reference to this table in the text (P22L12, P21L16).

12. 7/10: What does single level mean?

We replaced this with "surface level" (P8L15).

13. 8/7: What is meant by horizontally homogeneous? The lidar equation is one-dimensional (along the pointing of the laser beam, typically vertically).

We removed this part of the sentence because it is not strictly necessary. Horizontally homogeneous atmosphere assumption is commonly used in radiation transfer schemes and it simply means the atmospheric variables are a function of height only and horizontal dependence is neglected. This could be applicable with a lidar simulator if pointing at an off-zenith angle.

14. 8/21: usually lower : I would rather write much smaller . The backscatter coefficient of (low) clouds is several orders of magnitude larger than Rayleigh scattering. The lidar ratio of Rayleigh-scattering should be mentioned.

In our experience this is not the case. The molecular volume backscattering coefficient is about  $1.54 \times 10^{-6} \text{ m}^{-1} \text{ sr}^{-1}$  near the surface at 532 nm (Fig. 8)). Clouds in our dataset produce backscattering from about  $2 \times 10^{-6} \text{ m}^{-1} \text{ sr}^{-1}$ .

15. 12/31 ff: I don't know if it is necessary to consider calibration constant from past studies . Most ceilometers show calibration constants that are not really constant (temperature-dependent), and maybe some instruments have been set to a higher/lower sensitivity on purpose. If ALCF1.0 offers the option to calibrate signals (Rayleigh or cloud method) it is not required to rely on previous measurements anyway.

We include these values as default both as a first approximation and for the convenience of users who do not require a high accuracy of calibration (e.g. if they do not want to compare observations with a model). The default

values are probably still relatively good as a first approximation, Fig. 8 and 13 in Hopkin et al. (2019) show that most instruments have a calibration coefficient within an order of magnitude or less.

16. 14/24: It is certainly a good idea to define a absolute threshold for the cloud detection. However,  $2 \times 10^{-6} \text{ m}^{-1} \text{ sr}^{-1}$  corresponds to an extinction coefficient of  $0.04 \text{ km}^{-1}$  (assuming a lidar ratio = 20 sr), I would have expected a larger value. Maybe a few additional comments on the magnitude can be added (typical values for low water clouds).

$2 \times 10^{-6} \text{ m}^{-1} \text{ sr}^{-1}$  was a value which was the best compromise for false positives and misses in our datasets, which were affected by relatively little aerosol. Users are advised to tune the threshold for their situation. In our datasets some clouds produced backscattering below  $2 \times 10^{-6} \text{ m}^{-1} \text{ sr}^{-1}$ . We added text explaining this point (P21L6). Temporal and vertical subsampling (5-min and 50 m, respectively, by default in the ALCF), removal of molecular backscattering and 5 standard deviations of noise before cloud detection reduce the number of false detections even at a relatively low detection threshold.

17. 14/31: **simulated backscatter too crudely** : see comment 7/2.

We added reference to the new table (P21L16).

18. 15/25: **Only the altitude of Cass is given here. To be consistent the altitude of the other stations should be given as well in this section.**

We added altitude to the description of the two other stations (P24L1, P24L5).

19. 16/12: **Typo in unit.**

We changed the units to  $\text{m}^{-1} \text{sr}^{-1}$  (P24L20).

20. 18/3 ff: **When discussing Figs. 5–7, the panels of the lidar ratio are ignored. They are also not explained in the figure captions. What is the meaning of these panels? Can they be omitted?**

We replaced the lidar ratio panels with model cloud liquid water and ice mixing ratios and cloud fraction. We kept the lidar ratio panels under the observed attenuated volume backscattering coefficient (Fig. 4a, 5a, 6a). We also replaced it with effective lidar ratio. Previously true lidar ratio was calculated under the assumption of multiple scattering coefficient 0.7, but cannot be universally assumed for ice crystals, which are supported in the revised lidar simulator. In the revised manuscript, lidar ratio is discussed in Sect. 4.1. In addition, calibration using lidar ratio was already discussed in Sect. 5.2. We also added a paragraph at the end of Sect. 7.4 discussing the lidar ratio panels (P31L2). They are useful for determining if the attenuated volume backscattering coefficient is properly calibrated.

21. 18/5: **What is favourable ? Is it meant that individual intercomparisons of CBH between ALC and models are more or less impossible due to the different spatio-temporal resolution. So the focus must be on climatologies, and the provision of the true CBH and vertical cloud distribution for model-validation.**

As we note above, cloud base height is not evaluated in the manuscript but rather the cloud mask. By "favourable" we mean that even models with high spatiotemporal resolution (ERA5, UM) struggle to simulate clouds correctly on hourly time scales. Comparison of individual daily profiles is definitely possible and can uncover deficiencies in models (Kuma, 2020, Chapter 4). For models where only low resolution output is available (AMPS, JRA-55, MERRA-2), comparison of climatologies may be preferable. The individual daily profiles generated from instantaneous model output should still be comparable with observations if the limitations due to resolution are taken into consideration. The SCOPS subcolumn generator ensures that such a comparison is statistically unbiased.

22. 18/27: **Fig. 5l does not exist (typo). More examples like this are existing.**

We fixed the references to Fig. 5–7.

23. 20/32: **relatively poorly** . Here it should be stressed that the sample is limited, thus general conclusions might be difficult.

We added "in our time-limited local comparison" (P35L11).

24. 42: Figure caption should be revised: One description of day and night is sufficient.

We removed the definition.

## **Anonymous Referee #2**

### **General comments**

The paper deals primarily with comparing cloud distribution, the cloud base height and cloud layers from observations and models. However, the influence of aerosol on radiative transfer has to be considered in any case if quantitative statements are to be made. For the investigation of only the mentioned cloud parameters a calibration is not absolutely necessary except for the harmonisation of the data and to find a common cloud detection threshold. However, with a view to a next version of the framework, the already included parts such as multiple scattering, Mie and Rayleigh scattering calculations and the calibration itself provide already the basis to include the option of investigating aerosol optical properties as well.

The authors should emphasize that the focus in this version of the framework is on clouds and not on aerosol which has influence on the attenuated backscatter of observations and of simulated lidar signals.

As we note above in our response to Referee 1, we added clarifying text to the abstract and Sect. 4 (P1L20, P9L17).

Section 7: This is a rather long section and should be divided into subsections dealing with the different figures 4, 5-7, 8, 9-11.

As we note above in our response to Referee 1, we split this section into multiple subsections.

Fig. 5 it seems that for some reanalyses and models the cloud layer near the ground seem to fully attenuate the signal and no clouds above can be detected whereas higher clouds are visible in the observation. This is especially true for the models and reanalyses with very coarse resolution. However, the reanalyses and models may have calculated the clouds, but they are not visible in the simulated lidar signal. Have you compared the observed signal also with cloud masks or cloud fraction of the models? Is the picture different if you would plot other subcolumns than the first one?

As we note above in our response to Referee 1, the revised Fig. 5-7 contain panels with model cloud liquid water and ice mixing ratios and cloud fraction. If obscuration by low clouds happens in the model but not in observations, it means that the low clouds in the model are too opaque or cover too large area.

The lidar ratio subplots in Fig. 5-7 can be omitted since they don't appear neither in the text nor in the caption.

As we note above in our response to Referee 1, we replaced the lidar ratio panels with model cloud liquid water and ice mixing ratio and cloud fraction.

Fig. 8: Aerosol would have influence on the lidar ratio determined with the applied calibration method in this paper as stated on P.15 L15. Aerosol in the layers below the cloud would lead to an increase of the determined lidar ratio (decrease of LR -1) what can be seen in Fig. 8. Because aerosol is not considered in the used models here, they are not affected. Could you please comment on how you screened the data for aerosol and is its influence somehow considered in your calculations?

As we note in our response to Referee 1, we decided to remove Fig. 8 and the accompanying discussion because of the potentially systematic bias caused by the lack of aerosol in the lidar simulator.

### **Specific comments**

1. P.2 L.4: ash is particulate matter as well. Thus, just aerosol is sufficient.

We removed this word (P2L2).

2. P.2 L.8: aerosol optical depth is an integrated quantity relying on the lidar ratio. ALCs are used to determine the particle backscatter coefficient: Change to particle backscatter coefficient

As noted in our response to Referee 1, we replaced this with "particle volume backscattering coefficient" (P2L6).

3. P.2 L.15: EARLINET is not a network of ALCs. Most of the instruments are not operated autonomously. The measurements must meet stringent quality control criteria and are performed for selected times. A better example would be PollyNET:

Baars et al.: An overview of the first decade of PollyNET: an emerging network of automated Raman-polarization lidars for continuous aerosol profiling, <https://doi.org/10.5194/acp-16-5111-2016>

We removed EARLINET and included PollyNET (P2L15).

4. P.6 L.19: instrument can be housed in a protective enclosure : This is not only true for the MiniMPL systems but also for the other ALCs to allow a 24/7 operation at any location.

As we note above in our response to Referee 1, we mention this because the enclosure is optional with the MiniMPL and contains a scanning head. This is clarified in the revised manuscript (P7L23, P24L3, P24L12).

5. P.8 L.10: As already mentioned, it should always be emphasized that the focus here is on clouds and that aerosol is not taken into account. This however would be necessary to perform the correct radiative transfer calculations in order to appropriately simulate the lidar signal.

As we note above in our response to Referee 1, aerosol is unlikely to affect cloud evaluation due to the relatively low attenuation caused by aerosol compared to the strong backscattering from clouds. Aerosol simulation is only possible with those models which provide aerosol concentration. We added text clarifying that aerosol is currently not implemented in the simulator (P1L20, P9L17).

6. P.13 L.1 and 2: Is the calibration constant a unitless quantity?

The units depend on the units of backscattering recorded by the instrument. The calibration coefficient translates between arbitrary units of the instrument and the units of the attenuated volume backscattering coefficient ( $\text{m}^{-1}\text{sr}^{-1}$ ). We added an explanation (P18L20).

7. P.14 L.24: The cloud detection threshold of  $2 \times 10^{-6} \text{ m}^{-1} \text{ sr}^{-1}$  corresponds to a particle extinction coefficient of  $0.1 \text{ km}^{-1}$  if assuming a lidar ratio of 50 sr which is not unusual for aerosol. This threshold seems to me rather low and could lead to misclassification of aerosol as clouds. This is even more true in the case of the MiniMPL system which is operated at 532 nm. Could you please comment on the observed magnitudes within liquid and cirrus clouds? Why is no adaption to the wavelength needed? Would it make sense to apply a height dependent threshold with lower values at higher altitudes to account for the lower attenuated backscatter of cirrus clouds?

As we note in our response to Referee 1, the threshold of  $2 \times 10^{-6} \text{ m}^{-1} \text{ sr}^{-1}$  was the best compromise between false positives and misses in our dataset. Users are advised to choose a threshold suitable for their situation if the aerosol contamination is higher for example. We clarified this in the text (P21L6).

In the revised version we implemented the option to couple observed backscattering with simulated backscattering and remove molecular backscattering from the observed attenuated volume backscattering coefficient before cloud detection is applied (P21L3, P24L24). This improves the results, especially with the MiniMPL which is affected by the largest amount of molecular backscattering at 532 nm.

Our dataset does not show a strong dependence of cloud attenuated volume backscattering coefficient on height to justify a height-dependent threshold, especially if molecular backscattering is removed prior to cloud detection.

8. P.16 L.11: Please check the use of the different terms attenuated backscatter , total attenuated backscatter , total volume backscatter coefficient and particle backscatter coefficient in the text, the figure captions



and labels. The measured range corrected lidar signal calibrated with the calibration constant is normally referred to as the attenuated backscatter.

As we note above in our response to Referee 1, we replaced the occurrences of "backscatter" with "attenuated volume backscattering coefficient".

9. P.18 L.17, 19, 27 and 29: Please check figure labels for Fig. 5 in the text. There are no subfigures g, h, k, i, l.

We fixed the references to Fig. 5–7.

10. P.19 L.29: Again, as already mentioned, aerosol can not be neglected. Especially when using the molecular backscatter to determine the calibration coefficient, aerosol must be considered. As you can see in Fig. 9 and 10 the streak caused by molecular backscatter is too broad to retrieve an accurate calibration constant. This is mainly caused by aerosol and varying atmospheric conditions (temperature, pressure). I would omit here and in P.20 L.10 the proposal of a new method.

We disagree that the molecular backscatter is too broad to retrieve an accurate calibration coefficient. For the purpose of cloud detection, a calibration coefficient accuracy better than a factor of 2 is still satisfactory, which is within the uncertainty range in Fig. 9 and 10 for at least the MiniMPL, where the peak is centred between 1 and  $2 \times 10^{-6} \text{ m}^{-1} \text{ sr}^{-1}$  (Fig. 9e, i; Fig. 10e). In a more polluted atmospheric conditions it may not be possible to clearly identify the streak, but in general this method may still be useful in conjunction with the stratocumulus method, especially if a dataset does not contain suitable cases of stratocumulus clouds.

11. P.20 L.9: Fig. 10 instead of Fig. 9

We replaced the reference with the correct figure (Fig. 8 in the revised manuscript).

12. P.20 L.21 and 24: You are only addressing the difference between day and night here for the MiniMPL instrument. However, it should be possible to make the same observation with the other devices. It would be interesting to have the difference between day and night for the other instruments as well.

We added night and day for CL51, CL31 and CHM 15k. They differ relatively little due to low incoming solar radiation at near infrared wavelengths compared to 532 nm. We updated the text accordingly (P34L10, P35L27).

13. P.21 L.11-14: see comment above

We expanded the text to describe the nighttime and daytime results for CL51, CL31 and CHM 15k (P34L10).

14. Fig. 9, 10 and 11: Check caption and labels for consistent use of the term attenuated backscatter (comment P.16 L.11)

As we note above in our response to Referee 1, we replaced the occurrences of "backscatter" with "attenuated volume backscattering coefficient".

15. Supplement: Some links to the plots in the tutorial seem not to be working. In the online version it is correct.

We fixed this in the v1.0.0-beta.3 release.

## References

- Chan, K. L., Wiegner, M., Flentje, H., Mattis, I., Wagner, F., Gasteiger, J., and Geiß, A.: Evaluation of ECMWF-IFS (version 41R1) operational model forecasts of aerosol transport by using ceilometer network measurements, *Geoscientific Model Development*, 11, 3807–3831, <https://doi.org/10.5194/gmd-11-3807-2018>, 2018.
- Garnier, A., Pelon, J., Vaughan, M. A., Winker, D. M., Trepte, C. R., and Dubuisson, P.: Lidar multiple scattering factors inferred from CALIPSO lidar and IIR retrievals of semi-transparent cirrus cloud optical depths over oceans, *Atmospheric Measurement Techniques*, 8, 2759–2774, <https://doi.org/10.5194/amt-8-2759-2015>, 2015.

- Hansen, A., Ament, F., Grützun, V., and Lammert, A.: Model evaluation by a cloud classification based on multi-sensor observations, *Geoscientific Model Development Discussions*, 2018, 1–20, <https://doi.org/10.5194/gmd-2018-259>, 2018.
- Hogan, R. J., Jakob, C., and Illingworth, A. J.: Comparison of ECMWF Winter-Season Cloud Fraction with Radar-Derived Values, *Journal of Applied Meteorology*, 40, 513–525, [https://doi.org/10.1175/1520-0450\(2001\)040<0513:COEWSC>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<0513:COEWSC>2.0.CO;2), 2001.
- Hopkin, E., Illingworth, A. J., Charlton-Perez, C., Westbrook, C. D., and Ballard, S.: A robust automated technique for operational calibration of ceilometers using the integrated backscatter from totally attenuating liquid clouds, *Atmospheric Measurement Techniques Discussions*, 2019, 1–28, <https://doi.org/10.5194/amt-2018-427>, 2019.
- Illingworth, A., Hogan, R., O’connor, E., Bouniol, D., Brooks, M., Delanoë, J., Donovan, D., Eastment, J., Gaussiat, N., Goddard, J., et al.: Cloudnet: Continuous evaluation of cloud profiles in seven operational models using ground-based observations, *Bulletin of the American Meteorological Society*, 88, 883–898, <https://doi.org/10.1175/BAMS-88-6-883>, 2007.
- Kuma, P.: Comparing remotely sensed observations of clouds and aerosols in the Southern Ocean with climate model simulations, Ph.D. thesis, University of Canterbury, Christchurch, New Zealand, <https://doi.org/10.5281/zenodo.3865850>, URL <https://doi.org/10.5281/zenodo.3865850>, submitted, 2020.
- Lamer, K., Fridlind, A. M., Ackerman, A. S., Kollias, P., Clothiaux, E. E., and Kelley, M.: (GO)2-SIM: a GCM-oriented ground-observation forward-simulator framework for objective evaluation of cloud and precipitation phase, *Geoscientific Model Development (Online)*, 11, <https://doi.org/10.5194/gmd-11-4195-2018>, 2018.
- Morcrette, C. J., O’Connor, E. J., and Petch, J. C.: Evaluation of two cloud parametrization schemes using ARM and Cloud-Net observations, *Quarterly Journal of the Royal Meteorological Society*, 138, 964–979, <https://doi.org/10.1002/qj.969>, 2012.
- O’Connor, E. J., Illingworth, A. J., and Hogan, R. J.: A technique for autocalibration of cloud lidar, *Journal of Atmospheric and Oceanic Technology*, 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.
- Warren, E., Charlton-Perez, C., Kotthaus, S., Lean, H., Ballard, S., Hopkin, E., and Grimmond, S.: Evaluation of forward-modelled attenuated backscatter using an urban ceilometer network in London under clear-sky conditions, *Atmospheric Environment*, 191, 532 – 547, <https://doi.org/https://doi.org/10.1016/j.atmosenv.2018.04.045>, 2018.

# Ground-based lidar processing and simulator framework for comparing models and observations (ALCF 1.0)

Peter Kuma<sup>1</sup>, Adrian J. McDonald<sup>1</sup>, Olaf Morgenstern<sup>2</sup>, Richard Querel<sup>3</sup>, Israel Silber<sup>4</sup>, and Connor J. Flynn<sup>5</sup>

<sup>1</sup>University of Canterbury, Christchurch, New Zealand

<sup>2</sup>National Institute of Water & Atmospheric Research (NIWA), Wellington, New Zealand

<sup>3</sup>National Institute of Water & Atmospheric Research (NIWA), Lauder, New Zealand

<sup>4</sup>Department of Meteorology and Atmospheric Science, Pennsylvania State University, PA, USA

<sup>5</sup>School of Meteorology, University of Oklahoma, Norman, OK, USA

**Correspondence:** Peter Kuma (peter@peterkuma.net)

**Abstract.** Automatic lidars and ceilometers provide valuable information on cloud and aerosols, but have not been used systematically in the evaluation of GCMs and NWP models. Obstacles associated with the diversity of instruments, a lack of standardisation of data products and open processing tools mean that the value of the large ALC networks worldwide is not being realised. We discuss a tool, called the Automatic Lidar and Ceilometer Framework (ALCF), that overcomes these problems and also includes a ground-based lidar simulator, which calculates the radiative transfer of laser radiation, and allows one-to-one comparison with models. Our ground-based lidar simulator is based on the Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (COSP) which has been used extensively for spaceborne lidar intercomparisons. The ALCF implements all steps needed to transform and calibrate raw ALC data and create simulated attenuated volume backscattering coefficient profiles for one-to-one comparison and complete statistical analysis of cloud. The framework supports multiple common commercial ALCs (Vaisala CL31, CL51, Lufft CHM 15k and Droplet Measurement Technologies MiniMPL), reanalyses (JRA-55, ERA5 and MERRA-2) and models (AMPS and the Unified Model). To demonstrate its capabilities, we present case studies evaluating cloud in the supported reanalyses and models using CL31, CL51, CHM 15k and MiniMPL observations at three sites in New Zealand. We show that the reanalyses and models generally underestimate cloud fraction. If sufficiently high temporal resolution model output is available (better than 6 hourly), a direct comparison of individual clouds is also possible. We demonstrate that the ALCF can be used as a generic evaluation tool to examine cloud occurrence and cloud properties in reanalyses, NWP models and GCMs, potentially utilising the large amounts of ALC data already available. This tool is likely to be particularly useful for the analysis and improvement of low-level cloud simulations which are not well monitored from space. This has previously been identified as a critical deficiency in contemporary models, limiting the accuracy of weather forecasts and future climate projections. While the current focus of the framework is on clouds, support for aerosol in the lidar simulator is planned in the future.

## 1 Introduction

Automatic lidars and ceilometers (ALC) are active ground-based instruments which emit laser pulses in the ultraviolet, visible or infrared (IR) part of the electromagnetic spectrum and measure radiation backscattered from atmospheric constituents such as cloud and fog liquid droplets and ice crystals, haze, aerosol and atmospheric gases (Emeis, 2010). Vertical profiles of attenuated backscattered radiation can be produced by measuring received power as a function of time elapsed between emitting the pulse and receiving the backscattered radiation. Quantities such as cloud base height (CBH) and a cloud mask (Pal et al., 1992; Wang and Sassen, 2001; Martucci et al., 2010; Costa-Surós et al., 2013; Van Tricht et al., 2014; Liu et al., 2015a, b; Lewis et al., 2016; Cromwell and Flynn, 2018; Silber et al., 2018), particle volume backscattering coefficient (Marenco et al., 1997; Welton et al., 2000, 2002; Wiegner and Geiß, 2012; Wiegner et al., 2014; Jin et al., 2015; Dionisi et al., 2018) and boundary layer height (Eresmaa et al., 2006; Münkler et al., 2007; Emeis et al., 2009; Tsaknakis et al., 2011; Milroy et al., 2012; Knepp et al., 2017) can be derived from the attenuated volume backscattering coefficient profile. Lidars equipped with polarisation or multiple wavelengths can also provide depolarisation ratio or colour ratio, respectively, which can be used to infer cloud phase or particle types. Doppler lidars can measure wind speed in the direction of the lidar orientation. ALCs are commonly deployed at airports, where they provide CBH, fog and aerosol observations needed for air traffic control. Large networks of up to hundreds of lidars and ceilometers have been deployed worldwide: Cloudnet (Illingworth et al., 2007), E-PROFILE (Illingworth et al., 2018), PollyNET (Baars et al., 2016), ICENET (Cazorla et al., 2017), MPLNET (Welton et al., 2006) and ARM (Stokes and Schwartz, 1994; Campbell et al., 2002). The purpose of these networks is to observe cloud, fog, aerosol, air quality, visibility and volcanic ash, provide input to numerical weather prediction (NWP) model evaluation (Hogan et al., 2001; Illingworth et al., 2007; Morcrette et al., 2012; Warren et al., 2018; Lamer et al., 2018; Hansen et al., 2018b) and assimilation (Illingworth et al., 2015b, 2018) and for climate studies. These networks are usually composed of multiple types of ALCs, with Vaisala CL31, CL51, Lufft (formerly Jenoptik) CHM 15k and Droplet Measurement Technologies (formerly Sigma Space and Hexagon) MiniMPL the most common. Complex lidar data processing has been set up on some of these networks. Notably, at the SIRTAsite in France, lidar ratio (LR) comparable with a lidar simulator (Chiriaco et al., 2018) is calculated as part of their "ReOBS" processing method. Intercomparison and calibration campaigns such as CeiLinEx2015 (Mattis et al., 2016) and INTERACT-I(-II) (Rosoldi et al., 2018; Madonna et al., 2018) have been performed. Lidar data processing involves a number of tasks such as resampling, calibration, noise removal and cloud detection. Some of these are implemented in the instrument firmware of ALCs. This, however, means that lidar attenuated volume backscattering coefficient and detected cloud and cloud base are not comparable between different instruments. In most cases the algorithms are not publicly documented, making it impossible to compare the data with values from a model or a lidar simulator without a systematic bias.

Atmospheric model evaluation is an ongoing task, and a critical part of the model improvement process (Eyring et al., 2019; Hourdin et al., 2017; Schmidt et al., 2017). Traditionally, various types of observational and model datasets have been utilised – weather and climate station data, upper air soundings, ground-based and satellite remote sensing datasets, and high-resolution model simulations, amongst others. Clouds are one of the most problematic phenomena in atmospheric models due to their transient nature, high spatial and temporal variability, and sensitivity to a complex combination of conditions such as relative

humidity, aerosols (presence of cloud condensation nuclei and ice nuclei), thermodynamic and dynamic conditions. At the same time, clouds have a very substantial effect on the atmospheric shortwave and longwave radiation balance, and any cloud misrepresentation has a strong effect on other components of the model, limiting the ability to accurately represent past and present climate and predict future climate (Zadra et al., 2018). An improved understanding of clouds and cloud feedbacks is one of the focuses of the Coupled Model Intercomparison Project Phase 6 (CMIP6) (Eyring et al., 2016), and comparison of model cloud with observations is one of the key points of the Cloud Feedback Model Intercomparison Project (CFMIP) (Webb et al., 2017). Satellite observations make up the majority of the data used to evaluate model clouds. These include passive visible and IR low earth orbit and geostationary radiometers measuring, among others, features such as cloud cover, cloud top height (CTH), cloud top temperature; passive microwave instruments measuring total column water; active radars and lidars measuring cloud vertical profiles. Ground-based remote sensing instruments include radars, lidars, ceilometers, radiometers and sky cameras. As pointed out by Williams and Bodas-Salcedo (2017), using a wide range of different observational datasets including satellite and ground-based for general circulation model (GCM) evaluation is important due to limitations of each dataset.

Model cloud is commonly represented by the mixing ratio of liquid and ice and cloud fraction (CF) on every model grid cell and vertical level. In addition some models provide the cloud droplet effective radius used in radiative transfer calculations. Remote sensing observations do not match the representation of the atmospheric model fields directly because of their different resolutions, limited field of view (FOV) and attenuation by atmospheric constituents before reaching the instrument's receiver. Instrument simulators bridge this gap by converting the model fields to quantities which emulate those measured by the instrument, which can then be compared directly with observations. One such collection of instrument simulators is the CFMIP Observation Simulator Package (COSP) (Bodas-Salcedo et al., 2011; Swales et al., 2018), which has been used for more than a decade for evaluation of models using satellite, and more recently ground-based, observations. The simulators in COSP include active instruments such as spaceborne and ground-based radars: Cloud Profiling Radar (CPR) on CloudSat (Stephens et al., 2002), Ka-band ARM Zenith Radar (KAZR); lidars: Cloud-Aerosol Lidar Orthogonal Polarization (CALIOP) on CALIPSO (Winker et al., 2009), Cloud-Aerosol Transport System (CATS) on ISS (McGill et al., 2015), the Atmospheric Lidar (ATLID) on EarthCARE (Illingworth et al., 2015a); and spaceborne passive instruments: ISCCP (Rossow and Schiffer, 1991), MODIS (Parkinson, 2003) and MISR (Diner et al., 1998). The more recent addition of ground-based radar (Zhang et al., 2018) and lidar (Chiriaco et al., 2018; Bastin et al., 2018) opens up new possibilities to use the large amount of remote sensing data obtained from ground-based active remote sensing instruments. In practice, ground-based observational remote sensing data are not straightforward to use without a substantial amount of additional processing. Some previous studies have also compared models and ground-based radar and lidar observations without the use of an instrument simulator (Bouniol et al., 2010; Hansen et al., 2018a), though for the reasons identified above this is not advisable.

In this study we introduce a software package called the Automatic Lidar and Ceilometer Framework (ALCF) for evaluating model cloud using ALC observations. It extends and integrates the COSP lidar simulator (Chiriaco et al., 2006; Chepfer et al., 2007, 2008) with pre- and post-processing steps, and allows the simulator to be run offline on model output, instead of having to be integrated inside the model. This makes it possible to compare ALC data at any location without having to run the model

with a specific configuration. Multiple ALCs, reanalyses and model output formats are supported. The original COSP lidar simulator was extended with Rayleigh, Mie and ice crystal scattering at multiple lidar wavelengths. Observational ALC data from a number of common instruments can be processed by re-sampling to a common resolution, removing noise, detecting cloud and calculating statistics. The same steps can be performed on the simulated lidar data from the model (the output of running COSP on the model data), allowing for one-to-one comparison of model and observations. A particular focus of our work was on applying the same processing steps on the observed and simulated attenuated volume backscattering coefficient in order to avoid biases. The ALCF is made available under an open source license (MIT) at <https://alcf-lidar.github.io>, and as a permanent archive of code and technical documentation on Zenodo at <https://doi.org/10.5281/zenodo.4088217>.

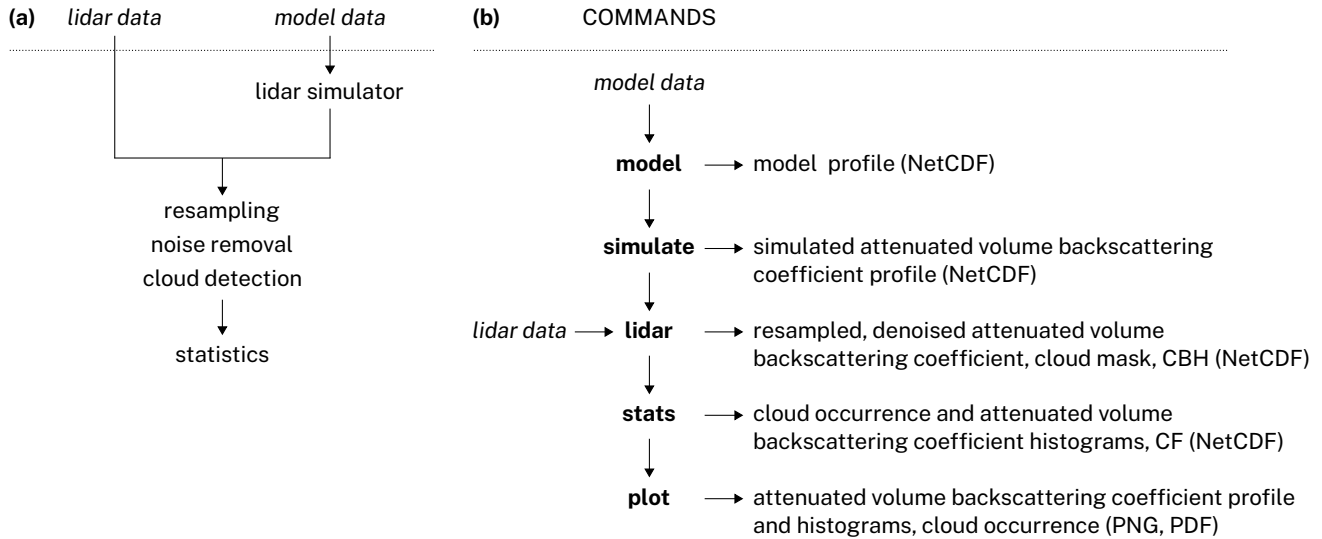
A relatively small amount of other open source code is available for ALC data processing. A lidar simulator has been developed as part of the Goddard Satellite Data Simulator Unit (G-SDSU) (G-S, 2019), a package based on the instrument simulator package SDSU (Masunaga et al., 2010). The Community Intercomparison Suite (CIS) (Watson-Parris et al., 2016) allows for subsetting, aggregation, co-location and plotting of mostly satellite data with a focus on model–observations intercomparison. The STRAT lidar data processing tools are a collection of tools for conversion of raw ALC data, visualisation and feature classification (Morille et al., 2007).

Here, we provide an overview of the ALCF (Sect. 2), describe the supported ALCs, reanalyses and models (Sect. 3), the lidar simulator (Sect. 4) and the observed and simulated lidar data processing steps (Sect. 5). Later, we present a set of case studies at three sites in New Zealand (NZ) (Sect. 6) to demonstrate the value of this new tool. Lastly, we present the results of the case studies in Sect. 7.

## 2 Overview of operation of the Automatic Lidar and Ceilometer Framework (ALCF 1.0)

The ALCF performs the necessary steps to simulate ALC attenuated volume backscattering coefficient based on 4-dimensional atmospheric fields from reanalyses, NWP models and GCMs, and to transform the observed raw ALC attenuated volume backscattering coefficient profiles to profiles comparable with the simulated profiles. It does so by extracting 2-dimensional (time  $\times$  height) profiles from the model data, performing radiative transfer calculations based on a modified COSP lidar simulator (Sect. 4), absolute calibration and resampling of the observed attenuated volume backscattering coefficient to common resolution and performing comparable cloud detection on the simulated and observed attenuated volume backscattering coefficient. The framework supports multiple common ALCs (Sect. 3.1), reanalyses and models (Sect. 3.2). The schematic in Fig. 1 illustrates this process as well as the ALCF commands which perform the individual steps. The following commands are implemented: **model**, **simulate**, **lidar**, **stats** and **plot**. The commands are normally executed in a sequence, which is also implemented by a meta-command **auto**, which is equivalent to executing a sequence of commands. The commands are described in detail in the technical documentation available online at <https://alcf-lidar.github.io>, on Zenodo at <https://doi.org/10.5281/zenodo.4088217> and in the Supplementary information. The physical basis is described here.

The **model** command extracts 2-dimensional profiles of cloud liquid and ice content (and other thermodynamic fields) from the supported NWP model, GCM and reanalysis data (*model data* in Fig. 1) at a geographical point, along a ship track or a



**Figure 1.** (a) Scheme showing the operation of the ALCF and (b) the processing commands.

flight path. The resulting profiles are recorded as NetCDF files. Sect. 3.2 describes the supported reanalyses and models. The *model data* can be either in one of the supported model output formats, or a new module for reading arbitrary model output can be written, providing that the required atmospheric fields are present in the model output. The required model fields are: per-level specific cloud liquid water content, specific cloud ice water content, cloud fraction, geopotential height, temperature, surface-level pressure and orography. No physical calculations are performed by this command. The atmospheric profiles are extracted by a nearest-neighbour selection.

The **simulate** command runs the lidar simulator described in Sect. 4 on the extracted model data (the output of the **model** command) and produces simulated attenuated volume backscattering coefficient profiles. This command runs the COSP-derived lidar simulator, which performs radiative transfer calculations of the laser radiation through the atmosphere. The resulting simulated attenuated volume backscattering coefficient profiles are the output of this command.

The **lidar** command applies various processing algorithms on either the simulated attenuated volume backscattering coefficient (the output of the **simulate** command) or the observed ALC coefficient (*lidar data* in Fig. 1) (Sect. 5). The data are resampled to increase the signal-to-noise ratio (SNR), noise is subtracted, LR is calculated, a cloud mask is calculated by applying a cloud detection algorithm and CBH is determined from the cloud mask. Absolute calibration (Sect. 5.2) can also be applied in this step by multiplying the observed attenuated volume backscattering coefficient by a calibration coefficient. This is important in order to obtain unbiased attenuated volume backscattering coefficient profiles comparable with the simulated profiles. Sect. 3.1 describes the supported instruments. The *lidar data* can be in one of the supported instrument formats. If the native instrument format is not NetCDF, it has to be converted from the native format with the auxiliary command **convert** or one of the conversion programs: *c12nc* (Vaisala CL31, CL51), *mpl2nc* or *SigmaMPL* (Sigma Space MiniMPL).

**Table 1.** Table of ALCs and their technical parameters. Power is calculated as Pulse  $\times$  Pulse Repetition Frequency (PRF).

Instrument	$\lambda$ (nm)	Laser	Rate <sup>1</sup> (s)	Res. <sup>2</sup> (m)	Depol. <sup>3</sup>	Pulse <sup>4</sup> ( $\mu$ J)	Range <sup>5</sup> (km)	PRF (kHz)	Overlap <sup>6</sup> (m)	Power (mW)	FOV <sup>7</sup> ( $\mu$ rad)
CHM 15k	1064	Nd:YAG	2–600	5	no	7–9	15.4	5–7	1000 <sup>8</sup>	48	450
CL31	910	InGaAs	2–120	10	no	1.2	7.7	10	70 <sup>8</sup>	12	830
CL51	910	InGaAs	6–120	10	no	3	15.4	6.5	230 <sup>9</sup>	20	560
MiniMPL	532	Nd:YAG	1–900	5–75	yes	3–4	30.0	2.5	2000 <sup>9</sup>	9	110

<sup>1</sup>Sampling rate. <sup>2</sup>Vertical (range) resolution. <sup>3</sup>Depolarisation. <sup>4</sup>Pulse energy. <sup>5</sup>Maximum range. <sup>6</sup>Range of full overlap. <sup>7</sup>Receiver field of view. <sup>8</sup>Hopkin et al. (2019). <sup>9</sup>Madonna et al. (2018).

The **stats** step calculates summary statistics from the output of the **lidar** command. These include CF, cloud occurrence by height, attenuated volume backscattering coefficient histograms and the averages of LR and backscattering coefficient.

The **plot** command plots attenuated volume backscattering coefficient profiles produced by the **lidar** command (Fig. 4, 5, 6), and the statistics produced by the **stats** command: cloud occurrence (Fig. 3), attenuated volume backscattering coefficient histograms (Fig. 7) and attenuated volume backscattering coefficient noise standard deviation histograms (Fig. 9).

### 3 Supported input data: instruments, reanalyses and models

#### 3.1 Instruments

The primary focus of the framework is to support common commercial ALCs. Ceilometers are considered the most basic type of lidar (Emeis, 2010; Kotthaus et al., 2016) intended as commercial products designed for unattended operation. They are used routinely to measure CBH, but most instruments also provide the full vertical profiles of attenuated volume backscattering coefficient. Therefore, they are suitable for model evaluation by comparing not only CBH, but also cloud occurrence as a function of height. Their compact size and low cost make it possible to deploy a large number of these instruments in different locations, or use them in unusual settings such as mounted on ships (Klekociuk et al., 2019; Kuma et al., 2020). Common off-the-shelf ceilometers are the Lufft CHM 15k, Vaisala CL31 and CL51. Some lidars offer higher power and therefore higher SNR, and capabilities not present in ceilometers such as dual polarisation, multiple wavelengths, Doppler shift measurement and Raman scattering. Below we describe ALCs supported by the framework and used in our case studies: Lufft CHM 15k, Vaisala CL31 and CL51 and Droplet Measurement Technologies MiniMPL. Table 1 lists selected parameters of the supported ALCs.

*Lufft CHM 15k* (previously Jenoptik CHM 15k) is a ceilometer operating at a wavelength of 1064 nm (near IR). The maximum range of the instrument is 15.4 km, vertical sampling resolution 5 m in the first 150 m and 15 m above and sampling rate 2 s. The total number of vertical levels is 1024. The wavelength in the near IR spectrum ensures low molecular backscattering. The instrument produces NetCDF files containing uncalibrated attenuated volume backscattering coefficient profiles and various derived variables, although the calibration coefficient is relatively consistent for different instruments of the model (Hopkin et al., 2019, Fig. 13).



**Table 2.** Reanalyses and models used in the case studies and some of their main properties. The temporal and horizontal grid resolution and vertical levels listed is the resolution of the model output available. The horizontal grid resolution is determined at 45°S. The internal resolution of the model may be different (see Sect. 3.2 for details). The reanalyses and the UM use regular longitude-latitude grids, while the AMPS horizontal grid is regular in the South Pole stereographic projection.

Model/Grid	Type	Time resolution	Horizontal grid resolution	Vertical levels
AMPS/D01	NWP	3 h	$0.27^\circ \times 0.19^\circ$ (21×21 km)	60
ERA5	Reanalysis	1 h	$0.25^\circ \times 0.25^\circ$ (20×28 km)	37
JRA-55	Reanalysis	6 h	$1.25^\circ \times 1.25^\circ$ (98×139 km)	37
MERRA-2	Reanalysis	3 h	$0.625^\circ \times 0.50^\circ$ (49×56 km)	72
UM (GA7.1)/N96	GCM	20 min.	$1.875^\circ \times 1.25^\circ$ (147×139 km)	85

*Vaisala CL31 and CL51* are ceilometers operating at a wavelength of 910 nm (near IR). The maximum range of CL31 and CL51 is 7.7 km and 15.4 km and the sampling rate is 2 and 6 s, respectively. The vertical resolution is 10 m. The total number of vertical levels is 770 and 1540, respectively. The wavelength is characterised by relatively low molecular backscattering (but higher than 1064 nm) and is affected by water vapour absorption (Wiegner and Gasteiger, 2015; Wiegner et al., 2019), which can cause additional absorption of about 20% in the mid-latitudes and 50% in the tropics (see also Sect. 5.4). The instruments produce data files containing uncalibrated attenuated volume backscattering coefficient which can be converted to NetCDF (see cl2nc in the Code and data availability section). The firmware configuration option "noise\_h2 off" results in backscatter range correction to be selectively applied under a certain critical range and above this range only if cloud is present (Kotthaus et al., 2016, Sect. 3.2). This was the case with our case study dataset (Sect. 6). We apply range correction on the uncorrected range gates during lidar data processing. The critical range in CL51 is not documented, but was determined as 6000 m based on an observed discontinuity.

*Droplet Measurement Technologies Mini Micro Pulse Lidar (MiniMPL)* (previously Sigma Space MiniMPL and Hexagon MiniMPL) (Spinhirne, 1993; Campbell et al., 2002; Flynn et al., 2007) is a dual-polarisation micro pulse lidar (meaning that it uses a high pulse repetition rate (PRF) and low pulse power) operating at a wavelength of 532 nm (green colour in the visible spectrum). The maximum range of the instrument is 30 km. The vertical resolution is 5–75 m and sampling rate 1 s. The shorter wavelength is affected by stronger molecular backscattering than 910 nm and 1064 nm. The instrument can be housed in an enclosure with a scanning head to provide configurable scanning by elevation angle and azimuth. The instrument produces data files containing raw attenuated volume backscattering coefficient which can be converted to NetCDF containing normalised relative backscatter (NRB) with a vendor-provided tool SigmaMPL (see also mpl2nc in the Code and data availability section).

### 3.2 Reanalyses and models

Below we briefly describe reanalyses and models<sup>1</sup> used in the case studies presented here (Sect. 6). We used publicly available output from three reanalyses and one NWP model. In addition, we performed nudged GCM simulations with high-temporal resolution output with the Unified Model (UM). Table 2 lists some of the main properties of the reanalyses and models.

5     *The Antarctic Mesoscale Prediction System (AMPS)* (Powers et al., 2003) is a limited-area NWP model based on the polar fifth-generation Pennsylvania State University-National Center for Atmospheric Research Mesoscale Model (Polar MM5), now known as the Polar Weather Research and Forecasting (WRF) model (Hines and Bromwich, 2008). The model serves operational and scientific needs in Antarctica, but its largest grid also covers the South Island of NZ. AMPS forecasts are publicly available on the Earth System Grid (Williams et al., 2009). The forecasts are produced on several domains. The largest  
10 domain D01 used in the presented analysis covers NZ and has horizontal grid spacing of approximately 21 km over NZ. The model uses 60 vertical levels. The model output is available in 3-hourly intervals and initialised at 00:00 and 12:00 UTC. The initial and boundary conditions are based on the Global Forecasting System (GFS) global NWP model. AMPS assimilates local Antarctic observations from human-operated stations, automatic weather stations (AWS), upper-air stations and satellites.

*ERA5* (ECMWF, 2019) is a reanalysis produced by the European Centre For Medium-Range Weather Forecasts (ECMWF)  
15 currently available for the time period 1979 to present, with a plan to extend the time period to 1950. The reanalysis is based on the global NWP model Integrated Forecast System (IFS) version CY41R2. It uses a 4D-Var assimilation of station, satellite, radiosonde, radar, aircraft, ship-based and buoy data. The model has 137 vertical levels. Atmospheric fields are interpolated from horizontal resolution equivalent of 31 km and 137 model levels on regular longitude-latitude grid of  $0.25^\circ$  and 37 pressure levels, and made available to the end-users. In this analysis we use the hourly data on pressure and surface levels.

20     *Japanese 55-year reanalysis (JRA-55)* (Ebita et al., 2011; Kobayashi et al., 2015; Harada et al., 2016) is a global reanalysis produced by the Japan Meteorological Agency (JMA) and the Central Research Institute of Electric Power Industry (CRIEPI) based on the JMA Global Spectral Model (GSM). The reanalysis is available from 1958 onward. The reanalysis is based on the JMA operational assimilation system. JRA-55 uses a 4D-Var assimilation of surface, upper-air, satellite, ship-based and aircraft observations. The model uses 60 vertical levels and a horizontal grid with resolution approximately 60 km. In this analysis we  
25 use the  $1.25^\circ$  isobaric analysis and forecast fields interpolated to 37 pressure levels.

*Modern-Era Retrospective analysis for Research and Applications (MERRA-2)* (Gelaro et al., 2017) is a reanalysis produced by the NASA Global Modeling and Assimilation Office (GMAO). The reanalysis is based on the Goddard Earth Observing System (GEOS) atmospheric model. The model has approximately  $0.5^\circ \times 0.65^\circ$  horizontal resolution and 72 vertical levels. It performs 3D-Var assimilation of station, upper-air, satellite, ship-based and aircraft data in 6-hourly cycles. In this analysis,  
30 we use the MERRA-2 3-hourly instantaneous model-level assimilated meteorological fields (M2I3NVASM) version 5.12.4 product.

---

<sup>1</sup>We use the term "reanalysis" when referring to ERA5, JRA-55 and MERRA-2 even though the reanalyses are based on atmospheric models. We use the term "model" when referring to AMPS and the UM, which are atmospheric models.

*The UK Met Office Unified Model (UM)* (Walters et al., 2019) is an atmospheric model for weather forecasting and climate projection developed by the UK Met Office and the Unified Model Partnership. The UM is the atmospheric component, called Global Atmosphere (GA), of the HadGEM3–GC3.1 GCM and the UKESM1 earth system model (ESM). In this analysis we performed custom nudged runs of the UM (Telford et al., 2008) in the GA7.1 configuration with 20 min. time step and output  
5 temporal resolution on a New Zealand eScience Infrastructure (NeSI)/National Institute of Water & Atmospheric Research (NIWA) supercomputer (Williams et al., 2016). The model was nudged to the ERA-Interim (Dee et al., 2011) atmospheric fields of horizontal wind speed and potential temperature and the HadISST sea surface temperature (SST) and sea ice dataset (Rayner et al., 2003). The model uses 85 vertical levels and a horizontal grid resolution of  $1.875^\circ \times 1.25^\circ$ .

#### 4 Lidar simulator

10 The COSP lidar simulator Active Remote Sensing Simulator (ACTSIM) was introduced by Chiriaco et al. (2006) for the purpose of deriving simulated CALIOP measurements (Chepfer et al., 2007, 2008). The simulation is implemented by applying the lidar equation on model levels. Scattering and absorption by cloud particles and air molecules is calculated using the Mie and Rayleigh theory, respectively. Scattering and absorption by aerosols is not implemented in the presented version, but support is planned in the future for models which provide concentration of aerosols. Therefore, the current focus of the  
15 simulator is solely on cloud evaluation. CALIOP operates at a wavelength of 532 nm, and calculations in the original COSP simulator use this wavelength. We implemented a small set of changes to the lidar simulator to support a number of ALCs with different operating wavelengths and developed parametrisation of backscattering from ice crystals based on temperature.

The lidar equation (Emeis, 2010) is based on the radiative transfer equation (Goody and Yung, 1995; Liou, 2002; Petty, 2006; Zdunkowski et al., 2007), which relates transmission of radiation to scattering, emission and absorption in media such as  
20 the atmosphere. The lidar equation assumes laser radiation passes through the atmosphere where it is absorbed and scattered. A fraction of laser radiation is scattered back to the instrument and reaches the receiver. Scattering and absorption in the atmosphere is determined by its constituents – gases, liquid droplets, ice crystals and aerosol particles. The focus of the current version of the simulator is on clouds. For this purpose, the atmospheric model output needed is 4-dimensional fields of mass mixing ratios of liquid and ice and CF. The lidar equation can be applied on these output fields to simulate the backscattered  
25 radiation received by the instrument. Table 3 lists physical quantities used in the following sections. Here, we use radiative transfer notation similar to Petty (2006) and the notation of the original lidar simulator (Chiriaco et al., 2006).

Below we provide a brief review of LR, Rayleigh and Mie scattering, calculate LR of cloud droplets at lidar wavelengths of the presented instruments and introduce an empirical parametrisation of LR and multiple scattering coefficient of ice crystals based on previous studies.

#### 30 4.1 Lidar ratio

Lidar ratio  $S$  is the extinction-to-backscattering ratio of atmospheric constituents at the lidar wavelength. It is an important quantity in lidar observations and the lidar simulator because it determines the amount of attenuation and backscattering.

**Table 3.** Table of physical quantities.

Symbol	Name	Units	Expression
$\Omega$	Solid angle	sr	
$z$	Height relative to the instrument	m	
$k_B$	Boltzmann constant	JK <sup>-1</sup>	$k_B \approx 1.38 \times 10^{-23} \text{ JK}^{-1}$
$p$	Atmospheric pressure	Pa	
$T$	Atmospheric temperature	K	
$\rho_{\text{air}}$	Air density	kg.m <sup>-3</sup>	
$\rho$	Liquid (or ice) density	kg.m <sup>-3</sup>	
$q$	Cloud liquid (or ice) mass mixing ratio	1	
$N$	Particle number concentration	m <sup>-3</sup>	
$\alpha_s$ ( $\alpha_e$ )	Volume scattering (extinction) coefficient	m <sup>-1</sup>	
$P_\pi(\theta)$	Scattering phase function at angle $\theta$	1	$\int_{4\pi} P_\pi(\theta) d\Omega = 4\pi$
$\beta$	Volume backscattering coefficient	m <sup>-1</sup> sr <sup>-1</sup>	$\beta = \alpha_s P_\pi(\pi) / (4\pi)$
$\beta_{\text{mol}}$	Volume backscattering coefficient for air molecules		
$\beta_p$	Volume backscattering coefficient for cloud particles		
$\eta$	Multiple scattering coefficient	1	
$\beta'$	Attenuated volume backscattering coefficient	m <sup>-1</sup> sr <sup>-1</sup>	$\beta' = \beta \exp(-2 \int_0^z \eta \alpha_e dz)$
$S$	Lidar ratio (extinction-to-backscatter ratio)	sr	$S = \alpha_e / \beta$
$S'$	Effective (apparent) lidar ratio	sr	$S' = S\eta$
$k$	Backscatter-to-extinction ratio	sr <sup>-1</sup>	$k = 1/S$
$n(r)$	Number distribution of particle size	m <sup>-4</sup>	$N = \int_0^\infty n(r) dr$
$Q_s$ ( $Q_e$ )	Scattering (extinction) efficiency of spherical particles	1	$\alpha_s = Q_s \pi r^2 N$ , $\alpha_e = Q_e \pi r^2 N$
$Q_b$	Backscattering efficiency of spherical particles	sr <sup>-1</sup>	$\beta = Q_b \pi r^2 N$
$r_{\text{eff}}$	Effective radius	m	$r_{\text{eff}} = \int_0^\infty r^3 n(r) dr / \int_0^\infty r^2 n(r) dr$
$\sigma_{\text{eff}}$	Effective standard deviation	m	$\sigma_{\text{eff}} = (\int_0^\infty (r - r_{\text{eff}})^2 r^2 n(r) dr) / (\int_0^\infty r^2 n(r) dr)$

LR is not explicitly known from the observed attenuated volume backscattering coefficient. For liquid cloud droplets at near IR wavelengths it is relatively constant at  $S \approx 19$  sr (Sect. 4.2), while for ice crystals (Sect. 4.3) and aerosol it is highly variable. When the lidar signal is fully attenuated, and under the assumption that cloud LR is constant and scattering from clouds is much stronger than molecular and aerosol scattering, LR can be determined from the observed attenuated volume backscattering coefficient by integrating it vertically (O'Connor et al., 2004):

$$S' = \eta S = \frac{1}{2 \int_0^\infty \beta' dz}, \quad (1)$$

where  $S'$  is effective (apparent) LR, a quantity which does not depend on the multiple scattering coefficient.

## 4.2 Rayleigh and Mie scattering

The Rayleigh volume backscattering coefficient  $\beta_{\text{mol}}$  ( $\text{m}^{-1}\text{sr}^{-1}$ ) in ACTSIM is parametrised by the following equation (Eq. (8) in Chiriaco et al. (2006)):

$$\beta_{\text{mol}} = \frac{p}{k_{\text{B}}T} (5.45 \times 10^{-32}) \left( \frac{\lambda}{550\text{nm}} \right)^{-4.09} = \frac{p}{k_{\text{B}}T} C_{\text{mol}}, \quad (2)$$

- 5 where for lidar wavelength  $\lambda = 532$  nm,  $C_{\text{mol}} = 6.2446 \times 10^{-32}$ ;  $k_{\text{B}}$  is the Boltzmann constant  $k_{\text{B}} \approx 1.38 \times 10^{-23}$   $\text{JK}^{-1}$ ,  $p$  is the atmospheric pressure and  $T$  is the atmospheric temperature. We multiply this equation by  $\exp(4.09(\log(532) - \log(\lambda)))$  (where the value of  $\lambda$  is in nm) to get molecular backscattering for wavelengths other than 532 nm, which allows us to support multiple commercially available instruments. The strength of molecular backscattering is usually lower than backscattering from clouds for the relevant wavelengths.
- 10 The lidar signal at visible or near IR wavelengths is scattered by cloud droplets in the Mie scattering regime (Mie, 1908). In the most simple approximation, one can assume spherical dielectric particles. The scattering from these particles depends on the relative size of the wavelength and the (spherical) particle radius  $r$ , expressed by the dimensionless size parameter  $x$ :

$$x = \frac{2\pi r}{\lambda}. \quad (3)$$

- While the wavelength is approximately constant during the operation of the lidar<sup>2</sup>, the particle size comes from a distribution of sizes, typically approximated in NWP models and GCMs by a Gamma or log-normal distribution with a given mean and standard deviation. Some models provide the mean as effective radius  $r_{\text{eff}}$ . If the effective radius is not provided by the model, the lidar simulator assumes a value  $r_{\text{eff}} = 10$   $\mu\text{m}$  by default, which is approximately consistent with global studies of effective radius (Bréon and Colzy, 2000; Bréon and Doutriaux-Boucher, 2005; Hu et al., 2007; Zhang and Platnick, 2011; Rausch et al., 2017; Fu et al., 2019). This is different from the default effective radius of 30  $\mu\text{m}$  in the original COSP lidar simulator.

- 20 In order to support multiple laser wavelengths, it is necessary to calculate backscattering efficiency due to scattering by a distribution of particle sizes. We use the computer code MIEV developed by Warren J Wiscombe (Wiscombe, 1979, 1980) to calculate backscattering efficiency for a range of the size parameter  $x$  and integrate for a distribution of particle sizes. The resulting pre-calculated LR (extinction-to-backscatter ratio) as a function of the effective radius is included in the lidar simulator for fast lookup during the simulation.

- 25 Cloud droplet size distribution parameters are an important assumption in lidar simulation due to the dependence of Mie scattering on the ratio of wavelength and particle size (the size parameter  $x$ ). NWP models and GCMs traditionally use the effective radius  $r_{\text{eff}}$  and effective standard deviation  $\sigma_{\text{eff}}$  (or an equivalent parameter such as effective variance  $\nu_{\text{eff}}$ ) to parametrise this distribution. Knowledge of the real distribution is likely highly uncertain due to a large variety of clouds occurring globally

---

<sup>2</sup>The actual lidar wavelength is not constant and is characterised by a central wavelength and width. The central wavelength may fluctuate with temperature (Wiegner and Gasteiger, 2015).

**Table 4.** Table of sensitivity tests of theoretical distribution assumption, effective radius  $r_{\text{eff}}$  and effective standard deviation  $\sigma_{\text{eff}}$  of the cloud droplet size distribution.  $\mu$  and  $\sigma$  are the mean and standard deviation of a normal distribution corresponding to the log-normal distribution, calculated numerically from  $r_{\text{eff}}$  and  $\sigma_{\text{eff}}$ .  $\mu_*$  and  $\sigma_*$  are the actual mean and standard deviation of the distribution (calculated numerically).

Distribution	$r_{\text{eff}}$ ( $\mu\text{m}$ )	$\sigma_{\text{eff}}$ ( $\mu\text{m}$ )	$\mu$	$\sigma$	$\mu_*$ ( $\mu\text{m}$ )	$\sigma_*$ ( $\mu\text{m}$ )
log-normal	20	10	2.44	0.47	12.76	6.26
log-normal	20	5	2.84	0.25	17.72	4.43
log-normal	10	5	1.74	0.47	6.40	3.20
Gamma	20	10			9.98	7.00
Gamma	20	5			17.50	4.68
Gamma	10	5			5.00	3.54

and the limited ability to predict microphysical cloud properties in models. In this section we introduce theoretical assumptions used in the lidar simulator based on established definitions of the effective radius and effective standard deviation and two common distributions. Edwards and Slingo (1996) discuss the effective radius in the context of model radiation schemes, and we will primarily follow the definitions detailed in Chang and Li (2001) and Petty and Huang (2011). The practical result of this section (and the corresponding offline code) is pre-calculated backscatter-to-extinction ratios as a function of the effective radius in the form of a lookup table included in the lidar simulator, and used in the online calculations. The offline code is provided and can be re-used for calculation of the necessary lookup tables for different lidar wavelengths, should the user of the code want to support another instrument.

The effective radius  $r_{\text{eff}}$  and effective standard deviation  $\sigma_{\text{eff}}$  are defined by:

$$r_{\text{eff}} = \frac{\int_0^\infty r^3 n(r) dr}{\int_0^\infty r^2 n(r) dr}, \quad \sigma_{\text{eff}}^2 = \frac{\int_0^\infty (r - r_{\text{eff}})^2 r^2 n(r) dr}{\int_0^\infty r^2 n(r) dr}, \quad (4)$$

where  $n(r)$  is the probability density function (PDF) of the distribution. Here, we follow Petty and Huang (2011), who define the effective variance  $\nu_{\text{eff}}$  which relates to  $\sigma_{\text{eff}}$  by  $\nu_{\text{eff}} = \sigma_{\text{eff}}^2 / r_{\text{eff}}^2$ . Due to lack of knowledge about the real distribution of particle radii, it has to be modelled by a theoretical distribution, such as a log-normal or Gamma distribution. The original ACTSIM simulator assumes a log-normal distribution (Chiriaco et al., 2006) with the PDF:

$$n(r) \propto \frac{1}{r} \exp\left(-\frac{(\log r - \mu)^2}{2\sigma^2}\right), \quad (5)$$

where  $\mu$  and  $\sigma$  are the mean and the standard deviation of the corresponding normal distribution, respectively. Chiriaco et al. (2006) use the value of  $\sigma = \log(1.2) = 0.18$  "for ice clouds" (the value for liquid cloud does not appear to be documented). In our parametrisation we used a combination of  $r_{\text{eff}}$  and  $\sigma_{\text{eff}}$  to constrain the theoretical distribution, where the effective standard deviation  $\sigma_{\text{eff}}$  was assumed to be one fourth of the effective radius  $r_{\text{eff}}$ . This choice is approximately consistent with

$\sigma = \log(1.2) = 0.18$  at  $r_{\text{eff}} = 20 \mu\text{m}$  (see Table 4, described below). In future updates, the values could be based on in situ studies of size distribution or taken from the atmospheric model output if available.

From the expression for the  $n$ -th moment of the log-normal distribution  $E[X^n] = \exp(n\mu + n^2\frac{\sigma^2}{2})$  and Equation (4) we calculate  $r_{\text{eff}}$  and  $\sigma_{\text{eff}}$  of the log-normal distribution:

$$5 \quad r_{\text{eff}} = \frac{E[r^3]}{E[r^2]} = \exp\left(\mu + \frac{5}{2}\sigma^2\right), \quad (6)$$

$$\begin{aligned} \sigma_{\text{eff}}^2 &= \frac{E[(r - r_{\text{eff}})^2 r^2]}{E[r^2]} = \frac{E[r^4] - 2E[r^3]r_{\text{eff}} + r_{\text{eff}}^2 E[r^2]}{E[r^2]} = \frac{\exp(4\mu + 8\sigma^2) - \exp(4\mu + 7\sigma^2)}{\exp(2\mu + 2\sigma^2)} = \\ &= \exp(2\mu + 6\sigma^2) - \exp(2\mu + 5\sigma^2). \end{aligned} \quad (7)$$

We find  $\mu$  and  $\sigma$  for given  $r_{\text{eff}}$  and  $\sigma_{\text{eff}}$  numerically by root-finding using the equations above. In practice, we find that the root-finding converges well for  $r_{\text{eff}}$  between 5 and 50  $\mu\text{m}$ , which is the range most likely to be applicable in practice.

10 The Gamma distribution follows the PDF:

$$n(r) \propto r^{(1-3\nu_{\text{eff}})/\nu_{\text{eff}}} \exp\left(-\frac{r}{r_{\text{eff}}\nu_{\text{eff}}}\right) \quad (8)$$

(see e.g. Eq. 13 in Petty and Huang (2011) or Eq. 1 in Bréon and Doutriaux-Boucher (2005)). In this case, the distribution depends explicitly on  $r_{\text{eff}}$  and  $\sigma_{\text{eff}}$ , and as such does not require numerical root-finding.

Figure 2a shows the log-normal and Gamma distributions calculated for a number of  $r_{\text{eff}}$  and  $\sigma_{\text{eff}}$  values, and Table 4  
15 summarises properties of these distributions. The actual mean and standard deviation of the distributions do not necessarily correspond well with the effective radius and effective standard deviation.

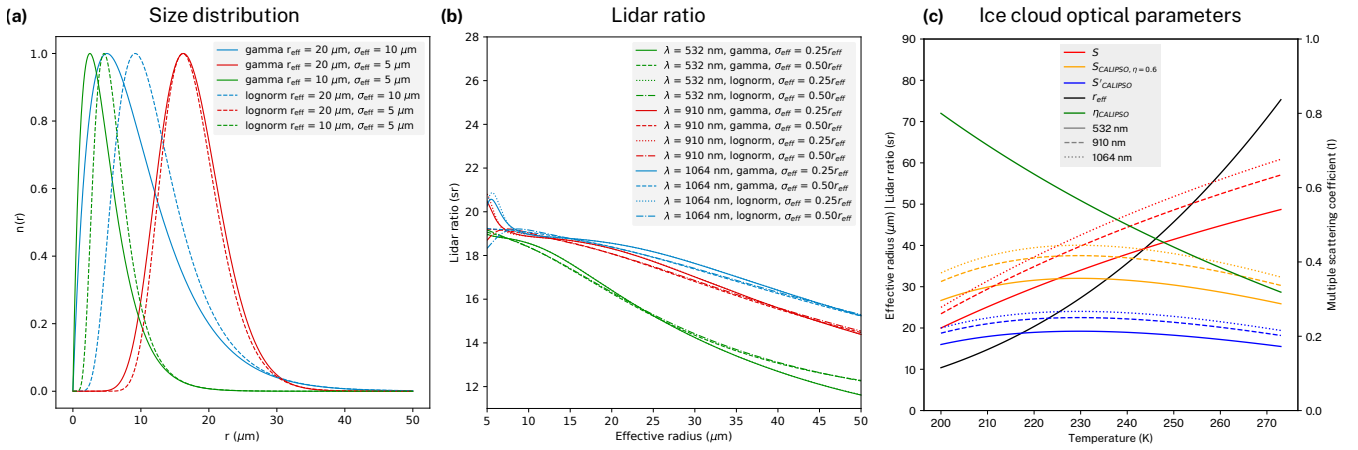
In ACTSIM, the volume extinction coefficient  $\alpha_e$  is calculated by integrating the extinction by individual particles over the particle size distribution:

$$\alpha_e = \int_0^{\infty} Q_e \pi r^2 n(r) dr \approx Q_e \pi \int_0^{\infty} r^2 n(r) dr = Q_e \frac{3q\rho_{\text{air}}}{4\rho r_{\text{eff}}}, \quad (9)$$

20 assuming approximately constant extinction efficiency  $Q_e \approx 2$  (which is approximately true for the interesting range of  $r_{\text{eff}}$  and laser wavelengths), and using the relationship between the cloud liquid mass mixing ratio  $q$  and  $\int_0^{\infty} r^2 n(r) dr$ :

$$q\rho_{\text{air}} = \int_0^{\infty} \frac{4}{3} \pi r^3 \rho n(r) dr = \frac{4}{3} \pi \rho \int_0^{\infty} r^3 n(r) dr = \frac{4}{3} \pi \rho r_{\text{eff}} \int_0^{\infty} r^2 n(r) dr, \quad (10)$$

where  $\rho$  and  $\rho_{\text{air}}$  are the densities of liquid water and air, respectively.



**Figure 2.** (a) Theoretical distributions of cloud droplet radius based on the log-normal and Gamma distributions parametrised by multiple choices of the effective radius  $r_{\text{eff}}$  and effective standard deviation  $\sigma_{\text{eff}}$ . (b) Lidar ratio (LR) as a function of effective radius calculated for different theoretical cloud droplet size distributions, laser wavelengths and effective standard deviation ratios. (c) Parametrisation of ice cloud optical properties as a function of temperature based on Garnier et al. (2015) and Heymsfield (2005). The plot shows LR ( $S$ ), LR of CALIPSO calculated using the constant standard processing multiple scattering coefficient  $\eta = 0.6$  ( $S_{\text{CALIPSO}, \eta=0.6}$ ), the effective LR of CALIPSO ( $S'_{\text{CALIPSO}}$ ), the effective radius ( $r_{\text{eff}}$ ), the multiple scattering coefficient of CALIPSO ( $\eta_{\text{CALIPSO}}$ ) determined by Garnier et al. (2015). LRs are calculated for three wavelengths of 532 nm (solid line), 910 nm (dashed line) and 1064 nm (dotted line) by scaling with colour ratio.

Likewise, the volume backscattering coefficient from particles  $\beta_p$  is calculated by integrating backscattering by individual particles over the particle size distribution:

$$\beta_p = \int_0^{\infty} Q_s \pi r^2 \frac{P_{\pi}(\pi)}{4\pi} n(r) dr. \quad (11)$$

where  $Q_s$  is scattering efficiency and  $P_{\pi}(\pi)$  is scattering phase function at  $180^\circ$ . Since the normalisation of  $n(r)$  is not known until the online phase of calculation, the backscatter-to-extinction ratio from particles  $k_p = \beta/\alpha_e$  can be calculated offline instead (the requirement for normalisation of  $n(r)$  is avoided by appearing in both the numerator and denominator):

$$k_p = \beta_p/\alpha_e = \frac{\int_0^{\infty} Q_s r^2 P_{\pi}(\pi)/(4\pi) n(r) dr}{\int_0^{\infty} Q_e r^2 n(r) dr}. \quad (12)$$

We pre-calculate this integral numerically for a permissible interval of  $r_{\text{eff}}$  (5–50  $\mu\text{m}$ ) at 500 evenly spaced wavelengths, and store the result as a lookup table for the online phase. The integral in the numerator is numerically hard to calculate due to strong dependency of  $P_{\pi}(\pi)$  on  $r$ . Figure 2b shows LR as a function of  $r_{\text{eff}}$ , calculated for log-normal and Gamma particle size distributions with  $\sigma_{\text{eff}} = 0.25r_{\text{eff}}$  and  $\sigma_{\text{eff}} = 0.5r_{\text{eff}}$ . This corresponds to the lookup table we use in the online phase of the lidar



simulator. As can be seen in Fig. 2, LR depends only weakly on the choice of the distribution type and the effective standard deviation ratio.

### 4.3 Backscattering from ice crystals

Simulation of backscattering from ice crystals is relatively complex compared to backscattering from liquid droplets due to the very high variability of ice crystal microphysical properties such as habit, size, orientation and surface roughness, all of which affect LR, extinction cross section, single-scattering albedo and multiple scattering coefficient. Common habits include hexagonal plates, hexagonal columns, hollow hexagonal columns, droxtals, bullet rosettes, hollow bullet rosettes and aggregates (Baran, 2009; van Diedenhoven, 2017). Size can be highly variable and bimodal with a dependence on temperature and relative humidity. Orientation is commonly random or horizontally oriented (often reported with hexagonal ice plates). Surface can vary between smooth and rough depending on supersaturation and crystal age. In general, the Mie theory cannot be used to simulate backscattering from ice crystals because of their irregular shape (Yang et al., 2014). While large crystals allow the use of the geometric optics approximation to estimate the optical properties, smaller crystals and diffraction by large crystals necessitate the use of more advanced techniques such as the T-matrix method, finite-difference time domain (FDTD), discrete dipole approximation (DDA) and others, which are generally computationally expensive. Current global atmospheric models do not normally parametrise the microphysical properties of cloud ice explicitly, and provide only very limited information such as ice mass concentration and in some cases the effective radius of ice crystals in the model output. Radiative transfer schemes of atmospheric models do not explicitly evaluate backscattering (the phase function at  $180^\circ$ ) and therefore cannot provide this information to the simulator. Instead the phase function is parametrised by the asymmetry factor, which is likely insufficient to give an accurate estimate of backscattering.

Because the model ice crystal microphysical and optical properties are not known, they have to be parametrised. A first option is to parametrise the microphysical properties such as habit and size and calculate optical properties theoretically. A second option is to parametrise the optical properties directly. This appears to be a more practical choice because of the broad availability of global remote sensing measurements of optical properties from satellites and ground-based lidars, compared to relatively scarce in situ measurements of ice crystals. Garnier et al. (2015) analysed CALIPSO lidar and co-located passive infrared data from the Imaging Infrared Radiometer (IIR) and determined a global relationship between temperature and LR and multiple scattering coefficient at the lidar wavelength of 532 nm. The multiple scattering coefficient is taken as a constant of 0.6 in the standard CALIPSO data processing, but they identified that it is in fact variable between about 0.4 and 0.8. Here, we parametrise LR based on their findings. LR varies with the lidar wavelength, a larger part of which is due to the change in the diffraction peak and a smaller part is due to the variation of the refractive index (Borovoi et al., 2014). We use colour ratio to estimate LR at lidar wavelengths other than 532 nm. Colour ratio of 1064 nm relative to 532 nm is commonly estimated for the dual-wavelength lidars such as CALIOP. Here, we use a value of 0.8 approximately consistent with the results of Bi et al. (2009) and Vaughan et al. (2010). The effective radius is defined for non-spherical particles as  $r_{\text{eff}} = \frac{3}{2} \frac{\text{IWC}}{\sigma}$ , where IWC is the ice water content, and  $\sigma$  is the volume extinction coefficient of ice. Heymsfield (2005) summarised ice crystal effective radius (related to "IWC/ $\sigma$ " by a factor of 1.64) parametrised as a function of temperature based on a number of field studies. We use

this relationship for determination of the effective radius. Figure 2c shows the true and effective LR based on Garnier et al. (2015) and the effective radius based on Heymsfield (2005), parametrised by the following equations:

$$S = \left( 20 + (34 - 20) \frac{1/T - 1/200}{1/230 - 1/200} \right) \text{sr},$$

$$\eta = 0.8 + (0.5 - 0.8) \frac{1/T - 1/200}{1/240 - 1/200},$$

$$r_{\text{eff}} = \exp \left( \log(16.4) + (\log(49.2) - \log(16.4)) \frac{1/T - 1/213.15}{1/253.15 - 1/213.15} \right) \mu\text{m},$$

where  $T$  is atmospheric temperature in K.  $S$  follows (Garnier et al., 2015, Figure 12b),  $\eta$  follows (Garnier et al., 2015, Figure 9a) and  $r_{\text{eff}}$  follows (Heymsfield, 2005, Figure 2), where the concave and convex shape (respectively) as approximated by using  $1/T$  as an argument of the linear approximation, and we use a logarithmic scale of  $r_{\text{eff}}$  in the expression for  $r_{\text{eff}}$  to avoid negative values at low temperature. Figure 2c also shows LR when calculated with the assumption of  $\eta = 0.6$  ( $S_{\text{CALIPSO}, \eta=0.6}$ ) as in the standard processing of CALIPSO data. This corresponds to the empirically found relationship in (Garnier et al., 2015, Figure 12a) and (Josset et al., 2012, Figure 9) with a local maximum at 225 K. LR at wavelengths other than 532 nm is approximated by  $0.8 \frac{\lambda - 532}{532}$ , where  $\lambda$  is lidar wavelength in  $\mu\text{m}$  and 0.8 is the approximate value of 1064 nm/532 nm colour ratio. The parametrisation of LR ( $S$  in Fig. 2c) spans about the same range of values as reported by (Hopkin, 2018, Figure 5.6) (20 to 60 sr) and Yorks et al. (2011) (10 to 60 sr). Based on CALIPSO observations, Hu (2007) identified that while effective LR of global ice clouds at a lidar wavelength of 532 nm is mostly clustered around 17 sr, horizontally oriented plates produce a much lower effective LR below 10 sr caused by specular reflection. These results are close to our parametrisation of effective LR ( $S'_{\text{CALIPSO}}$ ). In the current version of the lidar simulator we do not parametrise horizontally oriented plates, but in a future version they could be taken into account by parametrising their concentration based on temperature (Noel and Chepfer, 2010). For the ALCs we use the same constant value of the multiple scattering coefficient  $\eta = 0.7$  as for liquid cloud droplets (Sect. 4.5).

#### 4.4 Cloud overlap and cloud fraction

Model cloud is defined by the liquid and ice mass mixing ratio and cloud fraction in each atmospheric layer. The lidar simulator simulates radiation passing vertically at a random location within the grid cell. Therefore, it is necessary to generate a random vertical cloud overlap based on the cloud fraction in each layer, as the overlap is not defined explicitly in the model output. Two common methods of generating overlap are the random and maximum–random overlap (Geleyn and Hollingsworth, 1979). In the random overlap method, each layer is either cloudy or clear with a probability given by CF, independent of other layers. The maximum–random overlap assumes that adjacent layers with non-zero CF are maximally overlapped, whereas layers separated by zero CF layers are randomly overlapped. COSP implements cloud overlap generation in the Subgrid Cloud Overlap Profile Sampler (SCOPS) (Klein and Jakob, 1999; Webb et al., 2001; Chepfer et al., 2008). The ALC lidar simulator uses SCOPS to generate 10 random subcolumns for each profile, using the maximum–random overlap assumption as the default setting of a user-configurable option. The attenuated volume backscattering coefficient profile and cloud occurrence can be plotted for any

subcolumn. Due to the random nature of the overlap, the attenuated volume backscattering coefficient profile may differ from the observed profile even if the model is correct in its cloud simulation. The random overlap generation should, however, result in unbiased cloud statistics.

#### 4.5 Multiple scattering

5 Due to a finite FOV of the lidar receiver, a fraction of the laser radiation scattered forward will remain in FOV. Therefore, the effective attenuation is smaller than calculated with the assumption that all but the backscattered radiation is removed from FOV and cannot reach the receiver. The forward scattering can be repeated multiple times before a fraction of the radiation is backscattered, eventually reaching the receiver. To account for this multiple scattering effect, the COSP lidar simulator uses a multiple scattering correction coefficient  $\eta$ , by which the volume scattering coefficient is multiplied before calculating the layer optical thickness (Chiriaco et al., 2006; Chepfer et al., 2007, 2008). The theoretical value of  $\eta$  is between 0 and 1 and depends on the receiver FOV and optical properties of the cloud. For CALIOP at  $\lambda = 532$  nm a value of 0.7 is used in the COSP lidar simulator. Hogan (2006) implemented fast approximate multiple scattering code. This code has recently been used by Hopkin et al. (2019) in their ceilometer calibration method. They noted that  $\eta$  is usually between 0.7 and 0.85 for wavelengths between 905 and 1064 nm. The ALC simulator presented here does not use an explicit calculation of  $\eta$ , but retains the value of  $\eta = 0.7$  for cloud droplets. The code of Hogan (2006) "Multiscatter" is publicly available (<http://www.met.reading.ac.uk/clouds/multiscatter/>) and could be used in a later version of the framework to improve the accuracy of simulated attenuation and calibration.

### 5 Lidar data processing

Scheme in Fig. 1 outlines the processing done in the framework. The individual processing steps are described below.

#### 20 5.1 Noise and subsampling

ALC signal reception is affected by a number of sources of noise such as sunlight and electronic noise (Kotthaus et al., 2016). Range-independent noise can be removed by assuming that the attenuated volume backscattering coefficient at the highest range gate is dominated by noise. This is true if the highest range is not affected by clouds, aerosol, and if contributions from molecular scattering are negligible. The supported instruments have a range of approximately 8 (CL31), 15 (CL51, CHM 15k) and 30 km (MiniMPL). By assuming the distribution of noise at the highest level is approximately normal, the mean and standard deviation can be calculated from a sample over a period of time such as 5 minutes, which is short enough to assume the noise is constant over this period, and long enough to achieve accurate estimates of the standard deviation. The mean and standard deviation can then be scaled by the square of the range to estimate the distribution of range independent noise at each range bin. By subtracting the noise mean from the measured attenuated volume backscattering coefficient we get the expected attenuated volume backscattering coefficient. The result of the noise removal algorithm is the expected attenuated volume backscattering coefficient and its standard deviation at each range bin.

**Table 5.** Theoretical molecular volume backscattering coefficient calculated at pressure 1000 hPa and temperature 20 °C and the calibration coefficient, relative to the instrument native units, determined for the instrument based on the molecular volume backscattering coefficient and stratocumulus lidar ratio calibration methods.

Instrument	Wavelength (nm)	Molecular volume backscattering coefficient ( $\times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ )	Calibration coefficient
CHM 15k	1064	0.0906	0.34
CL31	910	0.172	$1.45 \times 10^{-3}$
CL51	910	0.172	$1.2 \times 10^{-3}$
MiniMPL	532	1.54	$3.75 \times 10^{-6}$

## 5.2 Backscatter calibration

ALCs often report attenuated volume backscattering coefficient in arbitrary units (a.u.) or as NRB (MiniMPL). If they report it in units of  $\text{m}^{-1} \text{sr}^{-1}$ , these values are often not calibrated to represent the true absolute attenuated volume backscattering coefficient. Assuming that range-dependent corrections (overlap, dead time and afterpulse) have been applied on attenuated volume backscattering coefficient in a. u., the reported attenuated volume backscattering coefficient is proportional to the true attenuated volume backscattering coefficient (inclusive of noise backscattering). In order to have a comparable quantity to the lidar simulator and consistent input to the subsequent processing (e.g. cloud detection), calibration by multiplying by a calibration coefficient is required. Formally, the units of the calibration coefficient depend on the units of backscattering recorded by the instrument, which are  $\text{m}^{-1} \text{sr}^{-1}$  in CL31, CL51, unitless in CHM 15k and  $\text{count} \cdot \mu\text{s}^{-1} \mu\text{J}^{-1} \text{km}^2$  in MiniMPL, i.e. the units of the calibration coefficient are  $\text{m}^{-1} \text{sr}^{-1} / (\text{instrument units})$ . In the following discussion, we leave out the units. Several methods of calibration have been described previously: calibration based on LR in fully attenuating liquid stratocumulus clouds (O’Connor et al., 2004; Hopkin et al., 2019), calibration based on molecular backscattering (Wiegner et al., 2014) and calibration based on a high spectral resolution lidar reference (Heese et al., 2010; Jin et al., 2015). In addition, calibration can be assisted by sunphotometer or radiosonde measurements (Wiegner et al., 2014).

A relatively large variability of the calibration coefficient has been determined for instruments of the same model (Hopkin et al., 2019). However, past studies can be useful for determining an approximate value of the coefficient before applying one of the calibration methods. For the CL51, Jin et al. (2015) reported a value of  $1.2 \pm 0.1$  based on a multi-wavelength lidar reference. Hopkin et al. (2019) reported mean values 1.4–1.5 for a number of CL31 instruments (software version 202). For CHM 15k, Hopkin et al. (2019) reported mean values between 0.3 and 0.8 for a majority of the instruments examined. The ALCF provides per-instrument default values of the calibration coefficient (Table 5), but a unit-specific coefficient should be determined for an analysed instrument during the lidar data processing step.

Calibration based on LR in fully opaque liquid stratocumulus clouds has been applied successfully on large networks of ALCs. It utilises the fact that given suitable conditions vertically integrated attenuated volume backscattering coefficient is proportional to LR of the cloud, which can be theoretically derived if the cloud droplet effective radius can be assumed. The theoretically derived value is about 18.8 sr for common ALC wavelengths and a relatively large range of effective radii

(O'Connor et al., 2004). Another factor which needs to be known or assumed is the multiple scattering coefficient, which tends to be about 0.7-1.0 in common ALCs. Due to its relatively simple requirements, this method is possibly the easiest ALC calibration method. The ALCF implements this calibration method by letting the user identify time periods with fully opaque liquid stratocumulus cloud, for which the mean LR is calculated. The ratio of the observed LR and the theoretical LR is equivalent to the calibration coefficient. This implementation, while very easy to perform, has multiple limitations, some of which are highlighted by Hopkin et al. (2019):

1. Aerosol can cause additional attenuation and scattering, which results in LR which is different from the theoretical value by an unknown factor. Therefore, a frequent re-calibration may be necessary.
2. The multiple scattering coefficient assumption may not be accurate for the given instrument.
3. The 910 nm wavelength of CL31 and CL51 is affected by water vapour absorption which causes additional attenuation, which is currently not taken into account in the calculation of LR.
4. Near-range attenuated volume backscattering coefficient retrieval is affected by receiver saturation and incomplete overlap. Therefore, using stratocumulus clouds above approximately 2 km for this calibration method is recommended. This range is instrument dependent.
5. The composition of the stratocumulus cloud may be uncertain. At temperature between 0 and -30°C these clouds may contain both liquid and ice which results in a different LR than expected.

These limitations could be addressed in the future by (1) using sunphotometer observations as an optional input to determine the aerosol optical depth (AOD), (2) calculating the multiple scattering coefficient more accurately (such as with the Multi-scatter package of Hogan (2006)), (3) calculating the water vapour absorption explicitly based on water vapour, temperature and pressure fields from a reanalysis or radiosonde profile data, (4) correcting the near-range backscatter based on the integrated attenuated volume backscattering coefficient distribution as a function of height of the maximum backscatter (Hopkin et al., 2019, Sect. 5.1), (5) combining the attenuated volume backscattering coefficient profile with temperature field from a reanalysis to exclude cold clouds.

Molecular (Rayleigh) backscattering can be accurately calculated if temperature and pressure of the atmospheric profile is known (Sect. 4.2). This can be employed for absolute calibration of ALCs. Given the low SNR of low-power ALCs, several hours of integration are required to identify the molecular backscattering (Wiegner et al., 2014). The molecular backscattering is attenuated by an unknown amount of aerosol with unknown LR, and the near-range backscattering is affected by a potentially inaccurate overlap correction. Therefore, this method alone produces calibration coefficient which depend on the atmospheric conditions. We found that all studied ALCs except for the CL31 are capable of observing the molecular backscattering (Section 7). Therefore, this method may be used in addition to the liquid stratocumulus LR method for cross-validation of the calibration.

### 5.3 Cloud detection

Cloud is the most strongly attenuating feature in ALC attenuated volume backscattering coefficient measurements. Due to this attenuation, the lidar signal is quickly attenuated in thick cloud and can fall below the noise level before reaching the top of the cloud. This means that the first cloud base can be detected reliably (unless the cloud is too thin or too high and obscured by noise), while the cloud top or multi-layer cloud cannot be observed reliably under all conditions. The opposite is true for spaceborne lidars, which can detect the cloud top reliably but cannot always detect the cloud base. Therefore, ALC observations can be regarded as complementary to spaceborne lidar observations. By applying a suitable algorithm, one can detect CBH, CTH and identify cloud layers. Instrument firmware often determines CBH and sometimes cloud layers as part of its internal processing, often using an undisclosed algorithm which is not comparable between different instruments and potentially not even different versions of the instrument firmware (Kotthaus et al., 2016). Mattis et al. (2016) compared a large number of ALCs and found differences of up to 70 m between the reported CBH, and others found relatively large differences as well (Liu et al., 2015b; Silber et al., 2018). Alternatively to the instrument reported CBH and cloud layers, it is possible to detect cloud based on the attenuated volume backscattering coefficient profile. A relatively large number of cloud detection algorithms have been proposed (Wang and Sassen, 2001; Morille et al., 2007; Martucci et al., 2010; Van Tricht et al., 2014; Silber et al., 2018; Cromwell and Flynn, 2019). We use a simple algorithm based on an attenuated volume backscattering coefficient threshold applied on the denoised backscatter, assuming that the noise can be represented by a normal distribution at the highest range, which is unlikely to contain cloud or aerosol if the instrument is pointing vertically (this may not be true, however, for CL31 which has a maximum range of just 7.7 km). This assumption neglects the range-dependent molecular backscattering, which is relatively small at the ceilometer wavelengths examined (910 nm and 1064 nm). A cloud mask is determined positive where attenuated volume backscattering coefficient is greater than a chosen threshold plus 5 standard deviations of noise at the given range. In addition, the observed attenuated volume backscattering coefficient can optionally be coupled with simulated attenuated molecular volume backscattering coefficient and molecular backscattering removed from the observed backscattering prior to cloud detection. This improves the results in the boundary layer, especially with instruments which operate in the visible range and therefore affected by large molecular backscattering (MiniMPL). A threshold of  $2 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  was found to be a good compromise between false detection and misses in our Southern Hemisphere data relatively unaffected by anthropogenic aerosol. Our observed and simulated results show that cloud backscatter is generally higher than  $1 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ , and a threshold below  $2 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  results in excessive false detection due to aerosol, molecular backscattering and noise from sunlight. The threshold is an adjustable option of the ALCF. Users are encouraged to change this value if, for example, they data are affected by a large amount of aerosol. This value is above the maximum molecular backscattering, which is approximately  $1.54 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  at the surface in the case of the MiniMPL (wavelength 532 nm). Noise is not simulated by the lidar simulator, but the cloud detection algorithm allows for coupling of simulated and observed profiles, whereby the noise standard deviation is taken from the corresponding location in the observed profile. With 5 minute averaging, when the standard deviation of noise is relatively low, we found that the coupling does not make substantial differences to the detected cloud (not shown). While the threshold-based algorithm is less sophisticated than other methods of

**Table 6.** Location of sites and instruments. The time periods are inclusive.

Site	Coordinates	Surface altitude (m)	Instruments	Time period	Missing period	Days
Cass	43.0346°S 171.7594°E	577	CL51	19 Sep–1 Oct 2014		13
Lauder	45.0379°S 169.6831°E	370	MiniMPL, CL31	12–24 Jan 2018		13
Christchurch	43.5225°S 172.5841°E	45	MiniMPL, CHM 15k	17 July–18 August 2019	22–31 July	23

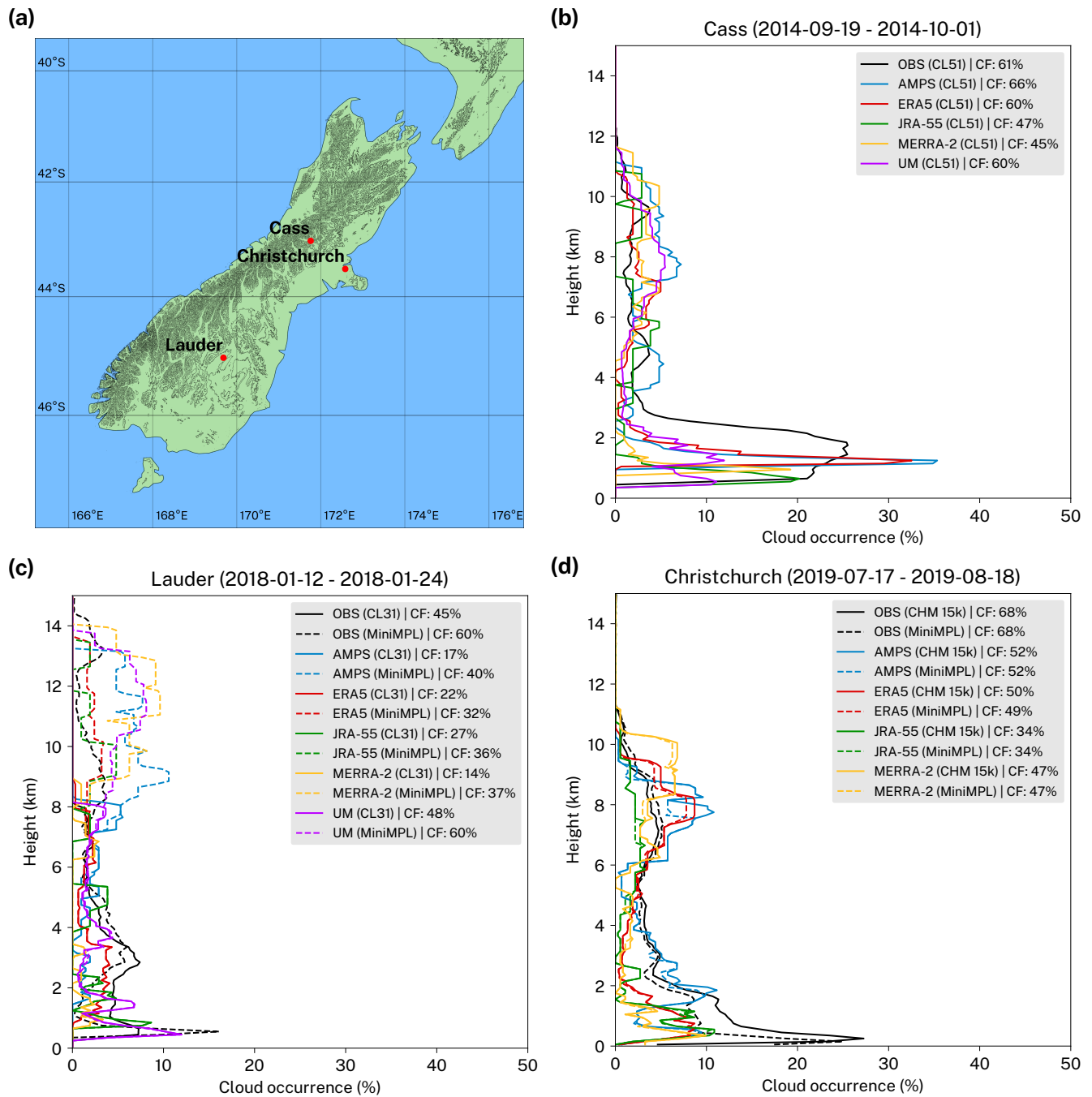
cloud detection, the vertical resolution of the simulated attenuated volume backscattering coefficient is likely too low and the vertical derivatives of the simulated attenuated volume backscattering coefficient too crudely represented (Table 7) to apply any algorithm based on the vertical derivatives of attenuated volume backscattering coefficient. Using the same cloud detection algorithm on the observed and simulated attenuated volume backscattering coefficient is essential for an unbiased one-to-one comparison of cloud.

#### 5.4 Water vapour absorption

Previous studies have noted that ceilometers which utilise the wavelength of 910 nm such as the Vaisala CL31 and CL51 are affected by additional absorption of laser radiation by water vapour (Wiegner and Gasteiger, 2015; Wiegner et al., 2019; Hopkin et al., 2019). The wavelength coincides with water vapour absorption bands between 900 and 930 nm, while the other common ceilometer wavelength of 1064 nm is not affected. Wiegner and Gasteiger (2015) reported that it can cause absorption of the order of 20% in the extratropics and 50% in the tropics. The lidar simulator does not currently account for this. However, as the water vapour concentration is available from the reanalyses and models, it should be possible to use a line-by-line model to calculate the water vapour volume absorption coefficient for each vertical layer during the integration process. Water vapour also affects calibration of the observed attenuated volume backscattering coefficient. In order to use the liquid stratocumulus LR calibration method, attenuated volume backscattering coefficient has to be corrected for water vapour absorption to achieve high accuracy of calibration. Hopkin et al. (2019) used a simplified approach based on a parametrised curve and reported a difference from explicit radiative transfer calculations of 2% in the United Kingdom atmosphere (Middle Wallop). In the future either approach should be used to include water vapour absorption in the simulator, or remove the effect of water vapour absorption from the observed lidar attenuated volume backscattering coefficient to achieve an improved one-to-one comparison between the observations, reanalyses and models.

## 6 Description of case studies

The case studies analysed here were selected to include all instruments supported by the framework. We compare four different instruments (CHM 15k, CL31, CL51, MiniMPL) deployed at three locations in NZ (Lauder, Christchurch, Cass) with three reanalyses (MERRA-2, ERA5, JRA-55), one NWP model (AMPS) and one GCM (UM). These case studies aim to demonstrate capability rather than to comprehensively evaluate cloud simulation in the models and reanalyses. The work detailed in Kuma



**Figure 3.** (a) Map showing the location of sites. Data at three sites in New Zealand were analysed: Cass, Lauder and Christchurch. (b), (c), (d) Cloud occurrence histograms as a function of height above the mean sea level observed at three sites and simulated by the lidar simulator based on atmospheric fields for five reanalyses and models. Shown is also the total cloud fraction (CF). The histogram is calculated from the cloud mask as determined by the cloud detection algorithm.



**Table 7.** Number of models levels and vertical resolution in the range of the instrument at the locations of the case studies. First number is the number of levels, followed by the minimum and maximum distance range between adjacent model levels in the lidar’s range (m).

	Cass (CL51)	Lauder (CL31)	Lauder (MiniMPL)	Christchurch (CHM 15k)	Christchurch (MiniMPL)
AMPS	42; 33–778	31; 35–528	59; 35–1021	43; 33–779	60; 33–870
ERA5	23; 222–1469	17; 220–950	30; 220–4748	25; 213–1425	31; 213–4107
JRA-55	23; 223–1479	17; 217–948	26; 217–1402	25; 213–1426	26; 213–1426
MERRA-2	34; 118–1080	26; 125–669	47; 125–1329	34; 124–1059	48; 124–1167
UM	44; 70–645	33; 32–449	65; 32–1181		

et al. (2020) provides a detailed evaluation of the UM and MERRA-2 relative to shipborne ceilometer observations. Figure 3a shows the location of the sites and Table 6 summarises the case studies, which are also described in greater detail below. The sites were chosen from available datasets to demonstrate the use of the framework with all supported instruments. Two of the sites also had co-located instruments: CL31 and MiniMPL in Lauder, and CHM 15k and MiniMPL in Christchurch. The MiniMPL in Lauder and Christchurch were two different units. The number of models levels within the range of each instrument and vertical resolution range are listed in Table 7.

Cass is a field station of the University of Canterbury located at an altitude of 577 m in the Southern Alps of the South Island of NZ. The station is located far from any settlements and likely affected little by anthropogenic aerosol relative to the other sites. We have analysed 13 days of observations with a CL51 at this station performed in September and October 2014.

Lauder is a field station of NIWA located inland in the Central Otago region on the South Island of NZ. The station is situated in a rural area relatively far from large human settlements at an altitude of 370 m. We have analysed 13 days of co-located MiniMPL and CL31 observations made in January 2018. The MiniMPL was operated in an enclosure with a scanning head set to a fixed vertical scanning mode during this period (elevation angle 90°).

Observations at the Christchurch site were performed at the University of Canterbury campus on the Ernest Rutherford building rooftop at an altitude of 45 m. Christchurch is located on the east coast of the South Island of NZ. Its climate is affected by the ocean, its proximity to the hilly area of the Banks Peninsula, the Canterbury Plains and föhn-type winds (Canterbury northwester) resulting from its position on the lee side of the Southern Alps. The city is affected by significant wintertime air pollution from domestic wood burning and transport. The orography of the city and the adjacent Canterbury Plains is very flat, making it prone to inversions. The Ernest Rutherford building is a 5 floor building situated in an urban area, surrounded by multiple buildings of similar height. We have analysed 23 days of co-located MiniMPL and CHM 15k observations performed in July and August 2019. The MiniMPL was operated in an enclosure with a scanning head set to a fixed vertical scanning mode (elevation angle 90°). The nudged run of the UM was only available up to year 2018. Therefore, it was not analysed for this site.

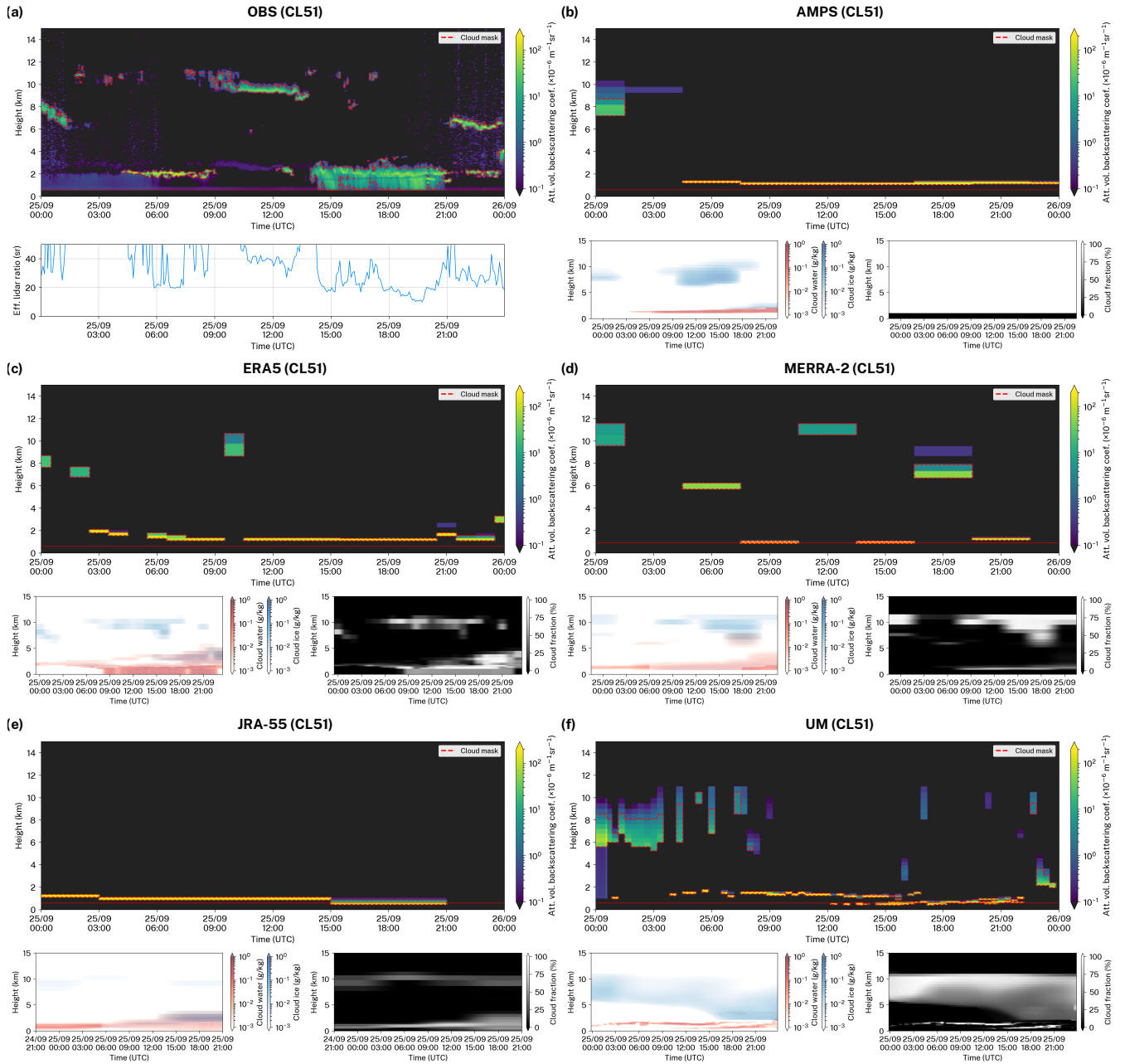
## 7 Results

To demonstrate the ways that the ALCF can be used we compared a total of 49 days of ALC observations with simulated lidar attenuated volume backscattering coefficient at three sites in NZ (Sect. 6). The observed attenuated volume backscattering coefficient was normalised to calibrated absolute range-corrected attenuated volume backscattering coefficient. The noise mean as determined at the furthest range was removed from attenuated volume backscattering coefficient. Cloud detection based on an attenuated absolute volume backscattering coefficient threshold of  $2 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ , after removing molecular backscattering and 5 noise standard deviations, was applied to derive a cloud mask and CBH. We compare the statistical cloud occurrence as a function of height above the mean sea level (ASL) (Fig. 3b, c, d) and individual attenuated volume backscattering coefficient profiles (selected profiles are shown in Fig. 4, 5 and 6) in this section. In these plots 5 standard deviations of the attenuated volume backscattering coefficient noise (Sect. 5.3) were removed. In addition, molecular backscattering was removed by coupling the observed data (Fig. 4a, 5a, 6a) with molecular attenuated volume backscattering coefficient calculated by the lidar simulator based the MERRA-2 reanalysis data. The same applies to model data (Fig. 4b–f, 5b–f, 6b–e), but molecular attenuated volume backscattering coefficient was calculated by the lidar simulator based the respective model data.

### 7.1 Cass

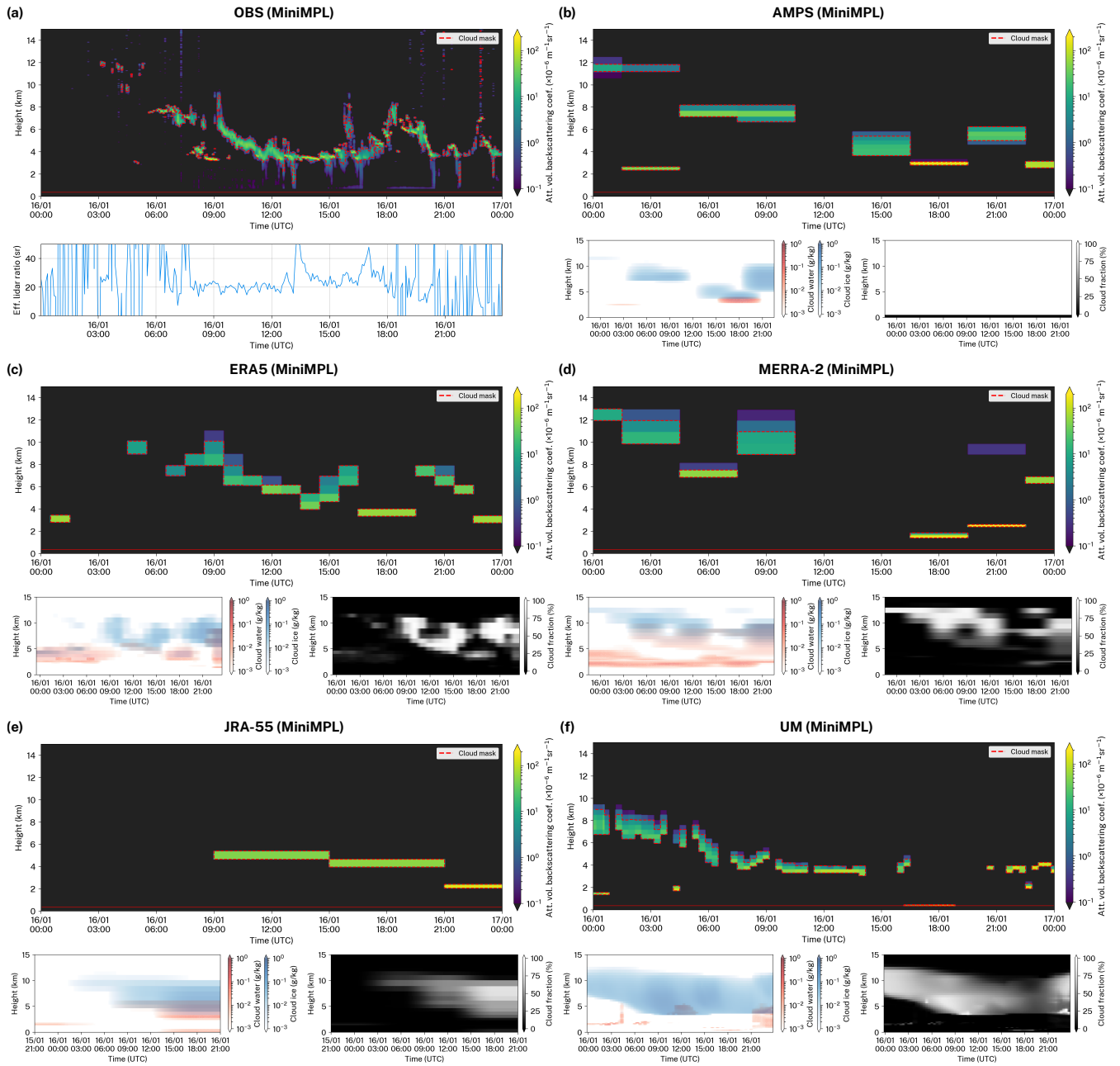
We analysed 13 days of CL51 observations at the Cass field station in late winter. Due to the location of the station at a relatively high altitude in a varied terrain of the Southern Alps, the models with their relatively coarse horizontal grid resolution do not represent the terrain and position accurately. The orography representation of the models meant that the virtual altitude of the station was 1115 m (AMPS), 1051 m (ERA5), 401 m (JRA-55), 914 m (MERRA-2) and 428 m (UM). The virtual position, which is the centre of the nearest model grid cell to the site location, ranged from relatively close in the Southern Alps (AMPS, ERA5, MERRA-2, UM) to relatively far on the West Coast of NZ (JRA-55) depending on the horizontal resolution of the grid. The time period examined was characterised by diverse cloud occurrence with periods of low cloud and precipitation, mid-level cloud, fog, high cloud and clear skies. Precipitation, currently not simulated by the lidar simulator, was present in about 18% of the observed attenuated volume backscattering coefficient profiles, as determined by visual inspection. Figure 3b shows that predominantly low cloud and precipitation between the ground and 3 km ASL in 25% of profiles was observed. Cloud between 3 and 12 km ASL was observed about evenly in 2% of profiles. While the reanalyses and models were able to partially reproduce the peak of cloud occurrence near 1 km ASL, the peak they displayed is less vertically broad than observed, and in the UM the peak was much weaker than observed. The lack of precipitation simulation might also have contributed to this apparent difference between observed and simulated cloud. Above 3 km ASL, the reanalyses and models tended to overestimate cloud, with only ERA5 and JRA-55 simulating close to the observed cloud occurrence. The observed total CF was 61%. AMPS overestimated this value by 5 percentage points (pp), ERA5 and the UM reproduced almost the exact value (within 1 pp), while the other reanalyses (JRA-55 and MERRA-2) underestimated CF by about 15 pp.

## Cass (2014-09-25)



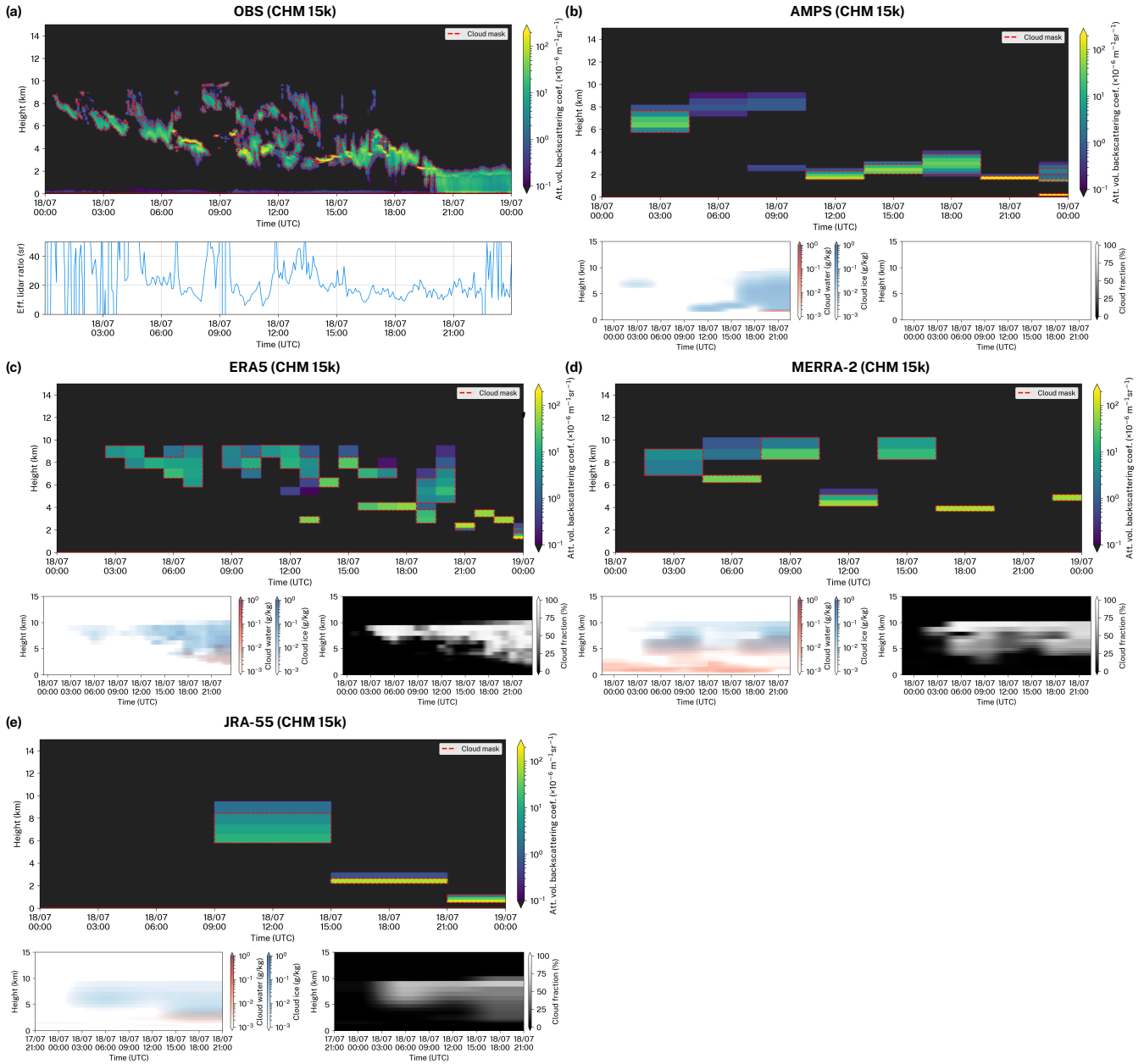
**Figure 4.** Examples of observed and simulated attenuated volume backscattering coefficient during 24 hours at Cass. The observed attenuated volume backscattering coefficient was normalised to absolute units and denoised. The first subcolumn generated by the Subgrid Cloud Overlap Profile Sampler (SCOPS) was used to make the plots. The red line is the station altitude. Shown is also (a) observed effective lidar ratio calculated by vertically integrating attenuated volume backscattering coefficient and (b–f) the corresponding model cloud liquid water, cloud ice and and cloud fraction.

## Lauder (2018-01-16)



**Figure 5.** The same as Fig. 4 but for the Lauder.

## Christchurch (2019-07-18)



**Figure 6.** The same as Fig. 4 but for the Christchurch.

## 7.2 Lauder

We also analysed 13 days of CL31 and MiniMPL observations at the Lauder station in summer. During the time period relatively diverse cloud was observed, with periods of low, mid- and high cloud, clear sky and a small fraction of profiles with precipitation (about 3%). The altitude of the station of 370 m ASL generally had a much higher equivalent in the reanalyses and models: 565 m (AMPS), 642 m (ERA5), 681 m (JRA-55) and 786 m (MERRA-2) due to the presence of hills in the surrounding region (the station is in a high valley), with the exception of the UM where the altitude was 385 m. The virtual station position in the reanalyses and models ranged from relatively close to the station in the same geographical region (AMPS, ERA5), a nearby location in a more hilly region (JRA-55), a relatively distant location in the adjacent Dunstan Mountains (MERRA-2) and a relatively distant location in Central Otago (UM). Figure 3c shows that the CL31 observed relatively even cloud occurrence between the ground and 3 km ASL at 8%, falling off to about 3% between 4 and 8 km ASL (the maximum lidar range of CL31 is 7.7 km). The MiniMPL observed much weaker attenuated volume backscattering coefficient than CL31 below 3 km ASL, which was identified as an overlap calibration issue in the MiniMPL. The MiniMPL observed substantial amounts of cloud above 8 km, not present in the CL31 observations due to its range limitation. Overall, the observed cloud occurrence had two peaks at ground to 3 km ASL and at about 9 km ASL. The simulated cloud occurrence was generally underestimated between the ground and 5 km ASL, with the exception of the UM which reproduced the lower half of the peak accurately, and ERA5 which reproduced the upper half of the peak accurately. Above 5 km ASL, the cloud occurrence was well reproduced in ERA5 and JRA-55, and strongly overestimated in AMPS, MERRA-2 and the UM. The reanalyses and models also tended to have two peaks at about 2 km ASL and 11 km ASL, but these were quite different from the observed peaks, with the lower peak underestimated by about 5 pp in the reanalyses and models and the higher peak overestimated by about 5–10 pp. The total CF was observed as 45% and 60% by CL31 and MiniMPL, respectively. CF observed by the MiniMPL was likely higher due to its higher maximum lidar range (CL31 missed a substantial amounts of high cloud due to this limitation). The total CF was strongly underestimated by the reanalyses and models by up to 31 pp (CL31) and 28 pp (MiniMPL), with the exception of the UM which simulated the correct CF within 3 pp.

## 7.3 Christchurch

The Christchurch observations were taken during a total of 23 days in mid- to late winter. The cloud situations were characterised by the frequent occurrence of low cloud and fog, with relatively diverse mid- and high level cloud and periods of clear sky also present (not shown). Precipitation was present in about 9% of profiles and fog in about 11% of profiles. As the site location is relatively flat (Canterbury Plains), the models did not have any difficulty in reproducing the altitude of the site, which was 32 m (AMPS), 72 m (ERA5), 143 m (JRA-55) and 76 m (MERRA-2). The virtual location was within the boundaries of the city (AMPS), on the Canterbury Plains close to the city boundaries (ERA5, MERRA-2), and over Lake Ellesmere about 20 km from the city (JRA-55). Figure 3d shows that the co-located CHM 15k and MiniMPL observed a strong peak of cloud occurrence of 26% (CHM 15k) at about 500 m ASL. This was likely due to the combined precipitation and fog as well as false detection of aerosol as cloud. The observed cloud occurrence had a local minimum of 2% at about 5 km ASL, a secondary peak

of 5% at 7 km ASL, and fell off 0% at 11 km ASL. The CHM 15k and MiniMPL observations showed inconsistencies of up to 4 pp. The reanalyses and models underestimated low cloud by 5–10 pp. With the exception of AMPS, they underestimated mid-level cloud by about 5 pp and represented high cloud relatively accurately. The total CF observed was 68%, while the reanalyses and models strongly underestimated CF by up to 34 pp (JRA-55), with underestimates around 20 pp common.

#### 5 7.4 Backscattering on daily scales

Figures 4, 5, 6 show images of attenuated volume backscattering coefficient for three separate days taken from the three case studies. The selected days represent some of the best-matching profiles and demonstrate how well the reanalyses and models can simulate cloud under favourable conditions. As can be seen in the figures, ERA5 and the UM perform the best in terms of temporal and height accuracy of the simulated cloud (Fig. 4c, 4f, 5c, 5f, 6c). This is likely due to the high output temporal resolution of the UM and ERA5 of 20 min. and 1 h, respectively. The UM and ERA5 were able to represent the relatively fine structure of cloud and to a lesser extent the optical thickness (inferred from the strength of backscattering) of the cloud. Deficiencies, however, are readily identifiable. The low cloud in the UM (Fig. 4f) covers too large area relative to observations (Fig. 4a) and the high cloud has a greater vertical extent in the UM. Likewise, the altocumulus cloud observed in Fig. 5a is shifted by several hours in the UM (Fig. 5f). The stratocumulus and nimbostratus cloud, identified visually based on the attenuated volume backscattering coefficient profiles, in ERA5 (Fig. 4c) is markedly lower than observed (Fig. 4a), as well as optically thicker than in reality. The mid-level cloud in ERA5 (Fig. 5c) was located about 2 km higher than observed (Fig. 5a). Precipitation observed in Fig. 6a towards the end of the analysed period was not present in the ERA5 simulated profile (Fig. 6c) due to lack of precipitation simulation in the current lidar simulator (even though rain and snow specific content is available from the reanalysis). AMPS and MERRA-2 had lower cloud representation accuracy. They managed to capture the overall structure of clouds (Fig. 4b, 4d, 5b, 5d, 6b, 6d), but substantial discrepancies were present, some of which were likely due to the relatively low temporal resolution of 3 h. AMPS has, however, relatively high horizontal grid resolution of 21 km. This demonstrates that other factors in the model than resolution have stronger influence on the quality of cloud simulation. JRA-55 was identified as the last in terms of cloud representation accuracy. JRA-55 has the lowest temporal resolution of the studied reanalyses and models of just 6 h, as well as the lowest horizontal grid resolution of 139 km. Therefore, it cannot be expected to capture any fine details of cloud. In the presented profiles (Fig. 4e, 5e, 6e) one can see that the cloud is only crudely represented. JRA-55 was able to represent the stratocumulus cloud of Fig. 4a, although its temporal extent and optical thickness were overestimated. The mid-level clouds of Fig. 5a and 6a were relatively well-represented in terms of height and optical thickness, given the low temporal resolution of the reanalysis. We stress that a direct attenuated volume backscattering coefficient profile intercomparison is highly dependent on the temporal resolution of the model output. The statistical intercomparison, however, should still give unbiased results if the cloud physics is accurately simulated by the atmospheric model.

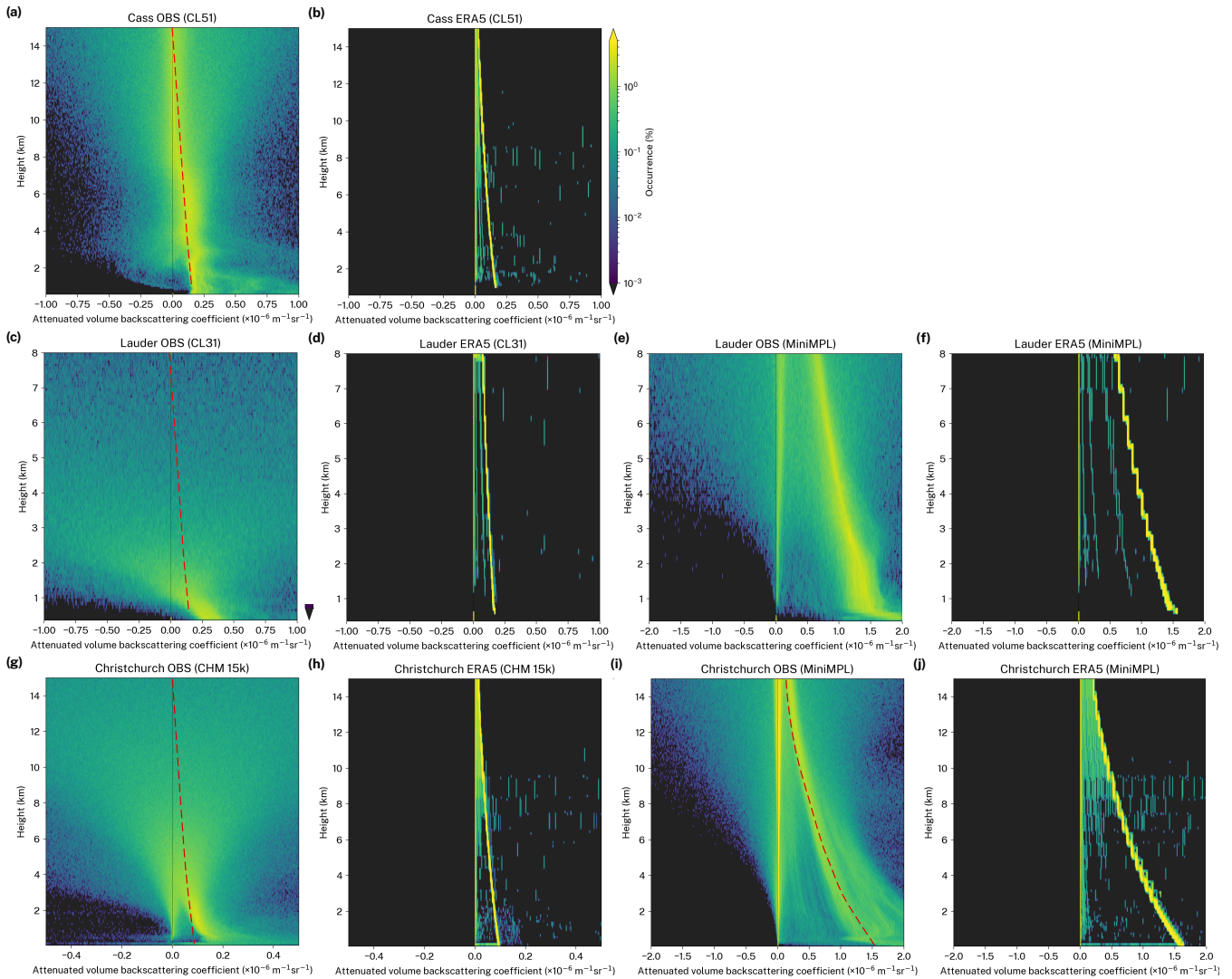
Figure 4a, 5a, 6a also show effective LR of observations calculated by integrating vertically attenuated volume backscattering coefficient (Sect. 4.1). If attenuated volume backscattering coefficient is properly calibrated, under fully attenuating cloud conditions effective LR converges to the theoretical value of LR of liquid cloud droplets (approximately 18.8 sr at near IR wavelengths) multiplied by the multiple scattering coefficient (approximately 0.7; Sect. 4.5).

## 7.5 Molecular backscattering, aerosol backscattering and noise

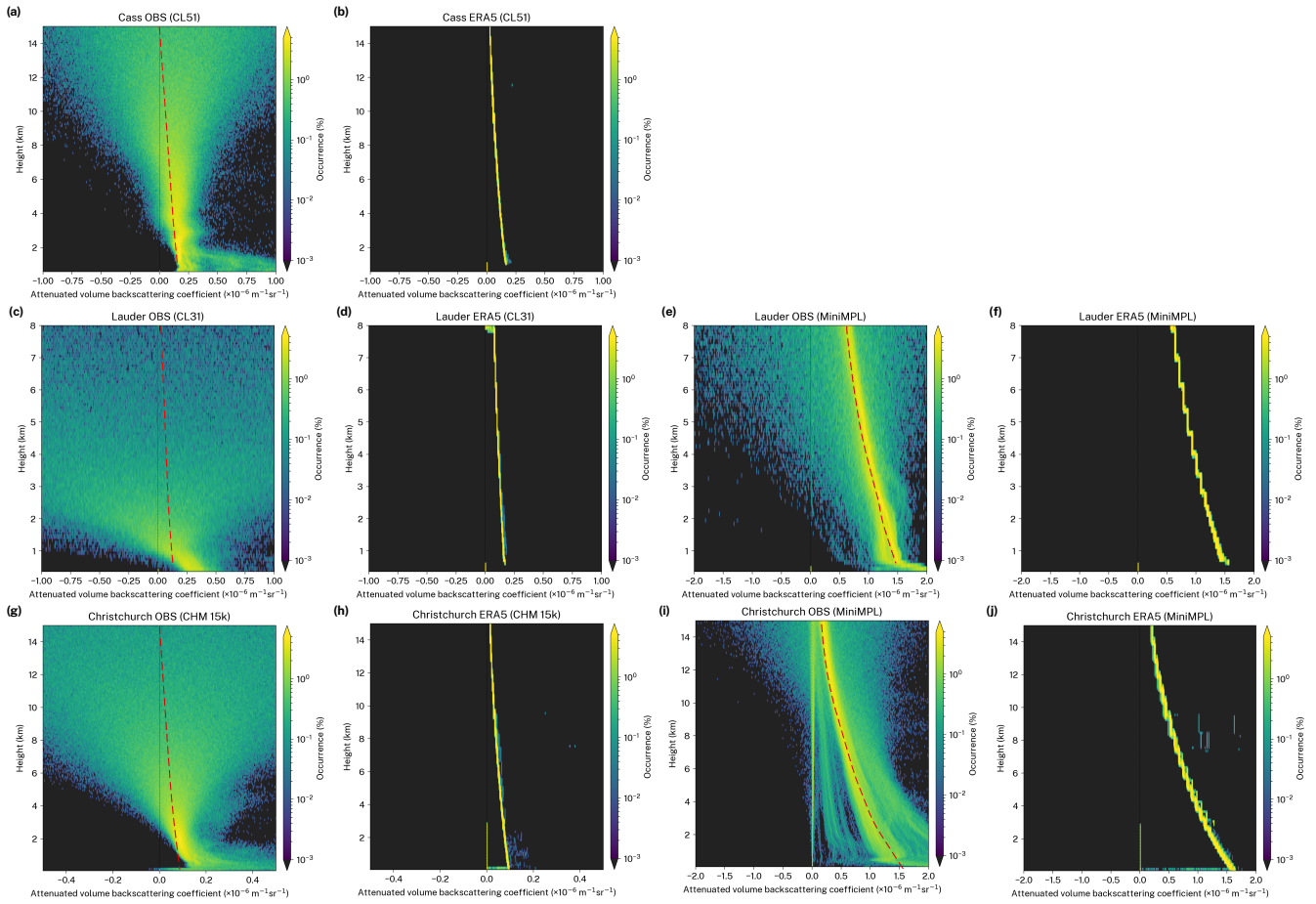
Figure 7 shows attenuated volume backscattering coefficient histograms as a function of height for small values of the coefficient (up to  $2 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ ) observed and simulated at the sites of the case studies, calculated for the entire time period of each case study. The scale of values is below cloud backscattering, and therefore shows backscattering which results from molecular and aerosol scattering and noise. Molecular backscattering depends on the atmospheric pressure and temperature as well as the lidar wavelength. It causes the main "streak" (a local maximum) visible in each of the histograms. The observed molecular attenuated volume backscattering coefficient at the surface approximately corresponds to the theoretically calculated value at each wavelength:  $0.0906 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  ( $\lambda = 1064 \text{ nm}$ ),  $0.172 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  ( $\lambda = 910 \text{ nm}$ ) and  $1.54 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  ( $\lambda = 532 \text{ nm}$ ) at 1000 hPa and  $20^\circ \text{C}$  (Table 5). The molecular backscattering in the boundary layer is, however, superimposed on backscattering by aerosol and cloud. In the case of the MiniMPL observations at the Christchurch site (Fig. 7i), the molecular attenuated volume backscattering coefficient streak has multiple secondary streaks. These are caused by different levels of attenuation by cloud and aerosol during the period of the observations. These secondary streaks were also partially reproduced by the simulator (Fig. 7j). A smaller portion of the width of the streak is also caused by fluctuations of atmospheric temperature and pressure. Under suitable conditions, molecular attenuated volume backscattering coefficient can be used for absolute calibration of an instrument. With the exception of CL31 (Fig. 7c), the molecular backscattering can be identified in the observed attenuated volume backscattering coefficient in each case. Therefore, it is possible to choose a calibration coefficient such that the observed and simulated molecular attenuated volume backscattering coefficients overlap. This can be considered a viable alternative to the liquid stratocumulus LR calibration method, or as a means of cross-validating the instrument calibration. However, it should be noted that the accuracy of this method is affected by an unknown amount of aerosol attenuation. Cloudy profiles can be filtered when calculating the histogram, and therefore the effect of cloud attenuation can be minimised. In addition to the molecular attenuated volume backscattering coefficient streak, there is a zero-centred streak visible in the histograms. This is caused by noise when the signal is fully attenuated by cloud. Lastly, a zero-centred "cone" of noise is visible in the observed attenuated volume backscattering coefficient, increasing with the square of range. The size of this cone is particularly large in the case of the CL31 (Fig. 7c), which is most likely the result of its low receiver sensitivity and low power compared to the other instruments. The standard deviation of the cone at the furthest range is used to determine the noise standard deviation used by the cloud detection algorithm (Sect. 5.3).

Figure 8 shows the same information as Fig. 7, but for clear sky profiles only. Here, it can be seen that the zero-centred peak caused by the complete attenuation by cloud is no longer present. There is a clear overlap between the centre of the noise cone with the simulated molecular attenuated volume backscattering coefficient; i.e. the noise cone is centred at the observed molecular attenuated volume backscattering coefficient. This is visible with all instruments including CL31 (Fig. 8c), where the overlap between the observed and simulated molecular attenuated volume backscattering coefficient is most clearly visible at about 1 km ASL. Below 1 km ASL, the effect of boundary layer aerosol distorts molecular attenuated volume backscattering coefficient by an unknown quantity. The clear sky histograms as shown in Fig. 8 may therefore be preferable to the all-sky histograms of Fig. 7 for calibration by fitting molecular attenuated volume backscattering coefficient. The dead





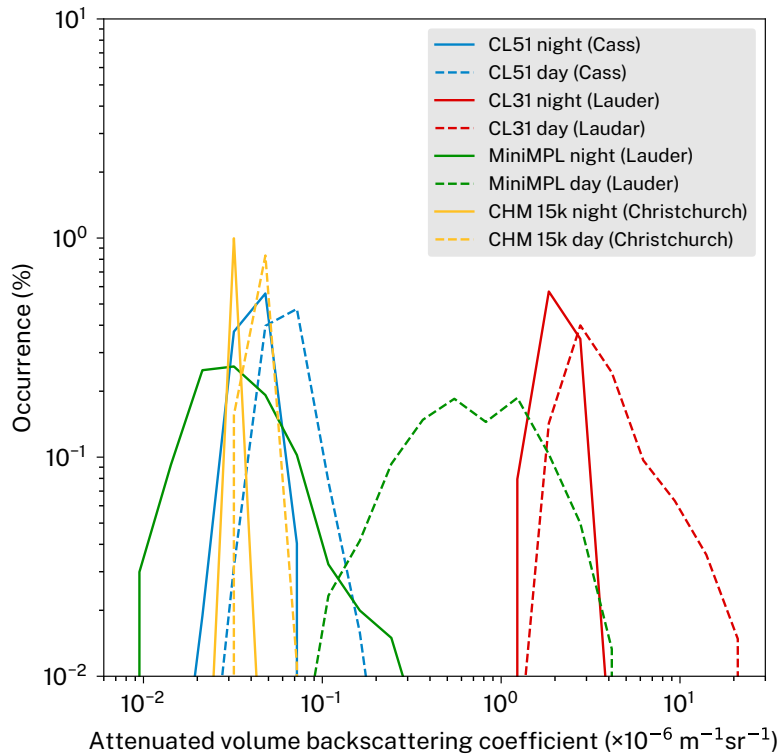
**Figure 7.** Attenuated volume backscattering coefficient histograms as a function of height observed and simulated at three different sites of the case studies calculated from all profiles. The plots show the distribution of attenuated volume backscattering coefficient for values which are on the scale of noise, molecular and aerosol backscattering ( $[-0.5, 0.5]$  for CHM 15k,  $[-1, 1]$  for CL31 and CL51 and  $[-2, 2] \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  for MiniMPL). The simulated attenuated volume backscattering coefficient is based on the ERA5 atmospheric fields. Visible in the plots is backscattering caused by molecular backscattering (the main "streak"), noise when signal is fully attenuated by cloud (the zero-centred "streak"), and the range-dependent noise (the zero-centred "cone"). The molecular backscattering is marked by a red dashed line on the observed attenuated volume backscattering coefficient plots, the shape of which is taken from the simulated molecular attenuated volume backscattering coefficient for the corresponding instrument and site.



**Figure 8.** The same as Fig. 7 but calculated from clear sky profiles only.

time, afterpulse and overlap MiniMPL calibration supplied by the vendor appears to be deficient and causes range-dependent bias in the attenuated volume backscattering coefficient profile.

We now examine the noise in each instrument using the ALCF. Figure 9 shows the distribution of standard deviation of backscatter noise determined at the highest observable range of each instrument and range-scaled to 8 km. It can be seen that the CL31 is affected by the greatest amount of noise, peaking at about  $2 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ . This is at the threshold of cloud detection of  $2 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ . Therefore, thin cloud may be obscured by noise at higher ranges with this instrument. The MiniMPL, operating in the visible spectral range, shows a strongly bimodal distribution of the attenuated volume backscattering coefficient noise depending on sunlight. During daytime, it peaks at about  $0.7 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ , which is the second highest of the analysed instruments. During nighttime, it peaks at about  $0.02 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ , which is the lowest of the analysed instruments. The CHM 15k and CL51 peak between the nighttime and daytime MiniMPL at about  $0.05 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ . All of CL31, CL51 and CHM 15k show a slight reduction of noise during nighttime, presumably because of a small amount of incoming solar radiation



**Figure 9.** Attenuated volume backscattering coefficient noise standard deviation histogram calculated for each instrument at sites of the case studies from clear sky profiles over the whole time period. The noise distribution is calculated at the furthest range. Shown is the range-scaled noise distribution at a range of 8 km. "Night" and "day" distributions are calculated separately from nighttime and daytime profiles only.

at near IR wavelengths. The difference between the nighttime and daytime attenuated volume backscattering coefficient noise in the MiniMPL has been previously analysed by Silber et al. (2018) (Fig. S3) and these results confirm their findings.

## 8 Discussion and conclusions

We presented the Automatic Lidar and Ceilometer Framework, which combines lidar processing and lidar simulation for the purpose of model evaluation. The lidar simulation is based on the COSP spaceborne lidar simulator by accounting for the different geometry and lidar wavelength. We calculated new lookup tables for Mie scattering for a number of ALC wavelengths, developed ice crystal backscattering parametrisation based on temperature and implemented noise removal and cloud detection algorithms. The framework supports the most common ALCs and reanalyses. We demonstrated the use of the framework on ALC observations at three different sites in New Zealand, and applied the lidar simulator to three reanalyses and two models. We found that while some reanalyses and models such as the UM and ERA5 show relatively good correspondence with observed cloud, others performed relatively poorly in our time-limited local comparison. All reanalyses and models

underestimated the total CF by up to 34 pp, with underestimation by 20 pp common. In some cases, the observed and simulated attenuated volume backscattering coefficient profiles matched relatively closely in terms of time and altitude, and a better match was observed with reanalyses with high output temporal resolution such as the UM and ERA5, while reanalyses with low temporal resolution did not allow for reliable direct (non-statistical) comparison of cloud. However, it is clear that more factors than the horizontal and vertical resolution influence the cloud simulation accuracy; especially the cloud, boundary layer and convection schemes employed by the atmospheric model. The reanalysis and model output temporal, horizontal grid resolution and vertical resolution are not always the same as the internal resolution of the underlying atmospheric model. Both have an impact on the comparison between the simulated and observed attenuated volume backscattering coefficient and cloud. While the output resolution should not have an impact on the long-term statistics, it can be a limiting factor for direct attenuated volume backscattering coefficient profile comparison. We demonstrated that the ALCF could be used to identify substantial differences in cloud attenuated volume backscattering coefficient which were present in all reanalyses and models. We showed that all the studied instruments except for the CL31 are capable of detecting molecular backscattering and that this can be used for calibration or for cross-validation of other calibration methods. We found that the nighttime MiniMPL was subject to the least amount of noise of all the instrument examined, followed by the CL51, CHM 15k, daytime MiniMPL and CL31. Noise in the MiniMPL, and to a lesser extent in the other ALCs, was shown to have a bimodal distribution due to day/nighttime. The ALCF can therefore be useful for testing the quality of collected data.

Currently the framework has several limitations which should be addressed in the future. The water vapour absorption at 910 nm likely affects the instrument calibration of the CL31 and CL51 ceilometers and limits the accuracy of the one-to-one comparison, even though due to the relatively high backscattering caused by cloud, the calculated cloud masks are unlikely to be strongly affected. The lidar simulator currently does not simulate backscattering from precipitation. Observed precipitation is generally detected as "cloud" by the cloud detection algorithm, while the simulated profile contains no backscattering at the location of precipitation (backscattering and attenuation by rain drops and snow should be implemented in the lidar simulator in the future). If desired, the attenuated volume backscattering coefficient profiles affected by precipitation can be excluded before the comparison or their fraction determined by visually inspecting the observed attenuated volume backscattering to assess their possible effect on the statistical results. Aerosol is also not currently implemented in the simulator. Previous studies (Chan et al., 2018) characterised optical parameters of different groups of aerosol, which could be used in a future version of the simulator with models which provide concentration of aerosol in their output. In our case studies aerosol volume backscattering coefficient was less than  $2 \times 10^{-6} \text{ m}^{-1} \text{ sr}^{-1}$  and below 4 km, which could result in worst-case two-way attenuation of about 50% assuming LR of 50 sr. This should not preclude cloud detection due to the large magnitude of typical cloud backscattering. The ALCs also suffer from various measurement deficiencies. Notably incomplete overlap, dead time and afterpulse corrections tend to give sub-optimal results at the near range. It is possible to use semi-automated methods to correct for these deficiencies, such as by calculating the integrated attenuated volume backscattering coefficient distribution by height of the maximum backscattering and correcting for the range-dependent bias (Hopkin et al., 2019, Sect. 5.1). This method could be implemented in the framework to enable range-dependent calibration of the observed attenuated volume backscattering coefficient.

The presented framework streamlines lidar data processing and tasks related to lidar simulation and model comparison. The framework was recently used by Kuma et al. (2020) for Southern Ocean model cloud evaluation in the GA7.1 model and MERRA-2 reanalysis. Considering the existing extensive ALC networks worldwide there is a wealth of global data. We therefore think that ALCs should have a greater role in model evaluation. Satellite observations have long been established in this respect due to their availability, spatial and temporal coverage and their well-developed derived products and tools. ALCs, with their diverse formats and decentralised nature, have so far lacked derived products and tools which would make them more accessible for model evaluation. We hope that this software will enable more model evaluation studies based on ALC observations. Development of lidar data processing is currently hampered by closed development of code. We note that code has very rarely been made available with past ALC studies. Continued improvement of publicly available code for lidar data processing is needed to achieve faster development of ground-based remote sensing and make it more attractive for GCM, NWP model and reanalysis evaluation.

*Code and data availability.* The *ALCF* is open source and available at <https://alcf-lidar.github.io> and as a permanent archive of code and technical documentation on Zenodo at <https://doi.org/10.5281/zenodo.4088217>. The technical documentation is also in the Supplementary information. A tool for converting Vaisala CL31 and CL51 data files to NetCDF *cl2nc* is open source and available at <https://github.com/peterkuma/cl2nc>. A tool for converting MiniMPL raw binary data files to NetCDF *mpl2nc* is open source and available at <https://github.com/peterkuma/mpl2nc>. The observational data used in the case studies are available upon request. The reanalyses data used in the case studies are publicly available online from the respective projects. The *Unified Model* data used in the case studies are available upon request. The Unified Model is proprietary to the UK Met Office and is made available under a licence. For more information, readers are advised to contact the UK Met Office.

*Author contributions.* Peter Kuma wrote the code of the framework, performed the data analysis of the case studies and wrote the text of the manuscript. Adrian McDonald and Olaf Morgenstern provided continuous scientific input on the code development, analysis and text of the manuscript. Richard Querel, Israel Silber and Connor J. Flynn provided calibration of the MiniMPL data and substantial discussion of the theoretical concepts. All authors reviewed the manuscript.

*Competing interests.* The authors declare that they have no competing interests.

*Acknowledgements.* We would like to acknowledge the New Zealand Deep South National Science Challenge project which provided funding for our work; the New Zealand eScience Infrastructure (NeSI) which provided supercomputing resources to run the Unified Model; Vidya Varma, Jonny Williams, Guang Zeng and Wolfgang Hayek for their contribution to setting up a nudged run of the Unified Model; Graeme Plank and Graeme MacDonald who participated on the installation of the Vaisala CL51 at the Cass field station; the COSP project

for the code which we used as the basis for the lidar simulator; the AMPS, JRA-55, ERA5, MERRA-2 models and reanalyses which provided public access to their data; the open source libraries numpy (Van Der Walt et al., 2011), scipy (Virtanen et al., 2019), matplotlib (Hunter, 2007), netCDF4 (Rew and Davis, 1990) and Astropy (Price-Whelan et al., 2018) and the Python programming language (Rossum, 1995) which we used in the implementation of our code; the R programming language (R Core Team, 2017), the Natural Earth dataset (https://www.naturalearthdata.com) and the Shuttle Radar Topography Mission (SRTM) Version 3 Global 1 arc second digital elevation model (Werner, 2001; NASA JPL, 2013) which we used to produce a map of sites; GitHub which provided free hosting of our code; and the Linux-based (Torvalds, 1997) operating systems Devuan GNU+Linux and Debian GNU/Linux on which we produced this analysis.

## References

- G-SDSU, <https://cloud.gsfc.nasa.gov/index.php?section=14>, 2019.
- Baars, H., Kanitz, T., Engelmann, R., Althausen, D., Heese, B., Komppula, M., Preißler, J., Tesche, M., Ansmann, A., Wandinger, U., Lim, J.-H., Ahn, J. Y., Stachlewska, I. S., Amiridis, V., Marinou, E., Seifert, P., Hofer, J., Skupin, A., Schneider, F., Bohlmann, S., Foth, A., Bley, S., Pfüller, A., Giannakaki, E., Lihavainen, H., Viisanen, Y., Hooda, R. K., Pereira, S. N., Bortoli, D., Wagner, F., Mattis, I., Janicka, L., Markowicz, K. M., Achtert, P., Artaxo, P., Pauliquevis, T., Souza, R. A. F., Sharma, V. P., van Zyl, P. G., Beukes, J. P., Sun, J., Rohwer, E. G., Deng, R., Mamouri, R.-E., and Zamorano, F.: An overview of the first decade of PollyNET: an emerging network of automated Raman-polarization lidars for continuous aerosol profiling, *Atmospheric Chemistry and Physics*, 16, 5111–5137, <https://doi.org/10.5194/acp-16-5111-2016>, 2016.
- 5
- 10 Baran, A. J.: A review of the light scattering properties of cirrus, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110, 1239–1260, <https://doi.org/10.1016/j.jqsrt.2009.02.026>, 2009.
- Bastin, S., Chiriaco, M., and Drobinski, P.: Control of radiation and evaporation on temperature variability in a WRF regional climate simulation: comparison with colocated long term ground based observations near Paris, *Climate dynamics*, 51, 985–1003, <https://doi.org/10.1007/s00382-016-2974-1>, 2018.
- 15 Bi, L., Yang, P., Kattawar, G. W., Baum, B. A., Hu, Y. X., Winker, D. M., Brock, R. S., and Lu, J. Q.: Simulation of the color ratio associated with the backscattering of radiation by ice particles at the wavelengths of 0.532 and 1.064  $\mu\text{m}$ , *Journal of Geophysical Research*, 114, <https://doi.org/10.1029/2009jd011759>, 2009.
- Bodas-Salcedo, A., Webb, M., Bony, S., Chepfer, H., Dufresne, J.-L., Klein, S., Zhang, Y., Marchand, R., Haynes, J., Pincus, R., et al.: COSP: Satellite simulation software for model assessment, *Bulletin of the American Meteorological Society*, 92, 1023–1043, <https://doi.org/10.1175/2011BAMS2856.1>, 2011.
- 20 Borovoi, A., Konoshonkin, A., and Kustova, N.: Backscatter ratios for arbitrary oriented hexagonal ice crystals of cirrus clouds, *Opt. Lett.*, 39, 5788–5791, <https://doi.org/10.1364/OL.39.005788>, 2014.
- Bouniol, D., Protat, A., Delanoë, J., Pelon, J., Piriou, J.-M., Bouysse, F., Tompkins, A. M., Wilson, D. R., Morille, Y., Haefelin, M., et al.: Using continuous ground-based radar and lidar measurements for evaluating the representation of clouds in four operational models, *Journal of Applied Meteorology and Climatology*, 49, 1971–1991, <https://doi.org/10.1175/2010JAMC2333.1>, 2010.
- 25 Bréon, F.-M. and Doutriaux-Boucher, M.: A comparison of cloud droplet radii measured from space, *IEEE Transactions on Geoscience and Remote Sensing*, 43, 1796–1805, <https://doi.org/10.1109/TGRS.2005.852838>, 2005.
- Bréon, F.-M. and Colzy, S.: Global distribution of cloud droplet effective radius from POLDER polarization measurements, *Geophysical Research Letters*, 27, 4065–4068, <https://doi.org/10.1029/2000GL011691>, 2000.
- 30 Campbell, J. R., Hlavka, D. L., Welton, E. J., Flynn, C. J., Turner, D. D., Spinhirne, J. D., Scott III, V. S., and Hwang, I.: Full-time, eye-safe cloud and aerosol lidar observation at atmospheric radiation measurement program sites: Instruments and data processing, *Journal of Atmospheric and Oceanic Technology*, 19, 431–442, [https://doi.org/10.1175/1520-0426\(2002\)019<0431:FTESCA>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<0431:FTESCA>2.0.CO;2), 2002.
- Cazorla, A., Casquero-Vera, J. A., Román, R., Guerrero-Rascado, J. L., Toledano, C., Cachorro, V. E., Orza, J. A. G., Cancillo, M. L., Serrano, A., Titos, G., Pandolfi, M., Alastuey, A., Hanrieder, N., and Alados-Arboledas, L.: Near-real-time processing of a ceilometer network assisted with sun-photometer data: monitoring a dust outbreak over the Iberian Peninsula, *Atmospheric Chemistry and Physics*, 17, 11 861–11 876, <https://doi.org/10.5194/acp-17-11861-2017>, 2017.
- 35

- Chan, K. L., Wiegner, M., Flentje, H., Mattis, I., Wagner, F., Gasteiger, J., and Geiß, A.: Evaluation of ECMWF-IFS (version 41R1) operational model forecasts of aerosol transport by using ceilometer network measurements, *Geoscientific Model Development*, 11, 3807–3831, <https://doi.org/10.5194/gmd-11-3807-2018>, 2018.
- Chang, F. and Li, Z.: The effect of droplet size distribution on the determination of cloud droplet effective radius, in: 11th ARM Science Team Meeting, Atlanta, Ga, pp. 19–23, 2001.
- Chepfer, H., Chiriaco, M., Vautard, R., and Spinhirne, J.: Evaluation of MM5 optically thin clouds over Europe in fall using ICESat lidar spaceborne observations, *Monthly weather review*, 135, 2737–2753, <https://doi.org/10.1175/MWR3413.1>, 2007.
- Chepfer, H., Bony, S., Winker, D., Chiriaco, M., Dufresne, J.-L., and Sèze, G.: Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model, *Geophysical Research Letters*, 35, <https://doi.org/10.1029/2008GL034207>, 2008.
- 10 Chiriaco, M., Vautard, R., Chepfer, H., Haefelin, M., Dudhia, J., Wanherdrick, Y., Morille, Y., and Protat, A.: The ability of MM5 to simulate ice clouds: Systematic comparison between simulated and measured fluxes and lidar/radar profiles at the SIRTa atmospheric observatory, *Monthly weather review*, 134, 897–918, <https://doi.org/10.1175/MWR3102.1>, 2006.
- Chiriaco, M., Dupont, J.-C., Bastin, S., Badosa, J., Lopez, J., Haefelin, M., Chepfer, H., and Guzman, R.: ReOBS: a new approach to synthesize long-term multi-variable dataset and application to the SIRTa supersite, *Earth System Science Data*, 10, 919, <https://doi.org/10.5194/essd-10-919-2018>, 2018.
- 15 Costa-Surós, M., Calbó, J., González, J., and Martin-Vide, J.: Behavior of cloud base height from ceilometer measurements, *Atmospheric Research*, 127, 64–76, <https://doi.org/10.1016/j.atmosres.2013.02.005>, 2013.
- Cromwell, E. and Flynn, D.: Lidar Cloud Detection with Fully Convolutional Networks, 2018.
- Cromwell, E. and Flynn, D.: Lidar cloud detection with fully convolutional networks, in: 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 619–627, IEEE, 2019.
- 20 Dee, D. P., Uppala, S. M., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, d. P., et al.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Quarterly Journal of the royal meteorological society*, 137, 553–597, <https://doi.org/10.1002/qj.828>, 2011.
- Diner, D. J., Beckert, J. C., Reilly, T. H., Bruegge, C. J., Conel, J. E., Kahn, R. A., Martonchik, J. V., Ackerman, T. P., Davies, R., Gerstl, S. A., et al.: Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview, *IEEE Transactions on Geoscience and Remote Sensing*, 36, 1072–1087, <https://doi.org/10.1109/36.700992>, 1998.
- 25 Dionisi, D., Barnaba, F., Diémoz, H., Di Liberto, L., and Gobbi, G. P.: A multiwavelength numerical model in support of quantitative retrievals of aerosol properties from automated lidar ceilometers and test applications for AOT and PM 10 estimation, *Atmospheric Measurement Techniques*, 11, 6013, <https://doi.org/10.5194/amt-11-6013-2018>, 2018.
- 30 Ebata, A., Kobayashi, S., Ota, Y., Moriya, M., Kumabe, R., Onogi, K., Harada, Y., Yasui, S., Miyaoka, K., Takahashi, K., et al.: The Japanese 55-year reanalysis “JRA-55”: an interim report, *Sola*, 7, 149–152, <https://doi.org/10.2151/jmsj.2016-015>, 2011.
- ECMWF: Copernicus Climate Change Service (C3S) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate, Copernicus Climate Change Service Climate Data Store (CDS), <https://doi.org/10.24381/cds.bd0915c6>, <https://cds.climate.copernicus.eu/cdsapp#!/home>, 2019.
- 35 Edwards, J. and Slingo, A.: Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model, *Quarterly Journal of the Royal Meteorological Society*, 122, 689–719, <https://doi.org/10.1002/qj.49712253107>, 1996.
- Emeis, S.: Surface-based remote sensing of the atmospheric boundary layer, vol. 40, Springer Science & Business Media, <https://doi.org/10.1007/978-90-481-9340-0>, 2010.



- Emeis, S., Schäfer, K., and Münkel, C.: Observation of the structure of the urban boundary layer with different ceilometers and validation by RASS data, *Meteorologische Zeitschrift*, 18, 149–154, <https://doi.org/10.1127/0941-2948/2009/0365>, 2009.
- Eresmaa, N., Karppinen, A., Joffe, S., Räsänen, J., and Talvitie, H.: Mixing height determination by ceilometer, *Atmospheric Chemistry and Physics*, 6, 1485–1493, <https://doi.org/10.5194/acp-6-1485-2006>, 2006.
- 5 Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geoscientific Model Development*, 9, 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016>, 2016.
- Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., Collins, W. D., Gier, B. K., Hall, A. D., Hoffman, F. M., et al.: Taking climate model evaluation to the next level, *Nature Climate Change*, p. 1, <https://doi.org/10.1038/s41558-018-0355-y>, 2019.
- 10 Flynn, C. J., Mendozaa, A., Zhengb, Y., and Mathurb, S.: Novel polarization-sensitive micropulse lidar measurement technique, *Optics express*, 15, 2785–2790, <https://doi.org/10.1364/OE.15.002785>, 2007.
- Fu, D., Di Girolamo, L., Liang, L., and Zhao, G.: Regional Biases in MODIS Marine Liquid Water Cloud Drop Effective Radius Deduced Through Fusion With MISR, *Journal of Geophysical Research: Atmospheres*, 124, 13 182–13 196, <https://doi.org/10.1029/2019JD031063>, 2019.
- 15 Garnier, A., Pelon, J., Vaughan, M. A., Winker, D. M., Treppe, C. R., and Dubuisson, P.: Lidar multiple scattering factors inferred from CALIPSO lidar and IIR retrievals of semi-transparent cirrus cloud optical depths over oceans, *Atmospheric Measurement Techniques*, 8, 2759–2774, <https://doi.org/10.5194/amt-8-2759-2015>, 2015.
- Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., et al.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), *Journal of Climate*, 30, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>, 2017.
- 20 Geleyn, J. and Hollingsworth, A.: An economical analytical method for the computation of the interaction between scattering and line absorption of radiation, *Contributions to Atmospheric Physics*, 52, 1979.
- Goody, R. M. and Yung, Y. L.: *Atmospheric radiation: theoretical basis*, Oxford University Press, 2 edn., 1995.
- Hansen, A., Ament, F., Grützun, V., and Lammert, A.: Model evaluation by a cloud classification based on multi-sensor observations, *Geoscientific Model Development Discussions*, 2018, 1–20, <https://doi.org/10.5194/gmd-2018-259>, 2018a.
- 25 Hansen, A., Ament, F., Grützun, V., and Lammert, A.: Model evaluation by a cloud classification based on multi-sensor observations, *Geoscientific Model Development Discussions*, 2018, 1–20, <https://doi.org/10.5194/gmd-2018-259>, 2018b.
- Harada, Y., Kamahori, H., Kobayashi, C., Endo, H., Kobayashi, S., Ota, Y., Onoda, H., Onogi, K., Miyaoka, K., and Takahashi, K.: The JRA-55 Reanalysis: Representation of atmospheric circulation and climate variability, *Journal of the Meteorological Society of Japan*. Ser. II, 94, 269–302, <https://doi.org/10.2151/jmsj.2016-015>, 2016.
- 30 Heese, B., Flentje, H., Althausen, D., Ansmann, A., and Frey, S.: Ceilometer lidar comparison: backscatter coefficient retrieval and signal-to-noise ratio determination, *Atmospheric Measurement Techniques*, 3, 1763–1770, <https://doi.org/10.5194/amt-3-1763-2010>, 2010.
- Heymsfield, A. J.: Extinction-ice water content-effective radius algorithms for CALIPSO, *Geophysical Research Letters*, 32, <https://doi.org/10.1029/2005gl022742>, 2005.
- 35 Hines, K. M. and Bromwich, D. H.: Development and testing of Polar Weather Research and Forecasting (WRF) model. Part I: Greenland ice sheet meteorology, *Monthly Weather Review*, 136, 1971–1989, <https://doi.org/10.1175/2007MWR2112.1>, 2008.
- Hogan, R. J.: Fast approximate calculation of multiply scattered lidar returns, *Applied Optics*, 45, 5984–5992, <https://doi.org/10.1364/AO.45.005984>, 2006.

- Hogan, R. J., Jakob, C., and Illingworth, A. J.: Comparison of ECMWF Winter-Season Cloud Fraction with Radar-Derived Values, *Journal of Applied Meteorology*, 40, 513–525, [https://doi.org/10.1175/1520-0450\(2001\)040<0513:COEWSC>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<0513:COEWSC>2.0.CO;2), 2001.
- Hopkin, E.: Use of a calibrated ceilometer network to improve high resolution weather forecasts, Ph.D. thesis, University of Reading, UK, 2018.
- 5 Hopkin, E., Illingworth, A. J., Charlton-Perez, C., Westbrook, C. D., and Ballard, S.: A robust automated technique for operational calibration of ceilometers using the integrated backscatter from totally attenuating liquid clouds, *Atmospheric Measurement Techniques Discussions*, 2019, 1–28, <https://doi.org/10.5194/amt-2018-427>, 2019.
- Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., Folini, D., Ji, D., Klocke, D., Qian, Y., et al.: The art and science of climate model tuning, *Bulletin of the American Meteorological Society*, 98, 589–602, <https://doi.org/10.1175/BAMS-D-15-00135.1>,  
10 2017.
- Hu, Y.: Depolarization ratio–effective lidar ratio relation: Theoretical basis for space lidar cloud phase discrimination, *Geophysical Research Letters*, 34, <https://doi.org/10.1029/2007GL029584>, 2007.
- Hu, Y., Vaughan, M., McClain, C., Behrenfeld, M., Maring, H., Anderson, D., Sun-Mack, S., Flittner, D., Huang, J., Wielicki, B., Minnis, P., Weimer, C., Trepte, C., and Kuehn, R.: Global statistics of liquid water content and effective number concentration of water  
15 clouds over ocean derived from combined CALIPSO and MODIS measurements, *Atmospheric Chemistry and Physics*, 7, 3353–3359, <https://doi.org/10.5194/acp-7-3353-2007>, 2007.
- Hunter, J. D.: Matplotlib: A 2D graphics environment, *Computing in science & engineering*, 9, 90, <https://doi.org/10.1109/MCSE.2007.55>, 2007.
- Illingworth, A., Hogan, R., O’connor, E., Bouniol, D., Brooks, M., Delanoë, J., Donovan, D., Eastment, J., Gaussiat, N., Goddard, J., et al.:  
20 Cloudnet: Continuous evaluation of cloud profiles in seven operational models using ground-based observations, *Bulletin of the American Meteorological Society*, 88, 883–898, <https://doi.org/10.1175/BAMS-88-6-883>, 2007.
- Illingworth, A., Cimini, D., Haeffele, A., Haeffelin, M., Hervo, M., Kotthaus, S., Löhnert, U., Martinet, P., Mattis, I., O’Connor, E., et al.: How can Existing Ground-Based Profiling Instruments Improve European Weather Forecasts?, *Bulletin of the American Meteorological Society*, <https://doi.org/10.1175/BAMS-D-17-0231.1>, 2018.
- 25 Illingworth, A. J., Barker, H., Beljaars, A., Ceccaldi, M., Chepfer, H., Clerbaux, N., Cole, J., Delanoë, J., Domenech, C., Donovan, D. P., et al.: The EarthCARE satellite: The next step forward in global measurements of clouds, aerosols, precipitation, and radiation, *Bulletin of the American Meteorological Society*, 96, 1311–1332, <https://doi.org/10.1175/BAMS-D-12-00227.1>, 2015a.
- Illingworth, A. J., Cimini, D., Gaffard, C., Haeffelin, M., Lehmann, V., Löhnert, U., O’Connor, E. J., and Ruffieux, D.: Exploiting existing ground-based remote sensing networks to improve high-resolution weather forecasts, *Bulletin of the American Meteorological Society*,  
30 96, 2107–2125, <https://doi.org/10.1175/BAMS-D-13-00283.1>, 2015b.
- Jin, Y., Kai, K., Kawai, K., Nagai, T., Sakai, T., Yamazaki, A., Uchiyama, A., Batdorj, D., Sugimoto, N., and Nishizawa, T.: Ceilometer calibration for retrieval of aerosol optical properties, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 153, 49–56, <https://doi.org/10.1016/j.jqsrt.2014.10.009>, 2015.
- Josset, D., Pelon, J., Garnier, A., Hu, Y., Vaughan, M., Zhai, P.-W., Kuehn, R., and Lucker, P.: Cirrus optical depth and lidar ratio retrieval  
35 from combined CALIPSO-CloudSat observations using ocean surface echo, *Journal of Geophysical Research: Atmospheres*, 117, n/a–n/a, <https://doi.org/10.1029/2011jd016959>, 2012.
- Klein, S. A. and Jakob, C.: Validation and sensitivities of frontal clouds simulated by the ECMWF model, *Monthly Weather Review*, 127, 2514–2531, [https://doi.org/10.1175/1520-0493\(1999\)127<2514:VASOFC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<2514:VASOFC>2.0.CO;2), 1999.

- Klekociuk, A. R., French, W. J. R., Alexander, S. P., Kuma, P., and McDonald, A. J.: The state of the atmosphere in the 2016 southern Kerguelen Axis campaign region, *Deep Sea Research Part II: Topical Studies in Oceanography*, <https://doi.org/10.1016/j.dsr2.2019.02.001>, 2019.
- 5 Knepp, T. N., Szykman, J. J., Long, R., Duvall, R. M., Krug, J., Beaver, M., Cavender, K., Kronmiller, K., Wheeler, M., Delgado, R., Hoff, R., Berkoff, T., Olson, E., Clark, R., Wolfe, D., Van Gilst, D., and Neil, D.: Assessment of mixed-layer height estimation from single-wavelength ceilometer profiles, *Atmospheric Measurement Techniques*, 10, 3963–3983, <https://doi.org/10.5194/amt-10-3963-2017>, 2017.
- Kobayashi, S., Ota, Y., Harada, Y., Ebata, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., et al.: The JRA-55 reanalysis: General specifications and basic characteristics, *Journal of the Meteorological Society of Japan. Ser. II*, 93, 5–48, <https://doi.org/10.2151/jmsj.2015-001>, 2015.
- 10 Kotthaus, S., O’Connor, E., Munkel, C., Charlton-Perez, C., Haeffelin, M., Gabey, A. M., and Grimmond, C. S. B.: Recommendations for processing atmospheric attenuated backscatter profiles from Vaisala CL31 ceilometers, *Atmospheric Measurement Techniques*, 9, 3769–3791, <https://doi.org/10.5194/amt-9-3769-2016>, 2016.
- Kuma, P., McDonald, A., Morgenstern, O., Alexander, S. P., Cassano, J. J., Garrett, S., Halla, J., Hartery, S., Harvey, M. J., Parsons, S., Plank, G., Varma, V., and Williams, J.: Evaluation of Southern Ocean cloud in the HadGEM3 general circulation model and MERRA-2 reanalysis using ship-based observations, *Atmospheric Chemistry and Physics*, 20, 6607–6630, <https://doi.org/10.5194/acp-20-6607-2020>, 2020.
- 15 Lamer, K., Fridlind, A. M., Ackerman, A. S., Kollias, P., Clothiaux, E. E., and Kelley, M.: (GO)2-SIM: a GCM-oriented ground-observation forward-simulator framework for objective evaluation of cloud and precipitation phase, *Geoscientific Model Development (Online)*, 11, <https://doi.org/10.5194/gmd-11-4195-2018>, 2018.
- Lewis, J. R., Campbell, J. R., Welton, E. J., Stewart, S. A., and Haftings, P. C.: Overview of MPLNET version 3 cloud detection, *Journal of Atmospheric and Oceanic Technology*, 33, 2113–2134, <https://doi.org/10.1175/JTECH-D-15-0190.1>, 2016.
- 20 Liou, K.-N.: An introduction to atmospheric radiation, vol. 84, Elsevier, 2 edn., 2002.
- Liu, J., Li, Z., Zheng, Y., and Cribb, M.: Cloud-base distribution and cirrus properties based on micropulse lidar measurements at a site in southeastern China, *Advances in Atmospheric Sciences*, 32, 991–1004, <https://doi.org/10.1007/s00376-014-4176-2>, 2015a.
- Liu, L., Sun, X.-j., Liu, X.-c., Gao, T.-c., and Zhao, S.-j.: Comparison of cloud base height derived from a ground-based infrared cloud measurement and two ceilometers, *Advances in Meteorology*, 2015, <https://doi.org/10.1155/2015/853861>, 2015b.
- 25 Madonna, F., Rosoldi, M., Lolli, S., Amato, F., Vande Hey, J., Dhillon, R., Zheng, Y., Brettle, M., and Pappalardo, G.: Intercomparison of aerosol measurements performed with multi-wavelength Raman lidars, automatic lidars and ceilometers in the framework of INTERACT-II campaign, *Atmospheric Measurement Techniques*, 11, 2459–2475, <https://doi.org/10.5194/amt-11-2459-2018>, 2018.
- Marengo, F., Santacesaria, V., Bais, A. F., Balis, D., di Sarra, A., Papayannis, A., and Zerefos, C.: Optical properties of tropospheric aerosols determined by lidar and spectrophotometric measurement (Photochemical Activity and Solar Ultraviolet Radiation campaign), *Applied Optics*, 36, 6875–6886, <https://doi.org/10.1364/AO.36.006875>, 1997.
- 30 Martucci, G., Milroy, C., and O’Dowd, C. D.: Detection of cloud-base height using Jenoptik CHM15K and Vaisala CL31 ceilometers, *Journal of Atmospheric and Oceanic Technology*, 27, 305–318, <https://doi.org/10.1175/2009JTECHA1326.1>, 2010.
- Masunaga, H., Matsui, T., Tao, W.-k., Hou, A. Y., Kummerow, C. D., Nakajima, T., Bauer, P., Olson, W. S., Sekiguchi, M., and Nakajima, T. Y.: Satellite data simulator unit: A multisensor, multispectral satellite simulator package, *Bulletin of the American Meteorological Society*, 91, 1625–1632, <https://doi.org/10.1175/2010BAMS2809.1>, 2010.
- Mattis, I., Begbie, R., Boyouk, N., Bravo-Aranda, J. A., Brettle, M., Cermak, J., Drouin, M.-A., Geiß, A., Görsdorf, U., Haeefe, A., Haeffelin, M., Hervo, M., Komínková, K., Leinweber, R., Müller, G., Munkel, C., Pattantyús-Ábrahám, M., Pönitz, K., Wagner, F., and Wiegner,

- M.: The ceilometer inter-comparison campaign CeiLinEx2015, in: EGU General Assembly Conference Abstracts, pp. EPSC2016–9687, 2016.
- McGill, M. J., Yorks, J. E., Scott, V. S., Kupchock, A. W., and Selmer, P. A.: The Cloud-Aerosol Transport System (CATS): A technology demonstration on the International Space Station, in: Lidar Remote Sensing for Environmental Monitoring XV, vol. 9612, p. 96120A, International Society for Optics and Photonics, 2015.
- Mie, G.: Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen, *Annalen der physik*, 330, 377–445, <https://doi.org/10.1002/andp.19083300302>, 1908.
- Milroy, C., Martucci, G., Lolli, S., Loac, S., Sauvage, L., Xueref-Remy, I., Lavrič, J. V., Ciaï, P., Feist, D. G., Biavati, G., et al.: An assessment of pseudo-operational ground-based light detection and ranging sensors to determine the boundary-layer structure in the coastal atmosphere, *Advances in Meteorology*, 2012, <https://doi.org/10.1155/2012/929080>, 2012.
- Morcrette, C. J., O’Connor, E. J., and Petch, J. C.: Evaluation of two cloud parametrization schemes using ARM and Cloud-Net observations, *Quarterly Journal of the Royal Meteorological Society*, 138, 964–979, <https://doi.org/10.1002/qj.969>, 2012.
- Morille, Y., Haeffelin, M., Drobinski, P., and Pelon, J.: STRAT: An automated algorithm to retrieve the vertical structure of the atmosphere from single-channel lidar data, *Journal of Atmospheric and Oceanic Technology*, 24, 761–775, <https://doi.org/10.1175/JTECH2008.1>, 2007.
- Münkel, C., Eresmaa, N., Räsänen, J., and Karppinen, A.: Retrieval of mixing height and dust concentration with lidar ceilometer, *Boundary-layer meteorology*, 124, 117–128, <https://doi.org/10.1007/s10546-006-9103-3>, 2007.
- NASA JPL: NASA Shuttle Radar Topography Mission Global 3 arc second [Data set], NASA EOSDIS Land Processes DAAC, <https://doi.org/https://doi.org/10.5067/MEaSURES/SRTM/SRTMGL3.003>, 2013.
- Noel, V. and Chepfer, H.: A global view of horizontally oriented crystals in ice clouds from Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO), *Journal of Geophysical Research*, 115, <https://doi.org/10.1029/2009jd012365>, 2010.
- O’Connor, E. J., Illingworth, A. J., and Hogan, R. J.: A technique for autocalibration of cloud lidar, *Journal of Atmospheric and Oceanic Technology*, 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.
- Pal, S. R., Steinbrecht, W., and Carswell, A. I.: Automated method for lidar determination of cloud-base height and vertical extent, *Applied optics*, 31, 1488–1494, <https://doi.org/10.1364/AO.31.001488>, 1992.
- Parkinson, C. L.: Aqua: An Earth-observing satellite mission to examine water and other climate variables, *IEEE Transactions on Geoscience and Remote Sensing*, 41, 173–183, <https://doi.org/10.1109/TGRS.2002.808319>, 2003.
- Petty, G. W.: *A First Course in Atmospheric Radiation*, Sundog Publishing, 2 edn., 2006.
- Petty, G. W. and Huang, W.: The modified gamma size distribution applied to inhomogeneous and nonspherical particles: Key relationships and conversions, *Journal of the Atmospheric Sciences*, 68, 1460–1473, <https://doi.org/10.1175/2011JAS3645.1>, 2011.
- Powers, J. G., Monaghan, A. J., Cayette, A. M., Bromwich, D. H., Kuo, Y.-H., and Manning, K. W.: Real-Time Mesoscale Modeling Over Antarctica: The Antarctic Mesoscale Prediction System, *Bulletin of the American Meteorological Society*, 84, 1533–1546, <https://doi.org/10.1175/BAMS-84-11-1533>, 2003.
- Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., Conseil, S., Shupe, D. L., Craig, M. W., Dencheva, N., and et al.: The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package, *The Astronomical Journal*, 156, 123, <https://doi.org/10.3847/1538-3881/aabc4f>, 2018.
- R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/>, 2017.

- Rausch, J., Meyer, K., Bennartz, R., and Platnick, S.: Differences in liquid cloud droplet effective radius and number concentration estimates between MODIS collections 5.1 and 6 over global oceans, *Atmospheric Measurement Techniques*, 10, 2105–2116, <https://doi.org/10.5194/amt-10-2105-2017>, 2017.
- Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *Journal of Geophysical Research: Atmospheres*, 108, 4407, <https://doi.org/10.1029/2002JD002670>, 2003.
- Rew, R. and Davis, G.: NetCDF: an interface for scientific data access, *IEEE computer graphics and applications*, 10, 76–82, <https://doi.org/10.1109/38.56302>, 1990.
- Rosoldi, M., Madonna, F., Pappalardo, G., Hey, J. V., and Zheng, Y.: The lesson learnt during interact-I and INTERACT-II actris measurement campaigns, in: *EPJ Web of Conferences*, vol. 176, p. 11002, EDP Sciences, 2018.
- Rossow, W. B. and Schiffer, R. A.: ISCCP cloud data products, *Bulletin of the American Meteorological Society*, [https://doi.org/10.1175/1520-0477\(1991\)072<0002:ICDP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1991)072<0002:ICDP>2.0.CO;2), 1991.
- Rossum, G.: *Python reference manual*, 1995.
- Schmidt, G. A., Bader, D., Donner, L. J., Elsaesser, G. S., Golaz, J.-C., Hannay, C., Molod, A., Neale, R., and Saha, S.: Practice and philosophy of climate model tuning across six US modeling centers, *Geoscientific Model Development*, 10, 3207–3223, <https://doi.org/10.5194/gmd-10-3207-2017>, 2017.
- Silber, I., Verlinde, J., Eloranta, E. W., Flynn, C. J., and Flynn, D. M.: Polar liquid cloud base detection algorithms for high spectral resolution or micropulse lidar data, *Journal of Geophysical Research: Atmospheres*, 123, 4310–4322, <https://doi.org/10.1029/2017JD027840>, 2018.
- Spinhirne, J. D.: Micro pulse lidar, *IEEE Transactions on Geoscience and Remote Sensing*, 31, 48–55, 1993.
- Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., Illingworth, A. J., O’Connor, E. J., Rossow, W. B., Durden, S. L., et al.: The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation, *Bulletin of the American Meteorological Society*, 83, 1771–1790, <https://doi.org/10.1175/BAMS-83-12-1771>, 2002.
- Stokes, G. M. and Schwartz, S. E.: The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the cloud and radiation test bed, *Bulletin of the American Meteorological Society*, 75, 1201–1222, [https://doi.org/10.1175/1520-0477\(1994\)075<1201:TARMPP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1994)075<1201:TARMPP>2.0.CO;2), 1994.
- Swales, D. J., Pincus, R., and Bodas-Salcedo, A.: The Cloud Feedback Model Intercomparison Project Observational Simulator Package: Version 2, *Geoscientific Model Development*, 11, 77–81, <https://doi.org/10.5194/gmd-11-77-2018>, 2018.
- Telford, P. J., Braesicke, P., Morgenstern, O., and Pyle, J. A.: Technical Note: Description and assessment of a nudged version of the new dynamics Unified Model, *Atmospheric Chemistry and Physics*, 8, 1701–1712, <https://doi.org/10.5194/acp-8-1701-2008>, 2008.
- Torvalds, L.: *Linux: a portable operating system*, Master’s thesis, University of Helsinki, 1997.
- Tsaknakis, G., Papayannis, A., Kokkalis, P., Amiridis, V., Kambezidis, H., Mamouri, R., Georgoussis, G., and Avdikos, G.: Inter-comparison of lidar and ceilometer retrievals for aerosol and Planetary Boundary Layer profiling over Athens, Greece, *Atmospheric Measurement Techniques*, 4, 1261–1273, <https://doi.org/10.5194/amt-4-1261-2011>, 2011.
- Van Der Walt, S., Colbert, S. C., and Varoquaux, G.: The NumPy array: a structure for efficient numerical computation, *Computing in Science & Engineering*, 13, 22–30, <https://doi.org/10.1109/MCSE.2011.37>, 2011.
- van Diedenhoven, B.: Remote Sensing of Crystal Shapes in Ice Clouds, in: *Springer Series in Light Scattering*, pp. 197–250, Springer International Publishing, [https://doi.org/10.1007/978-3-319-70808-9\\_5](https://doi.org/10.1007/978-3-319-70808-9_5), 2017.

- Van Tricht, K., Gorodetskaya, I., Lhermitte, S., Turner, D., Schween, J., and Van Lipzig, N.: An improved algorithm for polar cloud-base detection by ceilometer over the ice sheets, *Atmospheric Measurement Techniques*, 7, 1153–1167, <https://doi.org/10.5194/amt-7-1153-2014>, 2014.
- Vaughan, M. A., Liu, Z., McGill, M. J., Hu, Y., and Obland, M. D.: On the spectral dependence of backscatter from cirrus clouds: Assessing CALIOP's 1064 nm calibration assumptions using cloud physics lidar measurements, *Journal of Geophysical Research: Atmospheres*, 115, <https://doi.org/10.1029/2009JD013086>, 2010.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Jarrod Millman, K., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C., Polat, İ., Feng, Y., Moore, E. W., Vand erPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors: SciPy 1.0—Fundamental Algorithms for Scientific Computing in Python, arXiv e-prints, arXiv:1907.10121, 2019.
- Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whitall, M., Williams, K., and Zerroukat, M.: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, *Geoscientific Model Development*, 12, 1909–1963, <https://doi.org/10.5194/gmd-12-1909-2019>, 2019.
- Wang, Z. and Sassen, K.: Cloud type and macrophysical property retrieval using multiple remote sensors, *Journal of Applied Meteorology*, 40, 1665–1682, [https://doi.org/10.1175/1520-0450\(2001\)040<1665:CTAMPR>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<1665:CTAMPR>2.0.CO;2), 2001.
- Warren, E., Charlton-Perez, C., Kotthaus, S., Lean, H., Ballard, S., Hopkin, E., and Grimmond, S.: Evaluation of forward-modelled attenuated backscatter using an urban ceilometer network in London under clear-sky conditions, *Atmospheric Environment*, 191, 532 – 547, <https://doi.org/https://doi.org/10.1016/j.atmosenv.2018.04.045>, 2018.
- Watson-Parris, D., Schutgens, N., Cook, N., Kipling, Z., Kershaw, P., Gryspeerdt, E., Lawrence, B., and Stier, P.: Community Intercomparison Suite (CIS) v1. 4.0: a tool for intercomparing models and observations, *Geoscientific Model Development*, 9, 3093–3110, <https://doi.org/10.5194/gmd-9-3093-2016>, 2016.
- Webb, M., Senior, C., Bony, S., and Morcrette, J.-J.: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models, *Climate Dynamics*, 17, 905–922, <https://doi.org/10.1007/s003820100157>, 2001.
- Webb, M. J., Andrews, T., Bodas-Salcedo, A., Bony, S., Bretherton, C. S., Chadwick, R., Chepfer, H., Douville, H., Good, P., Kay, J. E., et al.: The cloud feedback model Intercomparison Project (CFMIP) contribution to CMIP6, *Geoscientific Model Development*, 10, 359–384, <https://doi.org/10.5194/gmd-10-359-2017>, 2017.
- Welton, E. J., Voss, K. J., Gordon, H. R., Maring, H., Smirnov, A., Holben, B., Schmid, B., Livingston, J. M., Russell, P. B., Durkee, P. A., et al.: Ground-based lidar measurements of aerosols during ACE-2: Instrument description, results, and comparisons with other ground-based and airborne measurements, *Tellus B*, 52, 636–651, <https://doi.org/10.1034/j.1600-0889.2000.00025.x>, 2000.
- Welton, E. J., Voss, K. J., Quinn, P. K., Flatau, P. J., Markowicz, K., Campbell, J. R., Spinhirne, J. D., Gordon, H. R., and Johnson, J. E.: Measurements of aerosol vertical profiles and optical properties during INDOEX 1999 using micropulse lidars, *Journal of Geophysical Research: Atmospheres*, 107, INX2–18, <https://doi.org/10.1029/2000JD000038>, 2002.
- Welton, E. J., Campbell, J. R., Berkoff, T. A., Valencia, S., Spinhirne, J. D., Holben, B., Tsay, S.-C., and Schmid, B.: The NASA Micro-Pulse Lidar Network (MPLNET): an overview and recent results, *Opt. Pur. Apl*, 39, 67–74, 2006.

- Werner, M.: Shuttle radar topography mission (SRTM) mission overview, *Frequenz*, 55, 75–79, <https://doi.org/10.1515/FREQ.2001.55.3-4.75>, 2001.
- Wiegner, M. and Gasteiger, J.: Correction of water vapor absorption for aerosol remote sensing with ceilometers, *Atmospheric Measurement Techniques*, pp. 3971–3984, <https://doi.org/10.5194/amt-8-3971-2015>, 2015.
- 5 Wiegner, M. and Geiß, A.: Aerosol profiling with the Jenoptik ceilometer CHM15kx, *Atmos. Meas. Tech.*, 5, 1953–1964, <https://doi.org/10.5194/amt-5-1953-2012>, 2012.
- Wiegner, M., Madonna, F., Biniotoglou, I., Forkel, R., Gasteiger, J., Geiß, A., Pappalardo, G., Schäfer, K., and Thomas, W.: What is the benefit of ceilometers for aerosol remote sensing? An answer from EARLINET, *Atmospheric Measurement Techniques*, 7, 1979–1997, <https://doi.org/10.5194/amt-7-1979-2014>, 2014.
- 10 Wiegner, M., Mattis, I., Pattantyús-Ábrahám, M., Bravo-Aranda, J. A., Poltera, Y., Haefele, A., Hervo, M., Görsdorf, U., Leinweber, R., Gasteiger, J., Haefelin, M., Wagner, F., Cermak, J., Komínková, K., Brettle, M., Munkel, C., and Pönitz, K.: Aerosol backscatter profiles from ceilometers: validation of water vapor correction in the framework of CeiLinEx2015, *Atmospheric Measurement Techniques*, 12, 471–490, <https://doi.org/10.5194/amt-12-471-2019>, 2019.
- Williams, D. N., Ananthakrishnan, R., Bernholdt, D., Bharathi, S., Brown, D., Chen, M., Chervenak, A., Cinquini, L., Drach, R., Foster, I.,  
15 et al.: The Earth System Grid: Enabling access to multimodel climate simulation data, *Bulletin of the American Meteorological Society*, 90, 195–206, <https://doi.org/10.1175/2008BAMS2459.1>, 2009.
- Williams, J., Morgenstern, O., Varma, V., Behrens, E., Hayek, W., Oliver, H., Dean, S., Mullan, B., and Frame, D.: Development of the New Zealand Earth System Model: NZESM, *Weather and Climate*, 36, 25–44, <https://doi.org/10.2307/26779386>, 2016.
- Williams, K. D. and Bodas-Salcedo, A.: A multi-diagnostic approach to cloud evaluation, *Geoscientific Model Development*, 10, 2547,  
20 <https://doi.org/10.5194/gmd-10-2547-2017>, 2017.
- Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO mission and CALIOP data processing algorithms, *Journal of Atmospheric and Oceanic Technology*, 26, 2310–2323, <https://doi.org/10.1175/2009JTECHA1281.1>, 2009.
- Wiscombe, W. J.: Mie scattering calculations: Advances in technique and fast, vector-speed computer codes, Tech. rep., National Center for  
25 Atmospheric Research Boulder, Colorado, 1979.
- Wiscombe, W. J.: Improved Mie scattering algorithms, *Applied optics*, 19, 1505–1509, 1980.
- Yang, P., Liou, K.-N., Bi, L., Liu, C., Yi, B., and Baum, B. A.: On the radiative properties of ice clouds: Light scattering, remote sensing, and radiation parameterization, *Advances in Atmospheric Sciences*, 32, 32–63, <https://doi.org/10.1007/s00376-014-0011-z>, 2014.
- Yorks, J. E., Hlavka, D. L., Hart, W. D., and McGill, M. J.: Statistics of Cloud Optical Properties from Airborne Lidar Measurements, *Journal*  
30 *of Atmospheric and Oceanic Technology*, 28, 869–883, <https://doi.org/10.1175/2011jtecha1507.1>, 2011.
- Zadra, A., Williams, K., Frassoni, A., Rixen, M., Adames, Á. F., Berner, J., Bouysse, F., Casati, B., Christensen, H., Ek, M. B., et al.: Systematic Errors in Weather and Climate Models: Nature, Origins, and Ways Forward, *Bulletin of the American Meteorological Society*, 99, ES67–ES70, <https://doi.org/10.1175/BAMS-D-17-0287.1>, 2018.
- Zdunkowski, W., Trautmann, T., and Bott, A.: Radiation in the atmosphere: a course in theoretical meteorology, Cambridge University Press,  
35 2007.
- Zhang, Y., Xie, S., Klein, S. A., Marchand, R., Kollias, P., Clothiaux, E. E., Lin, W., Johnson, K., Swales, D., Bodas-Salcedo, A., et al.: The ARM Cloud Radar Simulator for Global Climate Models: Bridging Field Data and Climate Models, *Bulletin of the American Meteorological Society*, 99, 21–26, <https://doi.org/10.1175/BAMS-D-16-0258.1>, 2018.

Zhang, Z. and Platnick, S.: An assessment of differences between cloud effective particle radius retrievals for marine water clouds from three MODIS spectral bands, *Journal of Geophysical Research: Atmospheres*, 116, <https://doi.org/10.1029/2011JD016216>, 2011.



# Ground-based lidar processing and simulator framework for comparing models and observations (ALCF 1.0)

Peter Kuma<sup>1</sup>, Adrian J. McDonald<sup>1</sup>, Olaf Morgenstern<sup>2</sup>, Richard Querel<sup>3</sup>, Israel Silber<sup>4</sup>, and Connor J. Flynn<sup>5</sup>

<sup>1</sup>University of Canterbury, Christchurch, New Zealand

<sup>2</sup>National Institute of Water & Atmospheric Research (NIWA), Wellington, New Zealand

<sup>3</sup>National Institute of Water & Atmospheric Research (NIWA), Lauder, New Zealand

<sup>4</sup>Department of Meteorology and Atmospheric Science, Pennsylvania State University, PA, USA

<sup>5</sup>School of Meteorology, University of Oklahoma, Norman, OK, USA

**Correspondence:** Peter Kuma (peter@peterkuma.net)

**Abstract.** Automatic lidars and ceilometers provide valuable information on cloud and aerosols, but have not been used systematically in the evaluation of GCMs and NWP models. Obstacles associated with the diversity of instruments, a lack of standardisation of data products and open processing tools mean that the value of the large ALC networks worldwide is not being realised. We discuss a tool, called the Automatic Lidar and Ceilometer Framework (ALCF), that overcomes these problems and also includes a ground-based lidar simulator, which calculates the radiative transfer of laser radiation, and allows one-to-one comparison with models. Our ground-based lidar simulator is based on the Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (COSP) which has been used extensively for spaceborne lidar intercomparisons. The ALCF implements all steps needed to transform and calibrate raw ALC data and create simulated ~~backscatter-attenuated~~ volume backscattering coefficient profiles for one-to-one comparison and complete statistical analysis of cloud. The framework supports multiple common commercial ALCs (Vaisala CL31, CL51, Lufft CHM 15k and ~~Sigma-Space-Droplet Measurement Technologies~~ MiniMPL), reanalyses (JRA-55, ERA5 and MERRA-2) and models (AMPS and the Unified Model). To demonstrate its capabilities, we present case studies evaluating cloud in the supported reanalyses and models using CL31, CL51, CHM 15k and MiniMPL observations at three sites in New Zealand. We show that the reanalyses and models generally underestimate cloud fraction ~~and overestimate cloud albedo, the common "too few too bright" problem~~. If sufficiently high temporal resolution model output is available (better than 6 hourly), a direct comparison of individual clouds is also possible. We demonstrate that the ALCF can be used as a generic evaluation tool to examine cloud occurrence and cloud properties in reanalyses, NWP models and GCMs, potentially utilising the large amounts of ALC data already available. This tool is likely to be particularly useful for the analysis and improvement of low-level cloud simulations which are not well monitored from space. This has previously been identified as a critical deficiency in contemporary models, limiting the accuracy of weather forecasts and future climate projections. While the current focus of the framework is on clouds, support for aerosol in the lidar simulator is planned in the future.

## 1 Introduction

Automatic lidars and ceilometers (ALC) are active ground-based instruments which emit laser pulses in the ultraviolet, visible or infrared (IR) part of the electromagnetic spectrum and measure radiation backscattered from atmospheric constituents such as cloud and fog liquid droplets and ice crystals, haze, aerosol (~~particulate matter, ash~~) and atmospheric gases (Emeis, 5 2010). Vertical profiles of attenuated backscattered radiation can be produced by measuring received power as a function of time elapsed between emitting the pulse and receiving the backscattered radiation. Quantities such as cloud base height (CBH) and a cloud mask (Pal et al., 1992; Wang and Sassen, 2001; Martucci et al., 2010; Costa-Surós et al., 2013; Van Tricht et al., 2014; Liu et al., 2015a, b; Lewis et al., 2016; Cromwell and Flynn, 2018; Silber et al., 2018), ~~aerosol optical depth~~ particle volume backscattering coefficient (Marenco et al., 1997; Welton et al., 2000, 2002; Wiegner and Geiß, 2012; Wiegner et al., 10 2014; Jin et al., 2015; Dionisi et al., 2018) and boundary layer height (Eresmaa et al., 2006; Munkel et al., 2007; Emeis et al., 2009; Tsaknakis et al., 2011; Milroy et al., 2012; Knepp et al., 2017) can be derived from the ~~backscatter~~ attenuated volume backscattering coefficient profile. Lidars equipped with polarisation or multiple wavelengths can also provide depolarisation ratio or colour ratio, respectively, which can be used to infer cloud phase or particle types. Doppler lidars can measure wind speed in the direction of the lidar orientation. ALCs are commonly deployed at airports, where they provide 15 CBH, fog and aerosol observations needed for air traffic control. Large networks of up to ~~thousands of ACLs~~ hundreds of lidars and ceilometers have been deployed worldwide: Cloudnet (Illingworth et al., 2007), E-PROFILE (Illingworth et al., 2018), ~~EARLINET (Pappalardo et al., 2014)~~ PollyNET (Baars et al., 2016), ICENET (Cazorla et al., 2017), MPLNET (Welton et al., 2006) and ARM (Stokes and Schwartz, 1994; Campbell et al., 2002). The purpose of these networks is to observe cloud, fog, aerosol, air quality, visibility and volcanic ash, provide input to numerical weather prediction (NWP) model evaluation 20 (Hogan et al., 2001; Illingworth et al., 2007; Morcrette et al., 2012; Warren et al., 2018; Lamer et al., 2018; Hansen et al., 2018b) and assimilation (Illingworth et al., 2015b, 2018) and for climate studies. These networks are usually composed of multiple types of ~~ACLs~~ ALCs, with Vaisala CL31, CL51, Lufft (formerly Jenoptik) CHM 15k and ~~Sigma Space Droplet Measurement Technologies (formerly Sigma Space and Hexagon)~~ MiniMPL the most common. Complex lidar data processing has been set up on some of these networks. Notably, at the SIRTa site in France, lidar ratio (LR) comparable with a lidar simulator (Chiri- 25 aco et al., 2018) is calculated as part of their "ReOBS" processing method. Intercomparison and calibration campaigns such as CeiLinEx2015 (Mattis et al., 2016) and INTERACT-I(-II) (Rosoldi et al., 2018; Madonna et al., 2018) have been performed. Lidar data processing involves a number of tasks such as resampling, calibration, noise removal and cloud detection. Some of these are implemented in the instrument firmware of ALCs. This, however, means that lidar ~~backscatter~~ attenuated volume backscattering coefficient and detected cloud and cloud base are not comparable between different instruments. In most cases 30 the algorithms are not publicly documented, making it impossible to compare the data with values from a model or a lidar simulator without a systematic bias.

Atmospheric model evaluation is an ongoing task, and a critical part of the model improvement process (Eyring et al., 2019; Hourdin et al., 2017; Schmidt et al., 2017). Traditionally, various types of observational and model datasets have been utilised – weather and climate station data, upper air soundings, ground-based and satellite remote sensing datasets, and high-resolution

model simulations, amongst others. Clouds are one of the most problematic phenomena in atmospheric models due to their transient nature, high spatial and temporal variability, and sensitivity to a complex combination of conditions such as relative humidity, aerosols (presence of cloud condensation nuclei and ice nuclei), thermodynamic and dynamic conditions. At the same time, clouds have a very substantial effect on the atmospheric shortwave and longwave radiation balance, and any cloud misrepresentation has a strong effect on other components of the model, limiting the ability to accurately represent past and present climate and predict future climate (Zadra et al., 2018). An improved understanding of clouds and cloud feedbacks is one of the focuses of the Coupled Model Intercomparison Project Phase 6 (CMIP6) (Eyring et al., 2016), and comparison of model cloud with observations is one of the key points of the Cloud Feedback Model Intercomparison Project (CFMIP) (Webb et al., 2017). Satellite observations make up the majority of the data used to evaluate model clouds. These include passive visible and IR low earth orbit and geostationary radiometers measuring, among others, features such as cloud cover, cloud top height (CTH), cloud top temperature; passive microwave instruments measuring total column water; active radars and lidars measuring cloud vertical profiles. Ground-based remote sensing instruments include radars, lidars, ceilometers, radiometers and sky cameras. As pointed out by Williams and Bodas-Salcedo (2017), using a wide range of different observational datasets including satellite and ground-based for general circulation model (GCM) evaluation is important due to limitations of each dataset.

Model cloud is commonly represented by the mixing ratio of liquid and ice and cloud fraction (CF) on every model grid cell and vertical level. In addition some models provide the cloud droplet effective radius used in radiative transfer calculations. Remote sensing observations do not match the representation of the atmospheric model fields directly because of their different resolutions, limited field of view (FOV) and attenuation by atmospheric constituents before reaching the instrument's receiver. Instrument simulators bridge this gap by converting the model fields to quantities which emulate those measured by the instrument, which can then be compared directly with observations. One such collection of instrument simulators is the CFMIP Observation Simulator Package (COSP) (Bodas-Salcedo et al., 2011; Swales et al., 2018), which has been used for more than a decade for evaluation of models using satellite, and more recently ground-based, observations. The simulators in COSP include active instruments such as spaceborne and ground-based radars: Cloud Profiling Radar (CPR) on CloudSat (Stephens et al., 2002), Ka-band ARM Zenith Radar (KAZR); lidars: Cloud-Aerosol Lidar Orthogonal Polarization (CALIOP) on CALIPSO (Winker et al., 2009), Cloud-Aerosol Transport System (CATS) on ISS (McGill et al., 2015), the Atmospheric Lidar (ATLID) on EarthCARE (Illingworth et al., 2015a); and spaceborne passive instruments: ISCCP (Rossow and Schiffer, 1991), MODIS (Parkinson, 2003) and MISR (Diner et al., 1998). The more recent addition of ground-based radar (Zhang et al., 2018) and lidar (Chiriaco et al., 2018; Bastin et al., 2018) opens up new possibilities to use the large amount of remote sensing data obtained from ground-based active remote sensing instruments. In practice, ground-based observational remote sensing data are not straightforward to use without a substantial amount of additional processing. Some previous studies have also compared models and ground-based radar and lidar observations without the use of an instrument simulator (Bouniol et al., 2010; Hansen et al., 2018a), though for the reasons identified above this is not advisable.

In this study we introduce a software package called the Automatic Lidar and Ceilometer Framework (ALCF) for evaluating model cloud using ALC observations. It extends and integrates the COSP lidar simulator (Chiriaco et al., 2006; Chepfer et al.,

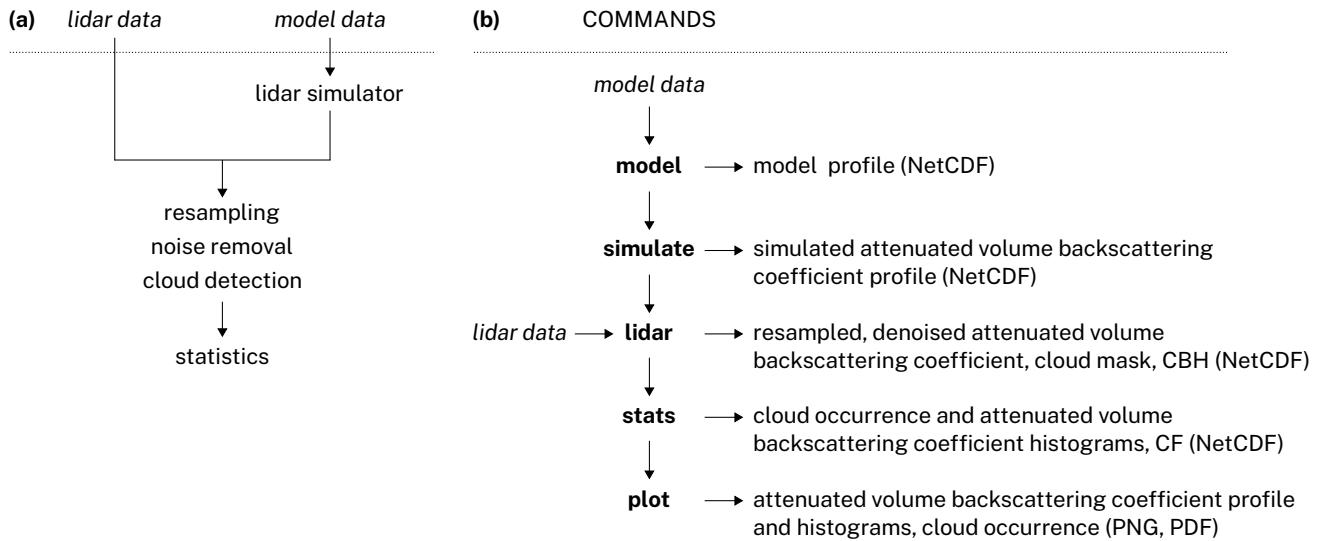
2007, 2008) with pre- and post-processing steps, and allows the simulator to be run offline on model output, instead of having to be integrated inside the model. This makes it possible to compare ALC data at any location without having to run the model with a specific configuration. Multiple ALCs, reanalyses and model output formats are supported. The original COSP lidar simulator was extended with Rayleigh and Mie, Mie and ice crystal scattering at multiple lidar wavelengths. Observational ALC data from a number of common instruments can be processed by re-sampling to a common resolution, removing noise, detecting cloud and calculating statistics. The same steps can be performed on the simulated lidar data from the model (the output of running COSP on the model data), allowing for one-to-one comparison of model and observations. A particular focus of our work was on applying the same processing steps on the observed and simulated ~~backscatter~~ attenuated volume backscattering coefficient in order to avoid biases. The ALCF is made available under an open source license (MIT) at <https://alcf-lidar.github.io>, and as a permanent archive of code and technical documentation on Zenodo at <https://doi.org/10.5281/zenodo.4088217>.

A relatively small amount of other open source code is available for ALC data processing. A lidar simulator has been developed as part of the Goddard Satellite Data Simulator Unit (G-SDSU) (G-S, 2019), a package based on the instrument simulator package SDSU (Masunaga et al., 2010). The Community Intercomparison Suite (CIS) (Watson-Parris et al., 2016) allows for subsetting, aggregation, co-location and plotting of mostly satellite data with a focus on model–observations intercomparison. The STRAT lidar data processing tools are a collection of tools for conversion of raw ALC data, visualisation and feature classification (Morille et al., 2007).

Here, we provide an overview of the ALCF (Sect. 2) ~~and~~, describe the supported ALCs, reanalyses and models (Sect. 3). ~~We~~, the lidar simulator (Sect. 4) and the observed and simulated lidar data processing steps (Sect. 5). Later, we present a set of case studies at three sites in New Zealand (NZ) (Sect. 6) to demonstrate the value of this new tool. ~~Furthermore, we describe~~ the lidar simulator (Sect. 4), observed and simulated lidar data processing steps (Sect. 5) and ~~Lastly, we present~~ the results of the case studies (in Sect. 7).

## 2 Overview of operation of the Automatic Lidar and Ceilometer Framework (ALCF 1.0)

The ALCF performs the necessary steps to simulate ALC ~~backscatter~~ attenuated volume backscattering coefficient based on 4-dimensional atmospheric fields from reanalyses, NWP models and GCMs, and to transform the observed raw ALC ~~backscatter~~ attenuated volume backscattering coefficient profiles to profiles comparable with the simulated profiles. It does so by extracting 2-dimensional (time × height) profiles from the model data, performing radiative transfer calculations based on a modified COSP lidar simulator (Sect. 4), absolute calibration and resampling of ~~observed backscatter~~ the observed attenuated volume backscattering coefficient to common resolution and performing comparable cloud detection on the simulated and observed ~~backscatter~~ attenuated volume backscattering coefficient. The framework supports multiple common ALCs (Sect. 3.1), reanalyses and models (Sect. 3.2). The schematic in Fig. 1 illustrates this process as well as the ALCF commands which perform the individual steps. The following commands are implemented: **model**, **simulate**, **lidar**, **stats** and **plot**. The commands are normally executed in a sequence, which is also implemented by a meta-command **auto**, which is equivalent to executing a sequence of commands. The commands are described in detail in the technical documentation available online at <https://>



**Figure 1.** (a) [Scheme showing the operation of the ALCF](#) and (b) [the processing commands](#).

alcf-lidar.github.io, on Zenodo at <https://doi.org/10.5281/zenodo.4088217> and in the Supplementary information. The physical basis is described here.

The **model** command extracts 2-dimensional profiles of cloud liquid and ice content (and other thermodynamic fields) from the supported NWP model, GCM and reanalysis data (*model data* in Fig. 1) at a geographical point, along a ship track or a flight path. The resulting profiles are recorded as NetCDF files. Sect. 3.2 describes the supported reanalyses and models. The *model data* can be either in one of the supported model output formats, or a new module for reading arbitrary model output can be written, providing that the required atmospheric fields are present in the model output. The required model fields are: per-level specific cloud liquid water content, specific cloud ice water content, cloud fraction, geopotential height, temperature, surface-level pressure and orography. No physical calculations are performed by this command. The atmospheric profiles are extracted by a nearest-neighbour selection.

The **simulate** command runs the lidar simulator described in Sect. 4 on the extracted model data (the output of the **model** command) and produces simulated [backscatter-attenuated volume backscattering coefficient](#) profiles. This command runs the COSP-derived lidar simulator, which performs radiative transfer calculations of the laser radiation through the atmosphere. The resulting simulated [backscatter-attenuated volume backscattering coefficient](#) profiles are the output of this command.

The **lidar** command applies various processing algorithms on either the simulated [backscatter-data-attenuated volume backscattering coefficient](#) (the output of the **simulate** command) or the observed ALC [data-coefficient](#) (*lidar data* in Fig. 1) (Sect. 5). The data are resampled to increase the signal-to-noise ratio (SNR), noise is subtracted, LR is calculated, a cloud mask is calculated by applying a cloud detection algorithm and CBH is determined from the cloud mask. Absolute calibration (Sect. 5.2) can also be applied in this step by multiplying the observed [backscatter-attenuated volume backscattering coefficient](#) by a

**Table 1.** Table of ALCs and their technical parameters. Power is calculated as Pulse  $\times$  Pulse Repetition Frequency (PRF).

Instrument	$\lambda$ (nm)	Laser	Rate <sup>1</sup> (s)	Res. <sup>2</sup> (m)	Depol. <sup>3</sup>	Pulse <sup>4</sup> ( $\mu$ J)	Range <sup>5</sup> (km)	PRF (kHz)	Overlap <sup>6</sup> (m)	Power (mW)	FOV <sup>7</sup> ( $\mu$ rad)
CHM 15k	1064	Nd:YAG	2–600	5	no	7–9	15.4	5–7	1000 <sup>8</sup>	48	450
CL31	910	InGaAs	2–120	10	no	1.2	7.7	10	70 <sup>8</sup>	12	830
CL51	910	InGaAs	6–120	10	no	3	15.4	6.5	230 <sup>9</sup>	20	560
MiniMPL	532	Nd:YAG	1–900	5–75	yes	3–4	30.0	2.5	2000 <sup>9</sup>	9	110

<sup>1</sup>Sampling rate. <sup>2</sup>Vertical (range) resolution. <sup>3</sup>Depolarisation. <sup>4</sup>Pulse energy. <sup>5</sup>Maximum range. <sup>6</sup>Range of full overlap. <sup>7</sup>Receiver field of view. <sup>8</sup>Hopkin et al. (2019). <sup>9</sup>Madonna et al. (2018).

calibration coefficient. This is important in order to obtain unbiased [backscatter-attenuated volume backscattering coefficient](#) profiles comparable with the simulated [backscatter](#) profiles. Sect. 3.1 describes the supported instruments. The *lidar data* can be in one of the supported instrument formats. If the native instrument format is not NetCDF, it has to be converted from the native format with the auxiliary command **convert** or one of the conversion programs: `cl2nc` (Vaisala CL31, CL51), `mpl2nc` or `SigmaMPL` (Sigma Space MiniMPL).

The **stats** step calculates summary statistics from the output of the **lidar** command. These include CF, cloud occurrence by height, [backscatter-attenuated volume backscattering coefficient](#) histograms and the averages of LR and [backscatter backscattering coefficient](#).

The **plot** command plots [the backscatter-attenuated volume backscattering coefficient](#) profiles produced by the **lidar** command (Fig. 4, 5, 6), and the statistics produced by the **stats** command: cloud occurrence (Fig. 3), [backscatter-attenuated volume backscattering coefficient](#) histograms (Fig. 7) and [backscatter-attenuated volume backscattering coefficient](#) noise standard deviation histograms (Fig. 9).

### 3 Supported input data: instruments, reanalyses and models

#### 3.1 Instruments

The primary focus of the framework is to support common commercial ALCs. Ceilometers are considered the most basic type of lidar (Emeis, 2010; Kotthaus et al., 2016) intended as commercial products designed for unattended operation. They are used routinely to measure CBH, but most instruments also provide the full vertical profiles of [backscatter-attenuated volume backscattering coefficient](#). Therefore, they are suitable for model evaluation by comparing not only CBH, but also cloud occurrence as a function of height. Their compact size and low cost make it possible to deploy a large number of these instruments in different locations, or use them in unusual settings such as mounted on ships (Klekociuk et al., 2019; Kuma et al., 2020). Common off-the-shelf ceilometers are the Lufft CHM 15k, Vaisala CL31 and CL51. Some lidars offer higher power and therefore higher SNR, and capabilities not present in ceilometers such as dual polarisation, multiple wavelengths, Doppler shift measurement and Raman scattering. Below we describe ALCs supported by the framework and used in our case studies: Lufft CHM 15k, Vaisala CL31 and CL51 and [Sigma-Space Droplet Measurement Technologies](#) MiniMPL. Table 1 lists selected parameters of the supported ALCs.

*Lufft CHM 15k* (previously Jenoptik CHM 15k) is a ceilometer operating at a wavelength of 1064 nm ~~in the near-IR spectrum.~~ (near IR). The maximum range of the instrument is 15.4 km, vertical sampling resolution 5 m in the first 150 m and 15 m above and sampling rate ~~up to~~ 2 s. The total number of vertical levels is 1024. The wavelength in the near IR spectrum ensures low molecular ~~backscatter~~ backscattering. The instrument produces NetCDF files containing uncalibrated attenuated ~~backscatter~~ volume backscattering coefficient profiles and various derived variables, although the calibration coefficient is relatively consistent for different instruments of the model (Hopkin et al., 2019, Fig. 13).

*Vaisala CL31 and CL51* are ceilometers operating at a wavelength of 910 nm ~~in the near-IR spectrum~~ (near IR). The maximum range of CL31 and CL51 is 7.7 km and 15.4 km and the sampling rate is 2 and 6 s, respectively. The vertical resolution is 10 m. The total number of vertical levels is 770 and 1540, respectively. The wavelength is characterised by relatively low molecular ~~backscatter~~ backscattering (but higher than 1064 nm) and is affected by water vapour absorption (Wiegner and Gasteiger, 2015; Wiegner et al., 2019), which can cause additional absorption of about 20% in the mid-latitudes and 50% in the tropics (see also Sect. 5.4). The instruments produce data files containing uncalibrated attenuated ~~backscatter~~ volume backscattering coefficient which can be converted to NetCDF (see c12nc in the Code and data availability section). The firmware configuration option "noise\_h2 off" results in backscatter range correction to be selectively applied under a certain critical range and above this range ~~only if cloud is present (Kotthaus et al., 2016, Sect. 3.2). This was the case with our case study dataset (Sect. 6). We apply range correction on the uncorrected range gates during lidar data processing. The critical range in CL51 is not documented, but was determined as 6000 m based on an observed discontinuity.~~

*Sigma Space (part of Hexagon) Droplet Measurement Technologies Mini Micro Pulse Lidar (MiniMPL) (previously Sigma Space MiniMPL and Hexagon MiniMPL)* (Spinhirne, 1993; Campbell et al., 2002; Flynn et al., 2007) is a dual-polarisation ~~micro pulse lidar (meaning that it uses a high pulse repetition rate (PRF) and low pulse power) operating at a wavelength of 532 nm (green )colour in the visible spectrum).~~ The maximum range of the instrument is 30 km. The vertical resolution is ~~up to 5–5–75 m~~ and sampling rate ~~up to~~ 1 s. The shorter wavelength is affected by stronger molecular ~~backscatter~~ backscattering than 910 nm and 1064 nm. The instrument can be housed in ~~a protective enclosure, which makes it suitable for deployment in a relatively harsh environment. A scanning head can be mounted on the enclosure which provides an enclosure with a scanning head to provide~~ configurable scanning by elevation angle and azimuth. The instrument produces data files containing raw ~~backscatter~~ attenuated volume backscattering coefficient which can be converted to NetCDF containing normalised relative backscatter (NRB) with a vendor-provided tool SigmaMPL (see also mpl2nc in the Code and data availability section).

### 3.2 Reanalyses and models

Below we briefly describe reanalyses and models<sup>1</sup> used in the case studies presented here (Sect. 6). We used publicly available ~~output from three reanalyses and one NWP model. In addition, we performed nudged GCM simulations with high-temporal resolution output with the Unified Model (UM). Table 2 lists some of the main properties of the reanalyses and models.~~

---

<sup>1</sup>We use the term "reanalysis" when referring to ERA5, JRA-55 and MERRA-2 even though the reanalyses are based on atmospheric models. We use the term "model" when referring to AMPS and the UM, which are atmospheric models.

**Table 2.** Reanalyses and models used in the case studies and some of their main properties. The temporal and horizontal grid resolution and vertical levels listed is the resolution of the model output available. The horizontal grid resolution is determined at 45°S. The internal resolution of the model may be different (see Sect. 3.2 for details). The reanalyses and the UM use regular longitude-latitude grids, while the AMPS horizontal grid is regular in the South Pole stereographic projection.

<u>Model/Grid</u>	<u>Type</u>	<u>Time resolution</u>	<u>Horizontal grid resolution</u>	<u>Vertical levels</u>
<u>AMPS/D01</u>	<u>NWP</u>	<u>3 h</u>	<u>0.27° × 0.19° (21 × 21 km)</u>	<u>60</u>
<u>ERA5</u>	<u>Reanalysis</u>	<u>1 h</u>	<u>0.25° × 0.25° (20 × 28 km)</u>	<u>37</u>
<u>JRA-55</u>	<u>Reanalysis</u>	<u>6 h</u>	<u>1.25° × 1.25° (98 × 139 km)</u>	<u>37</u>
<u>MERRA-2</u>	<u>Reanalysis</u>	<u>3 h</u>	<u>0.625° × 0.50° (49 × 56 km)</u>	<u>72</u>
<u>UM (GA7.1)/N96</u>	<u>GCM</u>	<u>20 min.</u>	<u>1.875° × 1.25° (147 × 139 km)</u>	<u>85</u>

*The Antarctic Mesoscale Prediction System (AMPS)* (Powers et al., 2003) is a limited-area NWP model based on the polar fifth-generation Pennsylvania State University-National Center for Atmospheric Research Mesoscale Model (Polar MM5), now known as the Polar Weather Research and Forecasting (WRF) model (Hines and Bromwich, 2008). The model serves operational and scientific needs in Antarctica, but its largest grid also covers the South Island of NZ. AMPS forecasts are publicly available on the Earth System Grid (Williams et al., 2009). The forecasts are produced on several domains. The largest domain D01 used in the presented analysis covers NZ and has horizontal grid spacing of approximately 21 km over NZ. The model uses 60 vertical levels. The model output is available in 3-hourly intervals and initialised at 00:00 and 12:00 UTC. The initial and boundary conditions are based on the Global Forecasting System (GFS) global NWP model. AMPS assimilates local Antarctic observations from human-operated stations, automatic weather stations (AWS), upper-air stations and satellites.

*ERA5* (ECMWF, 2019) is a reanalysis produced by the European Centre For Medium-Range Weather Forecasts (ECMWF) currently available for the time period 1979 to present, with a plan to extend the time period to 1950. The reanalysis is based on the global NWP model Integrated Forecast System (IFS) version CY41R2. It uses a 4D-Var assimilation of station, satellite, radiosonde, radar, aircraft, ship-based and buoy data. The model has 137 vertical levels. Atmospheric fields are interpolated from horizontal resolution equivalent of 31 km and 137 model levels on regular longitude-latitude grid of 0.25° and 37 pressure levels, and made available to the end-users. In this analysis we use the hourly data on pressure and single-surface levels.

*Japanese 55-year reanalysis (JRA-55)* (Ebita et al., 2011; Kobayashi et al., 2015; Harada et al., 2016) is a global reanalysis produced by the Japan Meteorological Agency (JMA) and the Central Research Institute of Electric Power Industry (CRIEPI) based on the JMA Global Spectral Model (GSM). The reanalysis is available from 1958 onward. The reanalysis is based on the JMA operational assimilation system. JRA-55 uses a 4D-Var assimilation of surface, upper-air, satellite, ship-based and aircraft observations. The model uses 60 vertical levels and a horizontal grid with resolution approximately 60 km. In this analysis we use the 1.25° isobaric analysis and forecast fields interpolated to 37 pressure levels.

*Modern-Era Retrospective analysis for Research and Applications (MERRA-2)* (Gelaro et al., 2017) is a reanalysis produced by the NASA Global Modeling and Assimilation Office (GMAO). The reanalysis is based on the Goddard Earth Observing



System (GEOS) atmospheric model. The model has approximately  $0.5^\circ \times 0.65^\circ$  horizontal resolution and 72 vertical levels. It performs 3D-Var assimilation of station, upper-air, satellite, ship-based and aircraft data in 6-hourly cycles. In this analysis, we use the MERRA-2 3-hourly instantaneous model-level assimilated meteorological fields (M2I3NVASM) version 5.12.4 product.

- 5     *The UK Met Office Unified Model (UM)* (Walters et al., 2019) is an atmospheric model for weather forecasting and climate projection developed by the UK Met Office and the Unified Model Partnership. The UM is the atmospheric component, called Global Atmosphere (GA), of the HadGEM3–GC3.1 GCM and the UKESM1 earth system model (ESM). In this analysis we performed custom nudged runs of the UM (Telford et al., 2008) in the GA7.1 configuration with 20 min. time step and output temporal resolution on a New Zealand eScience Infrastructure (NeSI)/National Institute of Water & Atmospheric Research
- 10 (NIWA) supercomputer (Williams et al., 2016). The model was nudged to the ERA-Interim (Dee et al., 2011) atmospheric fields of horizontal wind speed and potential temperature and the HadISST sea surface temperature (SST) and sea ice dataset (Rayner et al., 2003). The model uses 85 vertical levels and a horizontal grid resolution of  $1.875^\circ \times 1.25^\circ$ .

#### 4 Lidar simulator

The COSP lidar simulator Active Remote Sensing Simulator (ACTSIM) was introduced by Chiriaco et al. (2006) for the

15 purpose of deriving simulated CALIOP measurements (Chepfer et al., 2007, 2008). The simulation is implemented by applying the lidar equation on model levels. Scattering [and absorption](#) by cloud particles and air molecules is calculated using the Mie and Rayleigh theory, respectively. [Scattering and absorption by aerosols is not implemented in the presented version, but support is planned in the future for models which provide concentration of aerosols. Therefore, the current focus of the simulator is solely on cloud evaluation.](#) CALIOP operates at a wavelength of 532 nm, and calculations in the original COSP

20 simulator use this wavelength. We implemented a small set of changes to the lidar simulator to support a number of ALCs with different operating wavelengths [and developed parametrisation of backscattering from ice crystals based on temperature.](#)

The lidar equation (Emeis, 2010) is based on the radiative transfer equation (Goody and Yung, 1995; Liou, 2002; Petty, 2006; Zdunkowski et al., 2007), which relates transmission of radiation to scattering, emission and absorption in media such as the atmosphere. The lidar equation assumes laser radiation passes through ~~a horizontally homogeneous the~~ atmosphere where

25 it is absorbed and scattered. A fraction of laser radiation is scattered back to the instrument and reaches the receiver. Scattering and absorption in the atmosphere is determined by its constituents – gases, liquid droplets, ice crystals and aerosol particles. ~~Atmospheric model output typically contains~~ [The focus of the current version of the simulator is on clouds. For this purpose, the atmospheric model output needed is](#) 4-dimensional fields of mass mixing ratios of liquid and ice and CF. The lidar equation can be applied on these output fields to simulate the backscattered radiation received by the instrument. Table 3 lists physical

30 quantities used in the following sections. Here, we use radiative transfer notation similar to Petty (2006) and the notation of the original lidar simulator (Chiriaco et al., 2006).

**Table 3.** Table of physical quantities.

<u>Symbol</u>	<u>Name</u>	<u>Units</u>	<u>Expression</u>
$\Omega$	<u>Solid angle</u>	<u>sr</u>	
$z$	<u>Height relative to the instrument</u>	<u>m</u>	
$k_B$	<u>Boltzmann constant</u>	$\text{JK}^{-1}$	$k_B \approx 1.38 \times 10^{-23} \text{JK}^{-1}$
$p$	<u>Atmospheric pressure</u>	<u>Pa</u>	
$T$	<u>Atmospheric temperature</u>	<u>K</u>	
$\rho_{\text{air}}$	<u>Air density</u>	$\text{kg.m}^{-3}$	
$\rho$	<u>Liquid (or ice) density</u>	$\text{kg.m}^{-3}$	
$q$	<u>Cloud liquid (or ice) mass mixing ratio</u>	<u>1</u>	
$N$	<u>Particle number concentration</u>	$\text{m}^{-3}$	
$\alpha_s (\alpha_e)$	<u>Volume scattering (extinction) coefficient</u>	$\text{m}^{-1}$	
$P_\pi(\theta)$	<u>Scattering phase function at angle <math>\theta</math></u>	<u>1</u>	$\int_{4\pi} P_\pi(\theta) d\Omega = 4\pi$
$\beta$	<u>Volume backscattering coefficient</u>	$\text{m}^{-1}\text{sr}^{-1}$	$\beta = \alpha_s P_\pi(\pi)/(4\pi)$
$\beta_{\text{mol}}$	<u>Volume backscattering coefficient for air molecules</u>		
$\beta_p$	<u>Volume backscattering coefficient for cloud particles</u>		
$\eta$	<u>Multiple scattering coefficient</u>	<u>1</u>	
$\beta'$	<u>Attenuated volume backscattering coefficient</u>	$\text{m}^{-1}\text{sr}^{-1}$	$\beta' = \beta \exp(-2 \int_0^z \eta \alpha_e dz)$
$S$	<u>Lidar ratio (extinction-to-backscatter ratio)</u>	<u>sr</u>	$S = \alpha_e / \beta$
$S'$	<u>Effective (apparent) lidar ratio</u>	<u>sr</u>	$S' = S\eta$
$k$	<u>Backscatter-to-extinction ratio</u>	$\text{sr}^{-1}$	$k = 1/S$
$n(r)$	<u>Number distribution of particle size</u>	$\text{m}^{-4}$	$N = \int_0^\infty n(r) dr$
$Q_s (Q_e)$	<u>Scattering (extinction) efficiency of spherical particles</u>	<u>1</u>	$\alpha_s = Q_s \pi r^2 N, \alpha_e = Q_e \pi r^2 N$
$Q_b$	<u>Backscattering efficiency of spherical particles</u>	$\text{sr}^{-1}$	$\beta = Q_b \pi r^2 N$
$r_{\text{eff}}$	<u>Effective radius</u>	<u>m</u>	$r_{\text{eff}} = \int_0^\infty r^3 n(r) dr / \int_0^\infty r^2 n(r) dr$
$\sigma_{\text{eff}}$	<u>Effective standard deviation</u>	<u>m</u>	$\sigma_{\text{eff}} = (\int_0^\infty (r - r_{\text{eff}})^2 r^2 n(r) dr) / (\int_0^\infty r^2 n(r) dr)$

Below we provide a brief review of LR, Rayleigh and Mie scattering, calculate LR of cloud droplets at lidar wavelengths of the presented instruments and introduce an empirical parametrisation of LR and multiple scattering coefficient of ice crystals based on previous studies.

#### **4.1 Lidar ratio**

- Lidar ratio  $S$  is the extinction-to-backscattering ratio of atmospheric constituents at the lidar wavelength. It is an important quantity in lidar observations and the lidar simulator because it determines the amount of attenuation and backscattering. LR is not explicitly known from the observed attenuated volume backscattering coefficient. For liquid cloud droplets at near

IR wavelengths it is relatively constant at  $S \approx 19$  sr (Sect. 4.2), while for ice crystals (Sect. 4.3) and aerosol it is highly variable. When the lidar signal is fully attenuated, and under the assumption that cloud LR is constant and scattering from clouds is much stronger than molecular and aerosol scattering, LR can be determined from the observed attenuated volume backscattering coefficient by integrating it vertically (O'Connor et al., 2004):

$$5 \quad S' = \eta S = \frac{1}{2 \int_0^\infty \beta' dz}, \quad (1)$$

where  $S'$  is effective (apparent) LR, a quantity which does not depend on the multiple scattering coefficient.

## 4.2 Rayleigh and Mie scattering

The Rayleigh volume backscattering coefficient  $\beta_{\text{mol}}$  ( $\text{m}^{-1}\text{sr}^{-1}$ ) in ACTSIM is parametrised by the following equation (Eq. (8) in Chiriaco et al. (2006)):

$$10 \quad \beta_{\text{mol}} = \frac{p}{k_B T} \frac{p}{k_B T} (5.45 \times 10^{-32}) \left( \frac{\lambda}{550\text{nm}} \right)^{-4.09} = \frac{p}{k_B T} \frac{p}{k_B T} C_{\text{mol}}, \quad (2)$$

where for lidar wavelength  $\lambda = 532$  nm,  $C_{\text{mol}} = 6.2446 \times 10^{-32}$ ;  $k_B$  is the Boltzmann constant  $k_B \approx 1.38 \times 10^{-23}$   $k_B \approx 1.38 \times 10^{-23}$   $\text{JK}^{-1}$ ,  $p$  is the atmospheric pressure and  $T$  is the atmospheric temperature. We multiply this equation by  $\exp(4.09(\log(532) - \log(\lambda)))$  (where the value of  $\lambda$  is in nm) to get molecular backscatter-backscattering for wavelengths other than 532 nm, which allows us to support multiple commercially available instruments. The strength of molecular backscattering

15 is usually lower than backscattering from clouds for the relevant wavelengths.

The lidar signal at visible or near IR wavelengths is scattered by ~~atmospheric constituents of similar size~~ cloud droplets in the Mie scattering regime (Mie, 1908). In the most simple approximation, one can assume spherical dielectric particles. The scattering from these particles depends on the relative size of the wavelength and the (spherical) particle radius  $r$ , expressed by the dimensionless size parameter  $x$ :

$$20 \quad x = \frac{2\pi r}{\lambda}. \quad (3)$$

While the wavelength is approximately constant during the operation of the lidar<sup>2</sup>, the particle size comes from a distribution of sizes, typically approximated in NWP models and GCMs by a Gamma or log-normal distribution with a given mean and standard deviation. Some models provide the mean as effective radius  $r_{\text{eff}}$ . If the effective radius is not provided by the model, the lidar simulator assumes a value  $r_{\text{eff}} = 30\text{-}\mu\text{m}$   $10\text{-}\mu\text{m}$  by default, which is ~~the value assumed by the original ACTSIM~~

<sup>2</sup>The actual lidar wavelength is not constant and is characterised by a central wavelength and width. The central wavelength may fluctuate with temperature (Wiegner and Gasteiger, 2015).

approximately consistent with global studies of effective radius (Bréon and Colzy, 2000; Bréon and Doutriaux-Boucher, 2005; Hu et al., 2005). This is different from the default effective radius of 30  $\mu\text{m}$  in the original COSP lidar simulator.

In order to support multiple laser wavelengths, it is necessary to calculate backscattering efficiency due to scattering by a distribution of particle sizes. We use the computer code MIEV developed by Warren J Wiscombe (Wiscombe, 1979, 1980) to calculate backscattering efficiency for a range of the size parameter  $x$  and integrate for a distribution of particle sizes. The resulting pre-calculated LR (extinction-to-backscatter ratio) as a function of the effective radius is included in the lidar simulator for fast lookup during the simulation.

Cloud droplet ~~and ice crystal~~ size distribution parameters are an important assumption in lidar simulation due to the dependence of Mie scattering on the ratio of wavelength and particle size (the size parameter  $x$ ). NWP models and GCMs traditionally use the effective radius  $r_{\text{eff}}$  and effective standard deviation  $\sigma_{\text{eff}}$  (or an equivalent parameter such as effective variance  $\nu_{\text{eff}}$ ) to parametrise this distribution. Knowledge of the real distribution is likely highly uncertain due to a large variety of clouds occurring globally and the limited ability to predict microphysical cloud properties in models. In this section we introduce theoretical assumptions used in the lidar simulator based on established definitions of the effective radius and effective standard deviation and two common distributions. Edwards and Slingo (1996) discuss the effective radius in the context of model radiation schemes, and we will primarily follow the definitions detailed in Chang and Li (2001) and Petty and Huang (2011). The practical result of this section (and the corresponding offline code) is pre-calculated backscatter-to-extinction ratios as a function of the effective radius in the form of a lookup table included in the lidar simulator, and used in the online calculations. The offline code is provided and can be re-used for calculation of the necessary lookup tables for different lidar wavelengths, should the user of the code want to support another instrument.

The effective radius  $r_{\text{eff}}$  and effective standard deviation  $\sigma_{\text{eff}}$  are defined by:

$$r_{\text{eff}} = \frac{\int_0^\infty r^3 n(r) dr}{\int_0^\infty r^2 n(r) dr}, \quad \sigma_{\text{eff}}^2 = \frac{\int_0^\infty (r - r_{\text{eff}})^2 r^2 n(r) dr}{\int_0^\infty r^2 n(r) dr}, \quad (4)$$

where  $n(r)$  is the probability density function (PDF) of the distribution. Here, we follow Petty and Huang (2011), who define the effective variance  $\nu_{\text{eff}}$  which relates to  $\sigma_{\text{eff}}$  by  $\nu_{\text{eff}} = \sigma_{\text{eff}}^2 / r_{\text{eff}}^2$ . Due to lack of knowledge about the real distribution of particle radii, it has to be modelled by a theoretical distribution, such as a log-normal or Gamma distribution. The original ACTSIM simulator assumes a log-normal distribution (Chiriaco et al., 2006) with the PDF:

$$n(r) \propto \frac{1}{r} \exp\left(-\frac{(\log r - \mu)^2}{2\sigma^2}\right), \quad (5)$$

where  $\mu$  and  $\sigma$  are the mean and the standard deviation of the corresponding normal distribution, respectively. Chiriaco et al. (2006) use the value of  $\sigma = \log(1.2) = 0.18$  "for ice clouds" (the value for liquid cloud does not appear to be documented). In our parametrisation we used a combination of  $r_{\text{eff}}$  and  $\sigma_{\text{eff}}$  to constrain the theoretical distribution, where the effective standard deviation  $\sigma_{\text{eff}}$  was assumed to be one fourth of the effective radius  $r_{\text{eff}}$ . This choice is approximately consistent with  $\sigma = \log(1.2) = 0.18$  at  $r_{\text{eff}} = 20 \mu\text{m}$  (see Table 4, described below). In future updates, the values could be based on in situ studies of size distribution or taken from the atmospheric model output if available.

**Table 4.** Table of sensitivity tests of theoretical distribution assumption, effective radius  $r_{\text{eff}}$  and effective standard deviation  $\sigma_{\text{eff}}$  of the cloud droplet size distribution.  $\mu$  and  $\sigma$  are the mean and standard deviation of a normal distribution corresponding to the log-normal distribution, calculated numerically from  $r_{\text{eff}}$  and  $\sigma_{\text{eff}}$ .  $\mu_*$  and  $\sigma_*$  are the actual mean and standard deviation of the distribution (calculated numerically).

Distribution	$r_{\text{eff}} (\mu\text{m})$	$\sigma_{\text{eff}} (\mu\text{m})$	$\mu$	$\sigma$	$\mu_*$ ( $\mu\text{m}$ )	$\sigma_*$ ( $\mu\text{m}$ )
log-normal	20	10	2.44	0.47	12.76	6.26
log-normal	20	5	2.84	0.25	17.72	4.43
log-normal	10	5	1.74	0.47	6.40	3.20
Gamma	20	10			9.98	7.00
Gamma	20	5			17.50	4.68
Gamma	10	5			5.00	3.54

From the expression for the n-th moment of the log-normal distribution  $E[X^n] = \exp(n\mu + n^2\frac{\sigma^2}{2})$  and Equation (4) we calculate  $r_{\text{eff}}$  and  $\sigma_{\text{eff}}$  of the log-normal distribution:

$$r_{\text{eff}} = \frac{E[r^3]}{E[r^2]} = \exp\left(\mu + \frac{5}{2}\sigma^2\right), \quad (6)$$

$$\sigma_{\text{eff}}^2 = \frac{E[(r - r_{\text{eff}})^2 r^2]}{E[r^2]} = \frac{E[r^4] - 2E[r^3]r_{\text{eff}} + r_{\text{eff}}^2 E[r^2]}{E[r^2]} = \frac{\exp(4\mu + 8\sigma^2) - \exp(4\mu + 7\sigma^2)}{\exp(2\mu + 2\sigma^2)} =$$

$$= \exp(2\mu + 6\sigma^2) - \exp(2\mu + 5\sigma^2). \quad (7)$$

We find  $\mu$  and  $\sigma$  for given  $r_{\text{eff}}$  and  $\sigma_{\text{eff}}$  numerically by root-finding using the equations above. In practice, we find that the root-finding converges well for  $r_{\text{eff}}$  between 5 and 50  $\mu\text{m}$ , which is the range **mostly-most** likely to be applicable in practice.

The Gamma distribution follows the PDF:

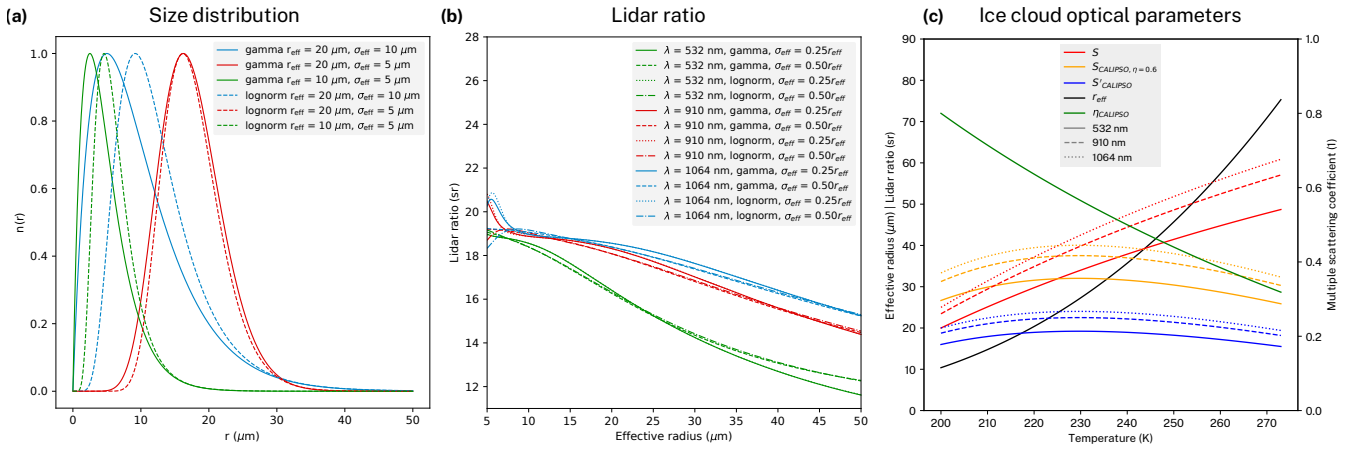
$$n(r) \propto r^{(1-3\nu_{\text{eff}})/\nu_{\text{eff}}} \exp\left(-\frac{r}{r_{\text{eff}}\nu_{\text{eff}}}\right) \quad (8)$$

(see e.g. Eq. 13 in Petty and Huang (2011) or Eq. 1 in Bréon and Doutriaux-Boucher (2005)). In this case, the distribution depends explicitly on  $r_{\text{eff}}$  and  $\sigma_{\text{eff}}$ , and as such does not require numerical root-finding.

Figure 2a shows the log-normal and Gamma distributions calculated for a number of  $r_{\text{eff}}$  and  $\sigma_{\text{eff}}$  values, and Table 4 summarises properties of these distributions. The actual mean and standard deviation of the distributions do **no-not** necessarily correspond well with the effective radius and effective standard deviation.

In ACTSIM, the volume extinction coefficient  $\alpha_e$  is calculated by integrating the extinction by individual particles over the particle size distribution:

$$\alpha_e = \int_0^\infty Q_e \pi r^2 n(r) dr \approx Q_e \pi \int_0^\infty r^2 n(r) dr = Q_e \frac{3q\rho_{\text{air}}}{4\rho r_{\text{eff}}}, \quad (9)$$



**Figure 2.** (a) Theoretical distributions of cloud droplet radius based on the log-normal and Gamma distributions parametrised by multiple choices of the effective radius  $r_{\text{eff}}$  and effective standard deviation  $\sigma_{\text{eff}}$ . (b) Lidar ratio (LR) as a function of effective radius calculated for different theoretical cloud droplet size distributions, laser wavelengths and effective standard deviation ratios. (c) Parametrisation of ice cloud optical properties as a function of temperature based on Garnier et al. (2015) and Heymsfield (2005). The plot shows LR ( $S$ ), LR of CALIPSO calculated using the constant standard processing multiple scattering coefficient  $\eta = 0.6$  ( $S_{\text{CALIPSO}, \eta=0.6}$ ), the effective LR of CALIPSO ( $S'_{\text{CALIPSO}}$ ), the effective radius ( $r_{\text{eff}}$ ), the multiple scattering coefficient of CALIPSO ( $\eta_{\text{CALIPSO}}$ ) determined by Garnier et al. (2015). LRs are calculated for three wavelengths of 532 nm (solid line), 910 nm (dashed line) and 1064 nm (dotted line) by scaling with colour ratio.

assuming approximately constant extinction efficiency  $Q_e \approx 2$  (which is approximately true for the interesting range of  $r_{\text{eff}}$  and laser wavelengths), and using the relationship between the cloud liquid (or ice) mass mixing ratio  $q$  and  $\int_0^\infty r^2 n(r) dr$ :

$$q\rho_{\text{air}} = \int_0^\infty \frac{4}{3}\pi r^3 \rho n(r) dr = \frac{4}{3}\pi\rho \int_0^\infty r^3 n(r) dr = \frac{4}{3}\pi\rho r_{\text{eff}} \int_0^\infty r^2 n(r) dr, \quad (10)$$

where  $\rho$  and  $\rho_{\text{air}}$  are the densities of liquid water (or ice) and air, respectively.

- 5 Likewise, the volume backscatter backscattering coefficient from particles  $\beta_p - \beta_D$  is calculated by integrating backscattering by individual particles over the particle size distribution:

$$\beta_{\underline{p}} = \int_0^\infty Q_s \pi r^2 \frac{P_\pi(\pi)}{4\pi} n(r) dr. \quad (11)$$

where  $Q_s$  is scattering efficiency and  $P_\pi(\pi)$  is scattering phase function at  $180^\circ$ . Since the normalisation of  $n(r)$  is not known until the online phase of calculation, the backscatter-to-extinction ratio from particles  $k_p = \beta/\alpha_e$  can be calculated offline instead (the requirement for normalisation of  $n(r)$  is avoided by appearing in both the numerator and denominator):

$$k_p = \beta_{\text{p}}/\alpha_e = \frac{\int_0^\infty Q_s r^2 P_\pi(\pi)/(4\pi) n(r) dr}{\int_0^\infty Q_e r^2 n(r) dr}. \quad (12)$$

- 5 We pre-calculate this integral numerically for a permissible interval of  $r_{\text{eff}}$  (5–50  $\mu\text{m}$ ) at 500 evenly spaced wavelengths, and store the result as a lookup table for the online phase. The integral in the numerator is numerically hard to calculate due to strong dependency of  $P_\pi(\pi)$  on  $r$ . Figure 2b shows LR as a function of  $r_{\text{eff}}$ , calculated for log-normal and Gamma particle size distributions with  $\sigma_{\text{eff}} = 0.25r_{\text{eff}}$  and  $\sigma_{\text{eff}} = 0.5r_{\text{eff}}$ . This corresponds to the lookup table we use in the online phase of the lidar simulator. As can be seen in Fig. 2, LR depends only weakly on the choice of the distribution type and the effective standard  
10 deviation ratio.

### 4.3 Backscattering from ice crystals

Simulation of backscattering from ice crystals is relatively complex compared to backscattering from liquid droplets due to the very high variability of ice crystal microphysical properties such as habit, size, orientation and surface roughness, all of which affect LR, extinction cross section, single-scattering albedo and multiple scattering coefficient. Common habits include  
15 hexagonal plates, hexagonal columns, hollow hexagonal columns, droxtals, bullet rosettes, hollow bullet rosettes and aggregates (Baran, 2009; van Diedenhoven, 2017). Size can be highly variable and bimodal with a dependence on temperature and relative humidity. Orientation is commonly random or horizontally oriented (often reported with hexagonal ice plates). Surface can vary between smooth and rough depending on supersaturation and crystal age. In general, the Mie theory cannot be used to simulate backscattering from ice crystals because of their irregular shape (Yang et al., 2014). While large crystals allow the use of the  
20 geometric optics approximation to estimate the optical properties, smaller crystals and diffraction by large crystals necessitate the use of more advanced techniques such as the T-matrix method, finite-difference time domain (FDTD), discrete dipole approximation (DDA) and others, which are generally computationally expensive. Current global atmospheric models do not normally parametrise the microphysical properties of cloud ice explicitly, and provide only very limited information such as ice mass concentration and in some cases the effective radius of ice crystals in the model output. Radiative transfer schemes  
25 of atmospheric models do not explicitly evaluate backscattering (the phase function at  $180^\circ$ ) and therefore cannot provide this information to the simulator. Instead the phase function is parametrised by the asymmetry factor, which is likely insufficient to give an accurate estimate of backscattering.

Because the model ice crystal microphysical and optical properties are not known, they have to be parametrised. A first option is to parametrise the microphysical properties such as habit and size and calculate optical properties theoretically. A  
30 second option is to parametrise the optical properties directly. This appears to be a more practical choice because of the broad availability of global remote sensing measurements of optical properties from satellites and ground-based lidars, compared to relatively scarce in situ measurements of ice crystals. Garnier et al. (2015) analysed CALIPSO lidar and co-located passive

infrared data from the Imaging Infrared Radiometer (IIR) and determined a global relationship between temperature and LR and multiple scattering coefficient at the lidar wavelength of 532 nm. The multiple scattering coefficient is taken as a constant of 0.6 in the standard CALIPSO data processing, but they identified that it is in fact variable between about 0.4 and 0.8. Here, we parametrise LR based on their findings. LR varies with the lidar wavelength, a larger part of which is due to the change in the diffraction peak and a smaller part is due to the variation of the refractive index (Borovoi et al., 2014). We use colour ratio to estimate LR at lidar wavelengths other than 532 nm. Colour ratio of 1064 nm relative to 532 nm is commonly estimated for the dual-wavelength lidars such as CALIOP. Here, we use a value of 0.8 approximately consistent with the results of Bi et al. (2009) and Vaughan et al. (2010). The effective radius is defined for non-spherical particles as  $r_{\text{eff}} = \frac{3}{2} \frac{\text{IWC}}{\sigma}$ , where IWC is the ice water content, and  $\sigma$  is the volume extinction coefficient of ice. Heymsfield (2005) summarised ice crystal effective radius (related to "IWC/ $\sigma$ " by a factor of 1.64) parametrised as a function of temperature based on a number of field studies. We use this relationship for determination of the effective radius. Figure 2c shows the true and effective LR based on Garnier et al. (2015) and the effective radius based on Heymsfield (2005), parametrised by the following equations:

$$S = \left( 20 + (34 - 20) \frac{1/T - 1/200}{1/230 - 1/200} \right) \text{sr},$$

$$\eta = 0.8 + (0.5 - 0.8) \frac{1/T - 1/200}{1/240 - 1/200},$$

$$r_{\text{eff}} = \exp \left( \log(16.4) + (\log(49.2) - \log(16.4)) \frac{1/T - 1/213.15}{1/253.15 - 1/213.15} \right) \mu\text{m},$$

where  $T$  is atmospheric temperature in K.  $S$  follows (Garnier et al., 2015, Figure 12b),  $\eta$  follows (Garnier et al., 2015, Figure 9a) and  $r_{\text{eff}}$  follows (Heymsfield, 2005, Figure 2), where the concave and convex shape (respectively) as approximated by using  $1/T$  as an argument of the linear approximation, and we use a logarithmic scale of  $r_{\text{eff}}$  in the expression for  $r_{\text{eff}}$  to avoid negative values at low temperature. Figure 2c also shows LR when calculated with the assumption of  $\eta = 0.6$  ( $S_{\text{CALIPSO}, \eta=0.6}$ ) as in the standard processing of CALIPSO data. This corresponds to the empirically found relationship in (Garnier et al., 2015, Figure 12a) and (Josset et al., 2012, Figure 9) with a local maximum at 225 K. LR at wavelengths other than 532 nm is approximated by  $0.8 \frac{\lambda - 532}{532}$ , where  $\lambda$  is lidar wavelength in  $\mu\text{m}$  and 0.8 is the approximate value of 1064 nm/532 nm colour ratio. The parametrisation of LR ( $S$  in Fig. 2c) spans about the same range of values as reported by (Hopkin, 2018, Figure 5.6) (20 to 60 sr) and Yorks et al. (2011) (10 to 60 sr). Based on CALIPSO observations, Hu (2007) identified that while effective LR of global ice clouds at a lidar wavelength of 532 nm is mostly clustered around 17 sr, horizontally oriented plates produce a much lower effective LR below 10 sr caused by specular reflection. These results are close to our parametrisation of effective LR ( $S'_{\text{CALIPSO}}$ ). In the current version of the lidar simulator we do not parametrise horizontally oriented plates, but in a future version they could be taken into account by parametrising their concentration based on temperature (Noel and Chepfer, 2010). For the ALCs we use the same constant value of the multiple scattering coefficient  $\eta = 0.7$  as for liquid cloud droplets (Sect. 4.5).



#### 4.4 Cloud overlap and cloud fraction

Model cloud is defined by the liquid and ice mass mixing ratio and cloud fraction in each atmospheric layer. The lidar simulator simulates radiation passing vertically at a random location within the grid cell. Therefore, it is necessary to generate a random vertical cloud overlap based on the cloud fraction in each layer, as the overlap is not defined explicitly in the model output. Two common methods of generating overlap are the random and maximum–random overlap (Geleyn and Hollingsworth, 1979). In the random overlap method, each layer is either cloudy or clear with a probability given by CF, independent of other layers. The maximum–random overlap assumes that adjacent layers with non-zero CF are maximally overlapped, whereas layers separated by zero CF layers are randomly overlapped. COSP implements cloud overlap generation in the Subgrid Cloud Overlap Profile Sampler (SCOPS) (Klein and Jakob, 1999; Webb et al., 2001; Chepfer et al., 2008). The ALC lidar simulator uses SCOPS to generate 10 random subcolumns for each profile, using the maximum–random overlap assumption as the default setting of a user-configurable option. The ~~backscatter~~-attenuated volume backscattering coefficient profile and cloud occurrence can be plotted for any subcolumn. Due to the random nature of the overlap, the ~~backscatter~~-attenuated volume backscattering coefficient profile may differ from the observed profile even if the model is correct in its cloud simulation. The random overlap generation should, however, result in unbiased cloud statistics.

#### 4.5 Multiple scattering

Due to a finite FOV of the lidar receiver, a fraction of the laser radiation scattered forward will remain in FOV. Therefore, the effective attenuation is smaller than calculated with the assumption that all but the backscattered radiation is removed from FOV and cannot reach the receiver. The forward scattering can be repeated multiple times before a fraction of the radiation is backscattered, eventually reaching the receiver. To account for this multiple scattering effect, the COSP lidar simulator uses a multiple scattering correction coefficient  $\eta$ , by which the volume scattering coefficient is multiplied before calculating the layer optical thickness (Chiriaco et al., 2006; Chepfer et al., 2007, 2008). The theoretical value of  $\eta$  is between 0 and 1 and depends on the receiver FOV and optical properties of the cloud. For CALIOP at  $\lambda = 532$  nm a value of 0.7 is used in the COSP lidar simulator. Hogan (2006) implemented fast approximate multiple scattering code. This code has recently been used by Hopkin et al. (2019) in their ceilometer calibration method. They noted that  $\eta$  is usually between 0.7 and 0.85 for wavelengths between 905 and 1064 nm. The ALC simulator presented here does not use an explicit calculation of  $\eta$ , but retains the value of  $\eta = 0.7$ . ~~This value is also used when calculating LR from the vertically integrated backscatter (Sect. 5.2) for cloud droplets.~~ The code of Hogan (2006) "Multiscatter" is publicly available (<http://www.met.reading.ac.uk/clouds/multiscatter/>) and could be used in a later version of the framework to improve the accuracy of simulated attenuation and calibration.

### 5 Lidar data processing

Scheme in Fig. 1 outlines the processing done in the framework. The individual processing steps are described below.

## 5.1 Noise and subsampling

ALC signal reception is affected by a number of sources of noise such as sunlight and electronic noise (Kotthaus et al., 2016). Range-independent noise can be removed by assuming that the attenuated volume ~~backscatter~~backscattering coefficient at the highest range gate is dominated by noise. This is true if the highest range is not affected by clouds, aerosol, and if contributions from molecular scattering are negligible. The supported instruments have a range of approximately 8 (CL31), 15 (CL51, CHM 15k) and 30 km (MiniMPL). By assuming the distribution of noise at the highest level is approximately normal, the mean and standard deviation can be calculated from a sample over a period of time such as 5 minutes, which is short enough to assume the noise is constant over this period, and long enough to achieve accurate estimates of the standard deviation. The mean and standard deviation can then be scaled by the square of the range to estimate the distribution of range independent noise at each range bin. By subtracting the noise mean from the measured ~~backscatter~~attenuated volume backscattering coefficient we get the expected ~~backscatter~~attenuated volume backscattering coefficient. The result of the noise removal algorithm is the expected ~~backscatter~~attenuated volume backscattering coefficient and its standard deviation at each range bin.

## 5.2 Backscatter calibration

ALCs often report ~~backscatter~~attenuated volume backscattering coefficient in arbitrary units (a.u.) or as NRB (MiniMPL). If they report it in units of  $\text{m}^{-1}\text{sr}^{-1}$ , these values are often not calibrated to represent the true absolute ~~backscatter~~attenuated volume backscattering coefficient. Assuming that range-dependent corrections (overlap, dead time and afterpulse) have been applied on ~~the backscatter~~attenuated volume backscattering coefficient in a. u., the reported ~~backscatter~~attenuated volume backscattering coefficient is proportional to the true attenuated ~~backscatter~~volume backscattering coefficient (inclusive of noise ~~backscatter~~backscattering). In order to have a comparable quantity to the lidar simulator and consistent input to the subsequent processing (e.g. cloud detection), calibration by multiplying by a calibration coefficient is required. Formally, the units of the calibration coefficient depend on the units of backscattering recorded by the instrument, which are  $\text{m}^{-1}\text{sr}^{-1}$  in CL31, CL51, unitless in CHM 15k and  $\text{count}\cdot\mu\text{s}^{-1}\mu\text{J}^{-1}\text{km}^2$  in MiniMPL, i.e. the units of the calibration coefficient are  $\text{m}^{-1}\text{sr}^{-1}/(\text{instrument units})$ . In the following discussion, we leave out the units. Several methods of calibration have been described previously: calibration based on LR in fully attenuating liquid stratocumulus clouds (O'Connor et al., 2004; Hopkin et al., 2019), calibration based on molecular ~~backscatter~~backscattering (Wiegner et al., 2014) and calibration based on a high spectral resolution lidar reference (Heese et al., 2010; Jin et al., 2015). In addition, calibration can be assisted by sunphotometer or radiosonde measurements (Wiegner et al., 2014).

A relatively large variability of the calibration coefficient has been determined for instruments of the same model (Hopkin et al., 2019). However, past studies can be useful for determining an approximate value of the coefficient before applying one of the calibration methods. For the CL51, Jin et al. (2015) reported a value of  $1.2\pm 0.1$  based on a multi-wavelength lidar reference. Hopkin et al. (2019) reported mean values 1.4–1.5 for a number of CL31 instruments (software version 202). For CHM 15k, Hopkin et al. (2019) reported mean values between 0.3 and 0.8 for a majority of the instruments examined. The

**Table 5.** Theoretical molecular volume backscattering coefficient calculated at pressure 1000 hPa and temperature 20 °C and the calibration coefficient, relative to the instrument native units, determined for the instrument based on the molecular volume backscattering coefficient and stratocumulus lidar ratio calibration methods.

<u>Instrument</u>	<u>Wavelength (nm)</u>	<u>Molecular volume backscattering coefficient (<math>\times 10^{-6} \text{m}^{-1} \text{sr}^{-1}</math>)</u>	<u>Calibration coefficient</u>
<u>CHM 15k</u>	<u>1064</u>	<u>0.0906</u>	<u>0.34</u>
<u>CL31</u>	<u>910</u>	<u>0.172</u>	<u><math>1.45 \times 10^{-3}</math></u>
<u>CL51</u>	<u>910</u>	<u>0.172</u>	<u><math>1.2 \times 10^{-3}</math></u>
<u>MiniMPL</u>	<u>532</u>	<u>1.54</u>	<u><math>3.75 \times 10^{-6}</math></u>

ALCF provides per-instrument default values of the calibration coefficient (Table 5), but a unit-specific coefficient should be determined for an analysed instrument during the lidar data processing step.

Calibration based on LR in fully opaque liquid stratocumulus clouds has been applied successfully on large networks of ALCs. It utilises the fact that given suitable conditions vertically integrated ~~backscatter-attenuated volume backscattering~~ coefficient is proportional to LR of the cloud, which can be theoretically derived if the cloud droplet effective radius can be assumed. The theoretically derived value is about 18.8 sr for common ALC wavelengths and a relatively large range of effective radii (O'Connor et al., 2004). Another factor which needs to be known or assumed is the multiple scattering coefficient, which tends to be about 0.7-1.0 in common ALCs. Due to its relatively simple requirements, this method is possibly the easiest ALC calibration method. The ALCF implements this calibration method by letting the user identify time periods with fully opaque liquid stratocumulus cloud, for which the mean LR is calculated. The ratio of the observed LR and the theoretical LR is equivalent to the calibration coefficient. This implementation, while very easy to perform, has multiple limitations, some of which are highlighted by Hopkin et al. (2019):

1. Aerosol can cause additional attenuation and scattering, which results in LR which is different from the theoretical value by an unknown factor. Therefore, a frequent re-calibration may be necessary.
2. The multiple scattering coefficient assumption may not be accurate for the given instrument.
3. The 910 nm wavelength of CL31 and CL51 is affected by water vapour absorption which causes additional attenuation, which is currently not taken into account in the calculation of LR.
4. Near-range ~~backscatter-attenuated volume backscattering~~ coefficient retrieval is affected by receiver saturation and incomplete overlap. Therefore, using stratocumulus clouds above approximately 2 km for this calibration method is recommended. This range is instrument dependent.
5. The composition of the stratocumulus cloud may be uncertain. At temperature between 0 and -30°C these clouds may contain both liquid and ice which results in a different LR than expected.

These limitations could be addressed in the future by (1) using sunphotometer observations as an optional input to determine the aerosol optical depth (AOD), (2) calculating the multiple scattering coefficient more accurately (such as with the Multiscatter package of Hogan (2006)), (3) calculating the water vapour absorption explicitly based on water vapour, temperature and pressure fields from a reanalysis or radiosonde profile data, (4) correcting the near-range backscatter based on the integrated ~~backscatter~~-attenuated volume backscattering coefficient distribution as a function of height of the maximum backscatter (Hopkin et al., 2019, Sect. 5.1), (5) combining the ~~backscatter~~-attenuated volume backscattering coefficient profile with temperature field from a reanalysis to exclude cold clouds.

Molecular (Rayleigh) backscattering can be accurately calculated if temperature and pressure of the atmospheric profile is known (Sect. 4.2). This can be employed for absolute calibration of ALCs. Given the low SNR of low-power ALCs, several hours of integration are required to identify the molecular ~~backscatter~~-backscattering (Wiegner et al., 2014). The molecular ~~backscatter~~-backscattering is attenuated by an unknown amount of aerosol with unknown LR, and the near-range ~~backscatter~~-backscattering is affected by a potentially inaccurate overlap correction. Therefore, this method alone produces calibration coefficient which depend on the atmospheric conditions. We found that all studied ALCs except for the CL31 are capable of observing the molecular ~~backscatter~~-backscattering (Section 7). Therefore, this method may be used in addition to the liquid stratocumulus LR method for cross-validation of the calibration.

### 5.3 Cloud detection

Cloud is the most strongly attenuating feature in ALC ~~backscatter~~-attenuated volume backscattering coefficient measurements. Due to this attenuation, the lidar signal is quickly attenuated in thick cloud and can fall below the noise level before reaching the top of the cloud. This means that the first cloud base can be detected reliably (unless the cloud is too thin or too high and obscured by ~~backscatter~~-noise), while the cloud top or multi-layer cloud cannot be observed reliably under all conditions. The opposite is true for spaceborne lidars, which can detect the cloud top reliably but cannot always detect the cloud base. Therefore, ALC observations can be regarded as complementary to spaceborne lidar observations. By applying a suitable algorithm, one can detect CBH, CTH and identify cloud layers. Instrument firmware often determines CBH and sometimes cloud layers as part of its internal processing, often using an undisclosed algorithm which is not comparable between different instruments and potentially not even different versions of the instrument firmware (Kotthaus et al., 2016). Mattis et al. (2016) compared a large number of ALCs and found differences of up to 70 m between the reported CBH, and others found relatively large differences as well (Liu et al., 2015b; Silber et al., 2018). Alternatively to the instrument reported CBH and cloud layers, it is possible to detect cloud based on the ~~backscatter~~-attenuated volume backscattering coefficient profile. A relatively large number of cloud detection algorithms have been proposed (Wang and Sassen, 2001; Morille et al., 2007; Martucci et al., 2010; Van Tricht et al., 2014; Silber et al., 2018; Cromwell and Flynn, 2019). We use a simple algorithm based on ~~a-backscatter~~-an attenuated volume backscattering coefficient threshold applied on the denoised ~~attenuated~~-backscatter, assuming that the noise can be represented by a normal distribution at the highest range, which is unlikely to contain cloud or aerosol if the instrument is pointing vertically (this may not be true, however, for CL31 which has a maximum range of just 7.7 km). This assumption neglects the range-dependent molecular ~~backscatter~~-backscattering, which is relatively small at the ceilometer

wavelengths examined (910 nm and 1064 nm). A cloud mask is determined positive where ~~the backscatter attenuated volume backscattering coefficient~~ is greater than a chosen threshold plus ~~three-5~~ standard deviations of noise at the given range. In addition, the observed attenuated volume backscattering coefficient can optionally be coupled with simulated attenuated molecular volume backscattering coefficient and molecular backscattering removed from the observed backscattering prior to

5 cloud detection. This improves the results in the boundary layer, especially with instruments which operate in the visible range and therefore affected by large molecular backscattering (MiniMPL). A threshold of  $2 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  was found to be a good compromise between false detection and misses in our Southern Hemisphere data relatively unaffected by anthropogenic aerosol. Our observed and simulated results show that cloud backscatter is generally higher than  $1 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ , and a threshold below  $2 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  results in excessive false detection due to aerosol, molecular backscattering and noise from

10 sunlight. The threshold is an adjustable option of the ALCF. Users are encouraged to change this value if, for example, they data are affected by a large amount of aerosol. This value is above the maximum molecular ~~backscatter~~backscattering, which is approximately  $1.54 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  at the surface in the case of the MiniMPL (wavelength 532 nm). Noise is not simulated by the lidar simulator, but the cloud detection algorithm allows for coupling of simulated and observed profiles, whereby the noise standard deviation is taken from the corresponding location in the observed profile. With 5 minute averaging, when

15 the standard deviation of noise is relatively low, we found that the coupling does not make substantial differences to the detected cloud (not shown). While the threshold-based algorithm is less sophisticated than other methods of cloud detection, the vertical resolution of the simulated ~~backscatter attenuated volume backscattering coefficient~~ is likely too low and the vertical derivatives of the simulated ~~backscatter attenuated volume backscattering coefficient~~ too crudely represented (~~as can be seen in Fig. 4b-f, 5b-f, 6b-e~~Table 7) to apply any algorithm based on the vertical derivatives of ~~backscatter attenuated volume~~

20 backscattering coefficient. Using the same cloud detection algorithm on the observed and simulated ~~backscatter attenuated volume backscattering coefficient~~ is essential for an unbiased one-to-one comparison of cloud.

#### 5.4 Water vapour absorption

Previous studies have noted that ceilometers which utilise the wavelength of 910 nm such as the Vaisala CL31 and CL51 are affected by additional absorption of laser radiation by water vapour (Wiegner and Gasteiger, 2015; Wiegner et al., 2019;

25 Hopkin et al., 2019). The wavelength coincides with water vapour absorption bands between 900 and 930 nm, while the other common ceilometer wavelength of 1064 nm is not affected. Wiegner and Gasteiger (2015) reported that it can cause absorption of the order of 20% in the extratropics and 50% in the tropics. The lidar simulator does not currently account for this. However, as the water vapour concentration is available from the reanalyses and models, it should be possible to use a line-by-line model to calculate the water vapour volume absorption coefficient for each vertical layer during the integration process. Water

30 vapour also affects calibration of ~~observed backscatter~~the observed attenuated volume backscattering coefficient. In order to use the liquid stratocumulus LR calibration method, ~~the backscatter attenuated volume backscattering coefficient~~ has to be corrected for water vapour absorption to achieve high accuracy of calibration. Hopkin et al. (2019) used a simplified approach based on a parametrised curve and reported a difference from explicit radiative transfer calculations of 2% in the United Kingdom atmosphere (Middle Wallop). In the future either approach should be used to include water vapour absorption in the

**Table 6.** Location of sites and instruments. The time periods are inclusive.

<u>Site</u>	<u>Coordinates</u>	<u>Surface altitude (m)</u>	<u>Instruments</u>	<u>Time period</u>	<u>Missing period</u>	<u>Days</u>
<u>Cass</u>	<u>43.0346°S 171.7594°E</u>	<u>577</u>	<u>CL51</u>	<u>19 Sep–1 Oct 2014</u>		<u>13</u>
<u>Lauder</u>	<u>45.0379°S 169.6831°E</u>	<u>370</u>	<u>MiniMPL, CL31</u>	<u>12–24 Jan 2018</u>		<u>13</u>
<u>Christchurch</u>	<u>43.5225°S 172.5841°E</u>	<u>45</u>	<u>MiniMPL, CHM 15k</u>	<u>17 July–18 August 2019</u>	<u>22–31 July</u>	<u>23</u>

**Table 7.** Number of models levels and vertical resolution in the range of the instrument at the locations of the case studies. First number is the number of levels, followed by the minimum and maximum distance range between adjacent model levels in the lidar’s range (m).

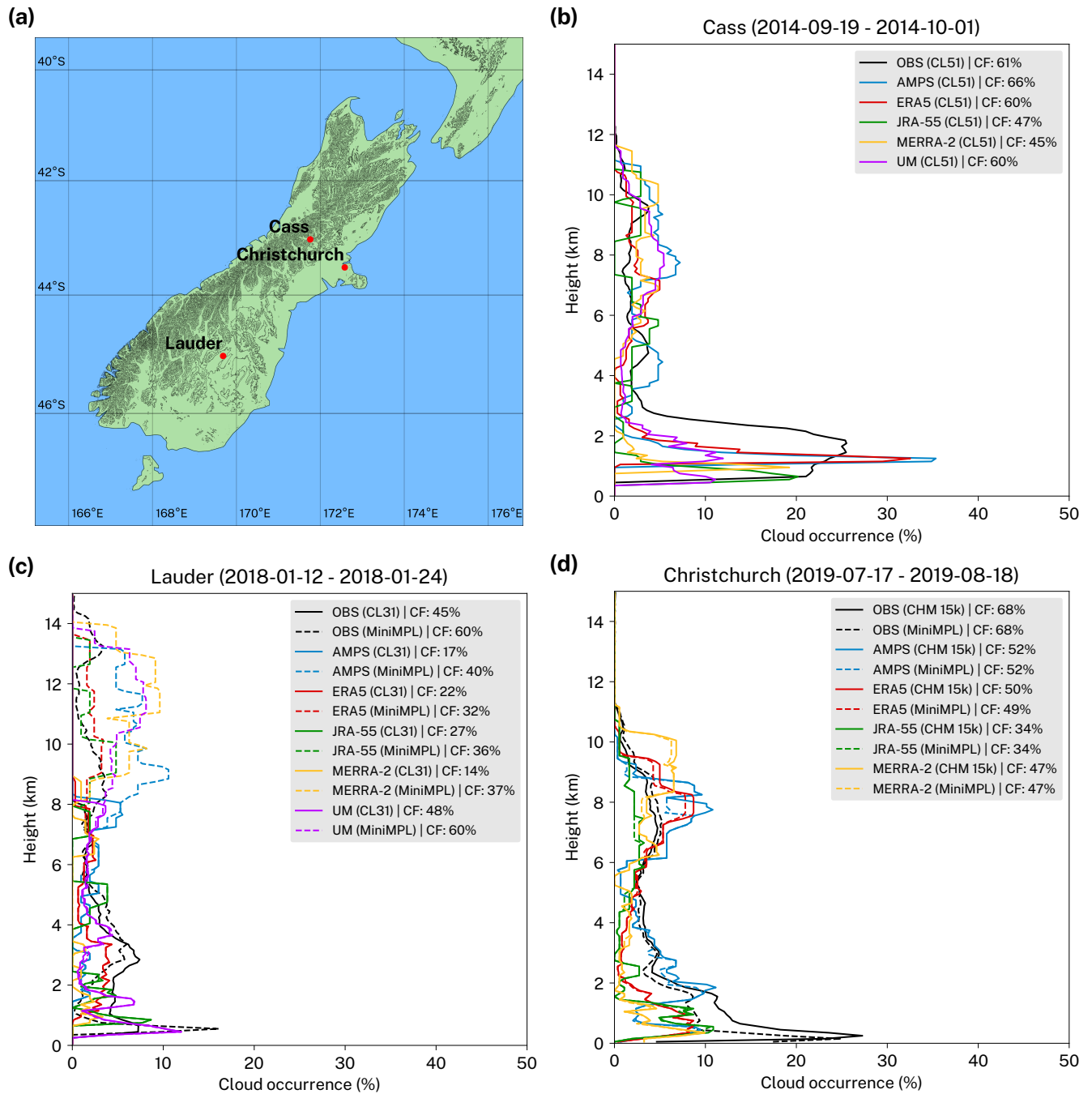
	<u>Cass (CL51)</u>	<u>Lauder (CL31)</u>	<u>Lauder (MiniMPL)</u>	<u>Christchurch (CHM 15k)</u>	<u>Christchurch (MiniMPL)</u>
<u>AMPS</u>	<u>42; 33–778</u>	<u>31; 35–528</u>	<u>59; 35–1021</u>	<u>43; 33–779</u>	<u>60; 33–870</u>
<u>ERA5</u>	<u>23; 222–1469</u>	<u>17; 220–950</u>	<u>30; 220–4748</u>	<u>25; 213–1425</u>	<u>31; 213–4107</u>
<u>JRA-55</u>	<u>23; 223–1479</u>	<u>17; 217–948</u>	<u>26; 217–1402</u>	<u>25; 213–1426</u>	<u>26; 213–1426</u>
<u>MERRA-2</u>	<u>34; 118–1080</u>	<u>26; 125–669</u>	<u>47; 125–1329</u>	<u>34; 124–1059</u>	<u>48; 124–1167</u>
<u>UM</u>	<u>44; 70–645</u>	<u>33; 32–449</u>	<u>65; 32–1181</u>		

simulator, or remove the effect of water vapour absorption from the observed lidar [backscatter attenuated volume backscattering coefficient](#) to achieve an improved one-to-one comparison between the observations, reanalyses and models.

## 6 Description of case studies

The case studies analysed here were selected to include all instruments supported by the framework. We compare four different instruments (CHM 15k, CL31, CL51, MiniMPL) deployed at three locations in NZ (Lauder, Christchurch, Cass) with three reanalyses (MERRA-2, ERA5, JRA-55), one NWP model (AMPS) and one GCM (UM). These case studies aim to demonstrate capability rather than to comprehensively evaluate cloud simulation in the models and reanalyses. The work detailed in Kuma et al. (2020) provides a detailed evaluation of the UM and MERRA-2 relative to shipborne ceilometer observations. Figure [?? 3a](#) shows the location of the sites and Table 6 summarises the case studies, which are also described in greater detail below. The sites were chosen from available datasets to demonstrate the use of the framework with all supported instruments. Two of the sites also had co-located instruments: CL31 and MiniMPL in Lauder, and CHM 15k and MiniMPL in Christchurch. The MiniMPL in Lauder and Christchurch were two different units. [The number of models levels within the range of each instrument and vertical resolution range are listed in Table 7.](#)

Cass is a field station of the University of Canterbury located at an altitude of 577 m in the Southern Alps of the South Island of NZ. The station is located far from any settlements and likely affected little by anthropogenic aerosol relative to the other sites. We have analysed 13 days of observations with a CL51 at this station performed in September and October 2014.



**Figure 3.** (a) Map showing the location of sites. Data at three sites in New Zealand were analysed: Cass, Lauder and Christchurch. (b), (c), (d) Cloud occurrence histograms as a function of height above the mean sea level observed at three sites and simulated by the lidar simulator based on atmospheric fields for five reanalyses and models. Shown is also the total cloud fraction (CF). The histogram is calculated from the cloud mask as determined by the cloud detection algorithm.

Lauder is a field station of NIWA located inland in the Central Otago region on the South Island of NZ. The station is situated in a rural area relatively far from large human settlements [at an altitude of 370 m](#). We have analysed 13 days of co-located MiniMPL and CL31 observations made in January 2018. The MiniMPL was operated in [an enclosure with a scanning head set to](#) a fixed vertical scanning mode during this period (elevation angle 90°).

5 Observations at the Christchurch site were performed at the University of Canterbury campus on the Ernest Rutherford building rooftop [at an altitude of 45 m](#). Christchurch is located on the east coast of the South Island of NZ. Its climate is affected by the ocean, its proximity to the hilly area of the Banks Peninsula, the Canterbury Plains and föhn-type winds (Canterbury northwester) resulting from its position on the lee side of the Southern Alps. The city is affected by significant wintertime air pollution from domestic wood burning and transport. The orography of the city and the adjacent Canterbury  
10 Plains is very flat, making it prone to inversions. The Ernest Rutherford building is a 5 floor building situated in an urban area, surrounded by multiple buildings of similar height. We have analysed [24-23](#) days of co-located MiniMPL and CHM 15k observations performed in July and August 2019. The MiniMPL was operated in [an enclosure with a scanning head set to](#) a fixed vertical scanning mode (elevation angle 90°). The nudged run of the UM was only available up to year 2018. Therefore, it was not analysed for this site.

## 15 7 ~~Case study results~~[Results](#)

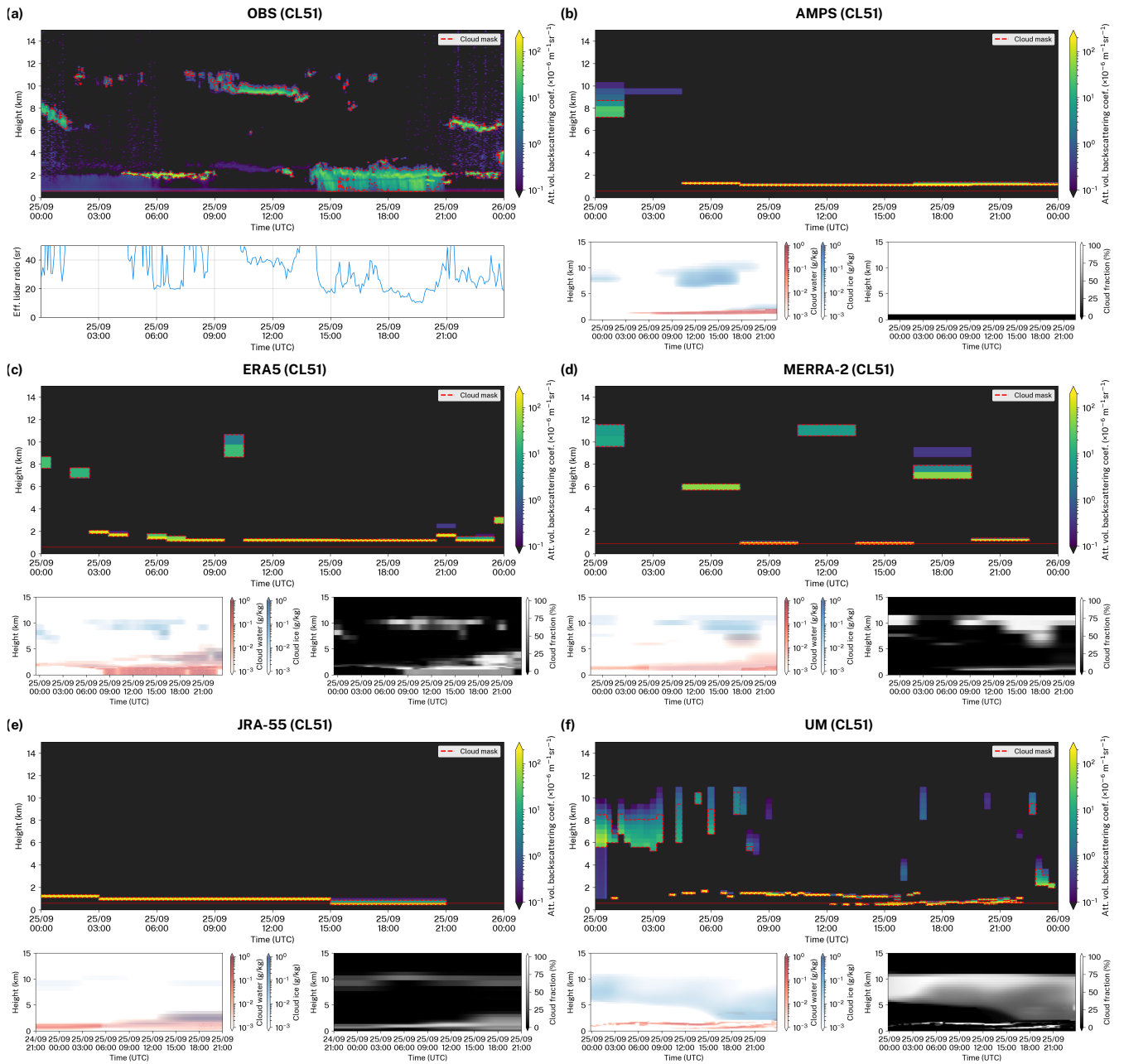
To demonstrate the ways that the ALCF can be used we compared a total of [50-49](#) days of ALC observations with simulated lidar ~~backscatter-attenuated volume backscattering coefficient~~ at three sites in NZ (Sect. 6). The observed ~~backscatter-attenuated volume backscattering coefficient~~ was normalised to calibrated absolute range-corrected ~~total-attenuated~~ volume backscattering coefficient. The noise mean as determined at the furthest range was removed from ~~the backscatter-attenuated~~  
20 ~~volume backscattering coefficient~~. Cloud detection based on an attenuated absolute ~~backscatter-volume backscattering~~ coefficient threshold of  $2 \text{ and three} \times 10^{-6} \text{ m}^{-1} \text{ sr}^{-1}$ , ~~after removing molecular backscattering and 5~~ noise standard deviations, was applied to derive a cloud mask and CBH. We compare the statistical cloud occurrence as a function of height above the mean sea level (ASL) (Fig. [3b, c, d](#)) and individual ~~backscatter-attenuated volume backscattering coefficient~~ profiles (selected profiles are shown in Fig. [4, 5](#) and [6](#)) in this section. [In these plots 5 standard deviations of the attenuated volume](#)  
25 [backscattering coefficient noise \(Sect. 5.3\) were removed. In addition, molecular backscattering was removed by coupling the observed data \(Fig. 4a, 5a, 6a\) with molecular attenuated volume backscattering coefficient calculated by the lidar simulator based the MERRA-2 reanalysis data. The same applies to model data \(Fig. 4b-f, 5b-f, 6b-e\), but molecular attenuated volume backscattering coefficient was calculated by the lidar simulator based the respective model data.](#)

### 7.1 [Cass](#)

30 We analysed 13 days of CL51 observations at the Cass field station in late winter. Due to the location of the station at a relatively high altitude in a varied terrain of the Southern Alps, the models with their relatively coarse horizontal grid resolution do not represent the terrain and position accurately. The orography representation of the models meant that the virtual altitude of the



## Cass (2014-09-25)



**Figure 4.** Examples of observed and simulated attenuated volume backscattering coefficient during 24 hours at Cass. The observed attenuated volume backscattering coefficient was normalised to absolute units and denoised. The first subcolumn generated by the Subgrid Cloud Overlap Profile Sampler (SCOPS) was used to make the plots. The red line is the station altitude. Shown is also (a) observed effective lidar ratio calculated by vertically integrating attenuated volume backscattering coefficient and (b-f) the corresponding model cloud liquid water, cloud ice and and cloud fraction.

## Lauder (2018-01-16)

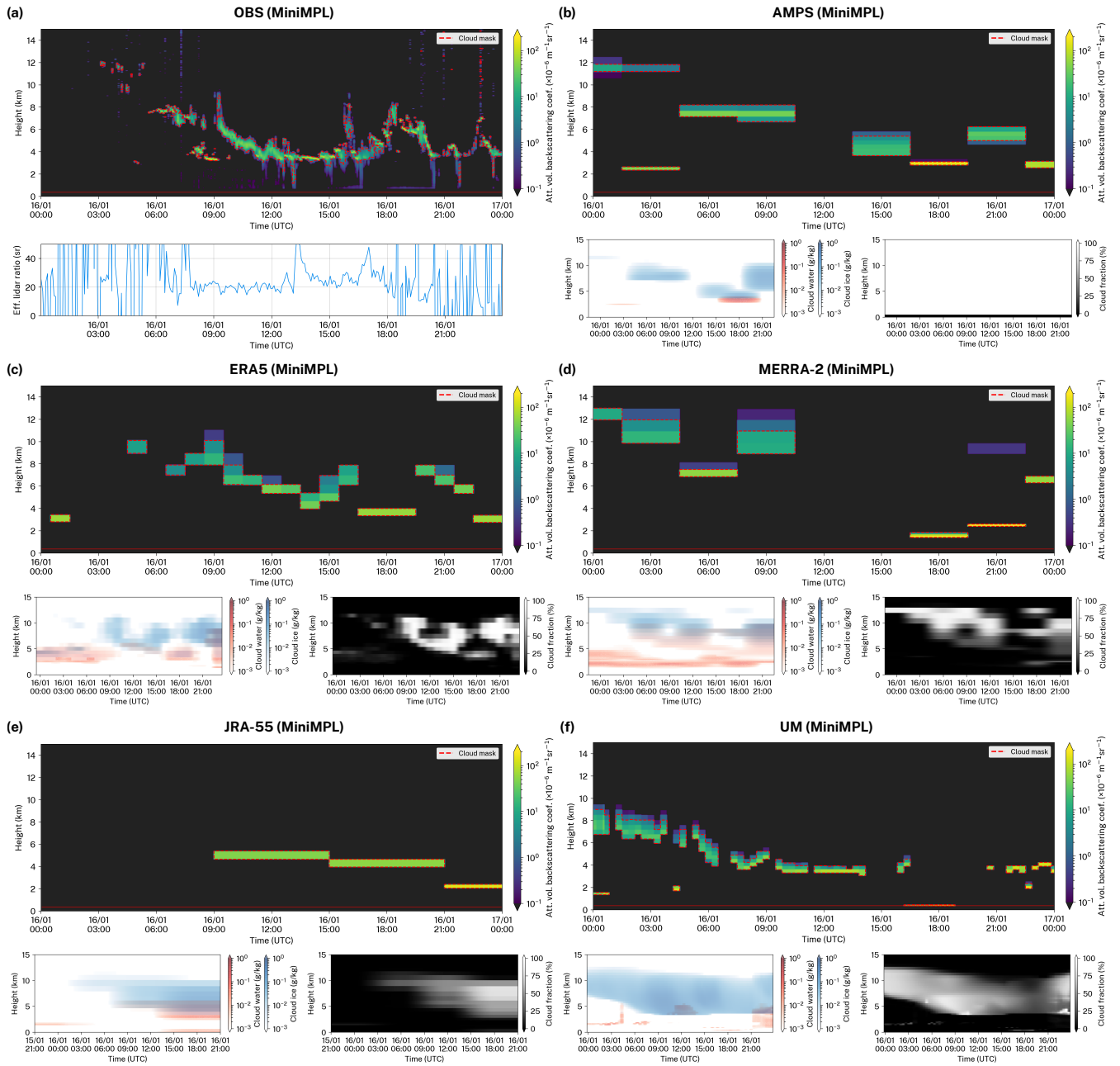


Figure 5. The same as Fig. 4 but for the Lauder.

## Christchurch (2019-07-18)

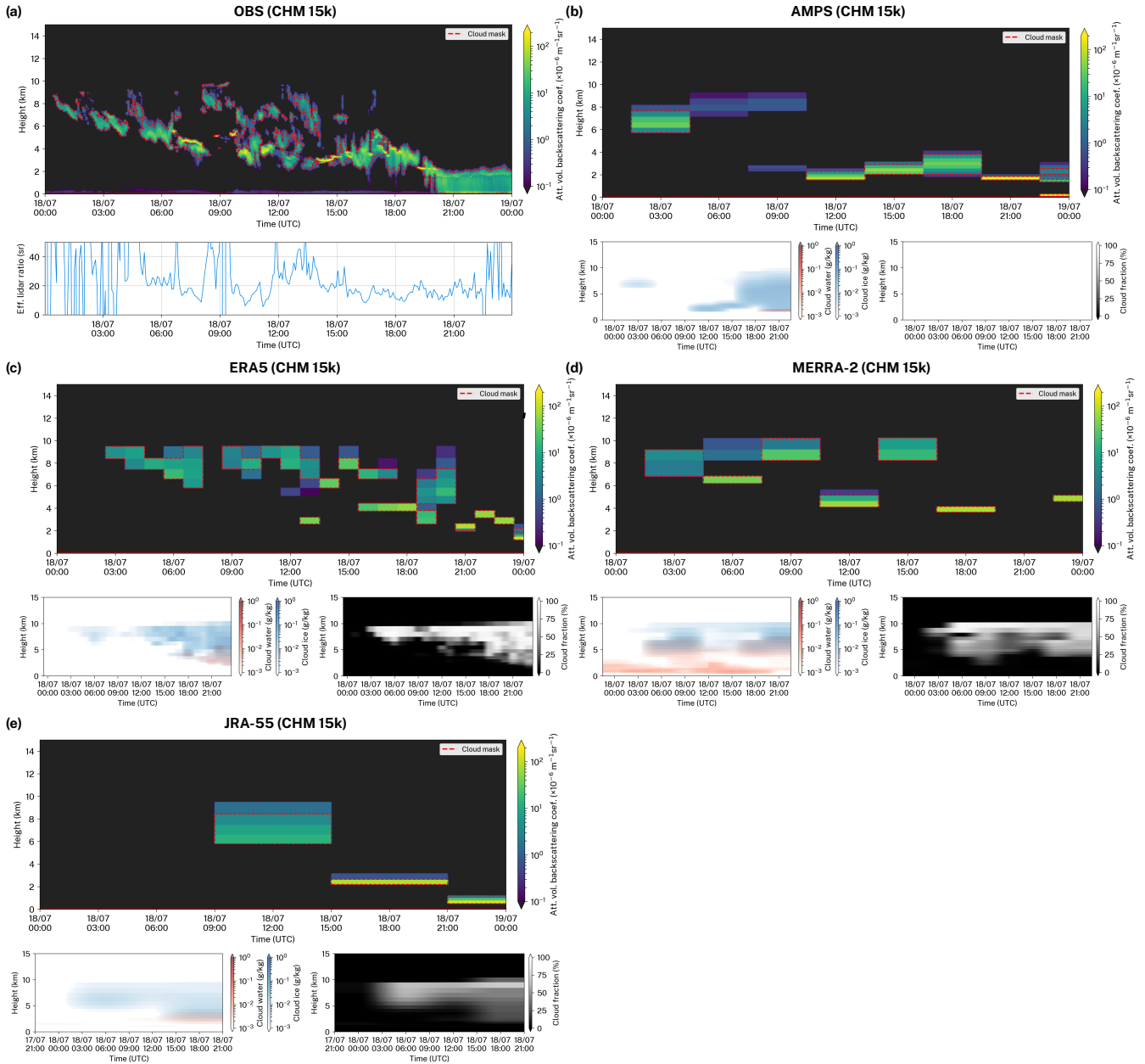


Figure 6. The same as Fig. 4 but for the Christchurch.

station was 1115 m (AMPS), 1051 m (ERA5), 401 m (JRA-55), 914 m (MERRA-2) and 428 m (UM). The virtual position, which is the centre of the nearest model grid cell to the site location, ranged from relatively close in the Southern Alps (AMPS, ERA5, MERRA-2, UM) to relatively far on the West Coast of NZ (JRA-55) depending on the horizontal resolution of the grid. The time period examined was characterised by diverse cloud occurrence with periods of low cloud and precipitation, mid-level cloud, fog, high cloud and clear skies. Precipitation, currently not simulated by the lidar simulator, was present in about 18% of the observed ~~backscatter-attenuated volume backscattering coefficient~~ profiles, as determined by visual inspection. Figure 3a-b shows that predominantly low cloud and precipitation between the ground and 2-3 km ASL in 25% of profiles was observed. Cloud between 3 and 12 km ASL was observed about evenly in 2% of profiles. While the reanalyses and models were able to partially reproduce the peak of cloud occurrence near 1 km ASL, the peak they displayed is less vertically broad than observed, and in the UM the peak was much weaker than observed. The lack of precipitation simulation might also have contributed to this apparent difference between observed and simulated cloud. Above 3 km ASL, the reanalyses and models tended to overestimate cloud, with only ERA5 and JRA-55 simulating close to the observed cloud occurrence. The observed total CF was 62.61%. AMPS overestimated this value by 4.5 percentage points (pp), ERA5 and the UM reproduced almost the exact value (within 2.1 pp), while the other reanalyses (JRA-55 and MERRA-2) underestimated CF by about 15 pp.

## 7.2 Lauder

We also analysed 13 days of CL31 and MiniMPL observations at the Lauder station in summer. During the time period relatively diverse cloud was observed, with periods of low, mid- and high cloud, clear sky and a small fraction of profiles with precipitation (about 3%). The altitude of the station of 370 m ASL generally had a much higher equivalent in the reanalyses and models: 565 m (AMPS), 642 m (ERA5), 681 m (JRA-55) and 786 m (MERRA-2) due to the presence of hills in the surrounding region (the station is in a high valley), with the exception of the UM where the altitude was 385 m. The virtual station position in the reanalyses and models ranged from relatively close to the station in the same geographical region (AMPS, ERA5), a nearby location in a more hilly region (JRA-55), a relatively distant location in the adjacent Dunstan Mountains (MERRA-2) and a relatively distant location in Central Otago (UM). Figure 3b-c shows that the CL31 observed relatively even cloud occurrence between the ground and 3 km ASL at 8%, falling off to about 3% between 4 and 8 km ASL (the maximum lidar range of CL31 is 7.7 km). The MiniMPL observed much weaker ~~backscatter-attenuated volume backscattering coefficient~~ than CL31 below 3 km ASL, which was identified as an overlap calibration issue in the MiniMPL. The MiniMPL observed substantial amounts of cloud above 8 km, not present in the CL31 observations due to its range limitation. Overall, the observed cloud occurrence had two peaks at ground to 3 km ASL and at about 9 km ASL. The simulated cloud occurrence was generally underestimated between the ground and 5 km ASL, with the exception of the UM which reproduced the lower half of the peak accurately, and ERA5 which reproduced the upper half of the peak accurately. Above 5 km ASL, the cloud occurrence was well reproduced in ERA5 and JRA-55, and strongly overestimated in AMPS, MERRA-2 and the UM. The reanalyses and models also tended to have two peaks at about 2 km ASL and 11 km ASL, but these were quite different from the observed peaks, with the lower peak underestimated by about 5 pp in the reanalyses and models and the higher peak overestimated by about 5–10 pp. The total CF was observed as ~~44% and 58.45%~~ and 60% by CL31 and MiniMPL, respectively. CF observed by the MiniMPL was

likely higher due to its higher maximum lidar range (CL31 missed a substantial amounts of high cloud due to this limitation). The total CF was strongly underestimated by the reanalyses and models by up to ~~34~~31 pp (CL31) and ~~38~~28 pp (MiniMPL), with the exception of the UM which simulated the correct CF within 3 pp.

### 7.3 Christchurch

5 The Christchurch observations were taken during a total of ~~24~~23 days in mid- to late winter. The cloud situations were characterised by the frequent occurrence of low cloud and fog, with relatively diverse mid- and high level cloud and periods of clear sky also present (not shown). Precipitation was present in about 9% of profiles and fog in about 11% of profiles. As the site location is relatively flat (Canterbury Plains), the models ~~didn't~~did not have any difficulty in reproducing the altitude of the site, which was 32 m (AMPS), 72 m (ERA5), 143 m (JRA-55) and 76 m (MERRA-2). The virtual location was within  
10 the boundaries of the city (AMPS), on the Canterbury Plains close to the city boundaries (ERA5, MERRA-2), and over Lake Ellesmere about 20 km from the city (JRA-55). Figure ~~3e~~d shows that the co-located CHM 15k and MiniMPL observed a strong peak of cloud occurrence of 26% (CHM 15k) ~~and 20% (MiniMPL)~~ at about 500 m ASL. This was likely due to the combined precipitation and fog as well as false detection of aerosol as cloud. The observed cloud occurrence had a local minimum of 2% at about 5 km ASL, a secondary peak of 5% at 7 km ASL, and fell off 0% at 11 km ASL. The CHM 15k and  
15 MiniMPL observations showed inconsistencies of up to ~~5~~4 pp. The reanalyses and models ~~strongly~~ underestimated low cloud by ~~about 10-20~~5-10 pp. With the exception of AMPS, they underestimated mid-level cloud by about 5 pp and represented high cloud relatively accurately. The total CF observed was ~~69% and 66% for the CHM 15k and MiniMPL, respectively~~68%, while the reanalyses and models strongly underestimated CF by up to ~~36~~34 pp (JRA-55), with underestimates around 20 pp common.

### 20 7.4 Backscattering on daily scales

Figures 4, 5, 6 show images of ~~backscatter~~attenuated volume backscattering coefficient for three separate days taken from the three case studies. The selected days represent some of the best-matching profiles and demonstrate how well the reanalyses and models can simulate cloud under favourable conditions. As can be seen in the figures, ~~the UM performs~~ERA5 and the UM perform the best in terms of temporal and height accuracy of the simulated cloud (Fig. ~~4f, 5f~~), ~~followed by ERA5 (Fig. 4e, 4f, 5c, 5f, 6c)~~. This is likely due to the high output temporal resolution of the UM and ERA5 of 20 min. and 1 h, respectively. The UM and ERA5 were able to represent the relatively fine structure of cloud and to a lesser extent the optical thickness (inferred from the strength of backscattering) of the cloud. Deficiencies, however, are readily identifiable. The low cloud in the UM (Fig. 4f) ~~appears to be shifted by several hours~~covers too large area relative to observations (Fig. 4a) and the high cloud has a greater vertical extent in the UM. Likewise, the altocumulus cloud observed in Fig. 5a is shifted by  
30 several hours in the UM (Fig. 5f). The stratocumulus and nimbostratus cloud, identified visually based on the ~~backscatter~~attenuated volume backscattering coefficient profiles, in ERA5 (Fig. 4c) is markedly lower than observed (Fig. 4a), as well as optically thicker than in reality. The mid-level cloud in ERA5 (Fig. 5c) was located about 2 km higher than observed (Fig. 5a). Precipitation observed in Fig. 6a towards the end of the analysed period was not present in the ERA5 simulated profile (Fig.

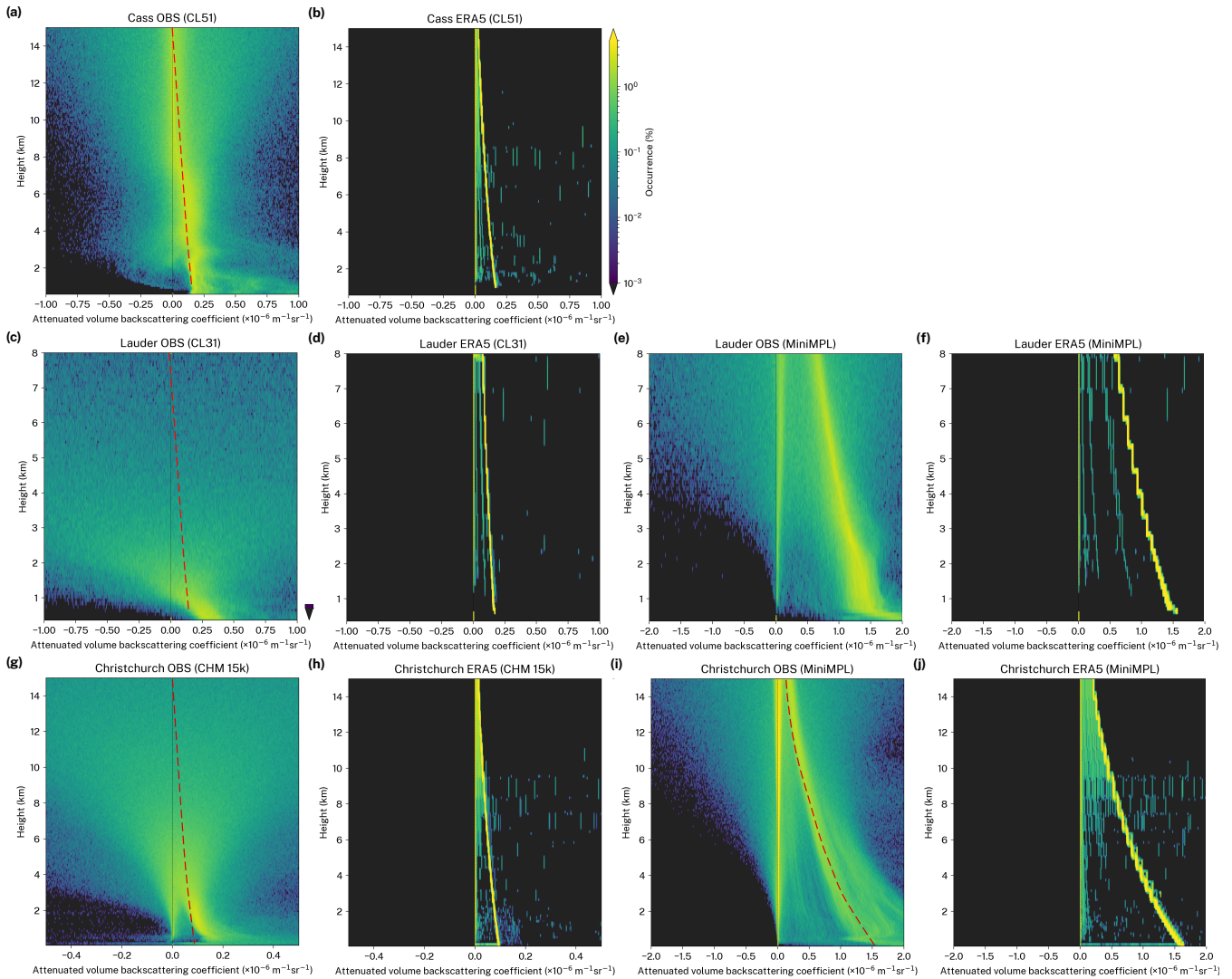
6c) due to lack of precipitation simulation in the current lidar simulator (even though rain and show snow specific content is available from the reanalysis). AMPS and MERRA-2 was the third-best performing reanalysis in terms of had lower cloud representation accuracy. The reanalysis-They managed to capture the overall structure of clouds (Fig. 4c, g, kb, 4d, 5b, 5d, 6b, 6d), but substantial discrepancies were present, some of which were likely due to the relatively low temporal resolution of the reanalysis (3 h). AMPS was identified as the fourth-best in terms of cloud representation accuracy despite the highest spatial resolution of the model. Fig. 4h shows that while the overall structure of the cloud is represented in the reanalysis, it was not able to represent any of the finer details. This is likely partially due to the temporal resolution of the model output of just 3 h (AMPS- AMPS has, however, relatively high horizontal grid resolution of 21 km). A visual comparison of AMPS and MERRA-2 with observations suggests that the the MERRA-2 simulations are more accurate than AMPS despite the fact that the horizontal grid resolution of AMPS is much finer than MERRA-2. This demonstrates that other factors in the model than resolution have stronger influence on the quality of cloud simulation. JRA-55 was identified as the last in terms of cloud representation accuracy. JRA-55 has the lowest temporal resolution of the studied reanalyses and models of just 6 h, as well as the lowest horizontal grid resolution of 139 km. Therefore, it cannot be expected to capture any fine details of cloud. In the presented profiles (Fig. 4d, 4e, 5e, 6e) one can see that the cloud is only crudely represented. JRA-55 was able to represent the stratocumulus cloud of Fig. 4a, although its temporal extent and optical thickness were overestimated. The mid-level cloud clouds of Fig. 4i was 5a and 6a were relatively well-represented in terms of height and optical thickness (Fig. 4i), given the low temporal resolution of the reanalysis. We stress that a direct backscatter attenuated volume backscattering coefficient profile intercomparison is highly dependent on the temporal resolution of the model output. The statistical intercomparison, however, should still give unbiased results if the cloud physics is accurately simulated by the atmospheric model.

The comparison between observations, reanalyses and models can be summarised by considering that the radiative effect of clouds is the product of cloud occurrence/fraction and their albedo. CF can be derived from ALC observations by determining the cloud mask and identifying cloudy profiles which contain detected cloud. Cloud albedo cannot be directly observed with ALCs due to the unknown backscatter-to-extinction ratio of the cloud layers, which depends on the cloud phase, the effective radius and the shape of ice crystals. However, we can use LR<sup>-1</sup>, calculated from the vertically integrated backscatter, as a proxy for cloud albedo. Clouds which are not fully opaque result in LR<sup>-1</sup> which is greater than the expected theoretical LR<sup>-1</sup> based on the Mie theory (Fig. 2). The inverse of LR<sup>-1</sup> (extinction-to-backscatter ratio) is therefore approximately proportional to the albedo of the cloud. Figure ??a, b shows an absolute and relative (respectively) scatter plot of the average cloud fraction and the average of the inverse of LR<sup>-1</sup> calculated from cloudy profiles. As already shown in Fig. 3, the reanalyses and models generally underestimate CF. However, Fig. ?? shows that this is compensated by overestimated albedo (higher LR<sup>-1</sup> indicates higher cloud albedo), which is higher in all reanalyses and models than the corresponding ALC observations. Further examination shows that the UM and ERA5 show the closest proximity to the corresponding observations in terms of the average CF and LR<sup>-1</sup>. The Lauder CL31 and MiniMPL also show a stark difference in CF, with the CL31 detecting much smaller CF due to its maximum vertical range limitation of 7.7 km, which causes a significant fraction of high cloud to be omitted. The results shown in Fig. ?? suggest that the "too few too bright" model cloud problem identified in previous studies (Nam et al., 2012; Klein et al., 2013; Wall et al., 2017; Kuma et al., 2020) is also observed in the current case studies;

whereby the simulated CF is too low, but its exaggerated albedo means that the top of atmosphere (TOA) shortwave radiative balance can be approximately correct. Figure 4a, 5a, 6a also show effective LR of observations calculated by integrating vertically attenuated volume backscattering coefficient (Sect. 4.1). If attenuated volume backscattering coefficient is properly calibrated, under fully attenuating cloud conditions effective LR converges to the theoretical value of LR of liquid cloud droplets (approximately 18.8 sr at near IR wavelengths) multiplied by the multiple scattering coefficient (approximately 0.7; Sect. 4.5).

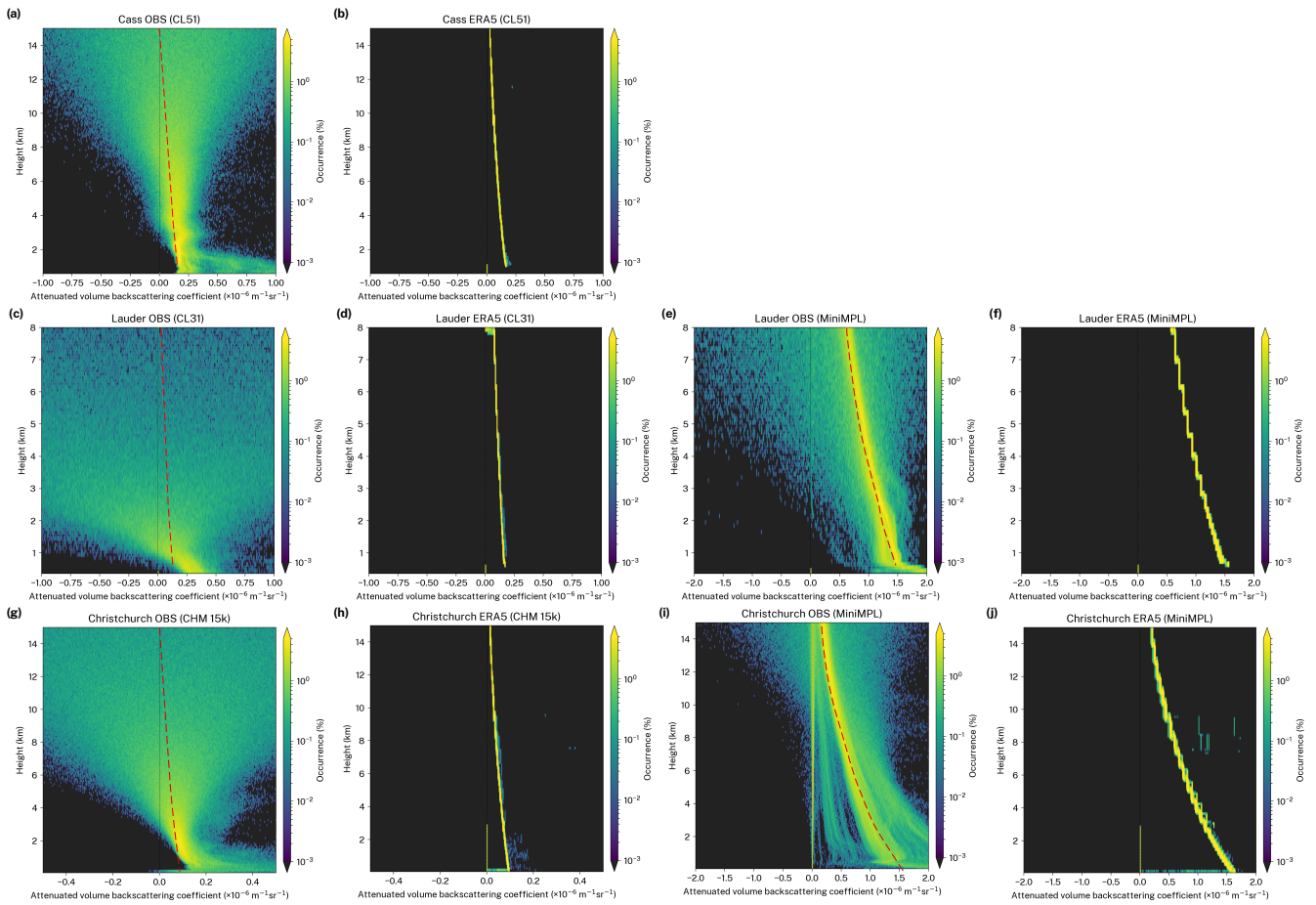
## 7.5 Molecular backscattering, aerosol backscattering and noise

Figure 7 shows ~~backscatter~~ attenuated volume backscattering coefficient histograms as a function of height for small values of ~~backscatter~~ the coefficient (up to  $2 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ ) observed and simulated at the sites of the case studies, calculated for the entire time period of each case study. The scale of values is below cloud ~~backscatter~~ backscattering, and therefore shows ~~backscatter~~ backscattering which results from molecular and aerosol scattering and noise. Molecular backscattering depends on the atmospheric pressure and temperature as well as the lidar wavelength. It causes the main "streak" (a local maximum) visible in each of the histograms. The observed molecular ~~backscatter~~ attenuated volume backscattering coefficient at the surface approximately corresponds to the theoretically calculated value at each wavelength:  $0.0906 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  ( $\lambda = 1064 \text{ nm}$ ),  $0.172 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  ( $\lambda = 910 \text{ nm}$ ) and  $1.54 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  ( $\lambda = 532 \text{ nm}$ ) at 1000 hPa and  $20^\circ \text{C}$  (Table 5). The molecular backscattering in the boundary layer is, however, superimposed on backscattering by aerosol and cloud. In the case of the MiniMPL observations at the Christchurch site (Fig. 7i), the molecular ~~backscatter~~ attenuated volume backscattering coefficient streak has multiple secondary streaks. These are caused by different levels of attenuation by cloud and aerosol during the period of the observations. These secondary streaks were also partially reproduced by the simulator (Fig. 7j). A smaller portion of the width of the streak is also caused by fluctuations of atmospheric temperature and pressure. Under suitable conditions, ~~the molecular backscatter~~ molecular attenuated volume backscattering coefficient can be used for absolute calibration of an instrument. With the exception of CL31 (Fig. 7c), the molecular ~~backscatter~~ backscattering can be identified in the observed ~~backscatter~~ attenuated volume backscattering coefficient in each case. Therefore, it is possible to choose a calibration coefficient such that the observed and simulated molecular ~~backscatter~~ attenuated volume backscattering coefficients overlap. This can be considered a viable alternative to the liquid stratocumulus LR calibration method, or as a means of cross-validating the instrument calibration. However, it should be noted that the accuracy of this method is affected by an unknown amount of aerosol attenuation. Cloudy profiles can be filtered when calculating the histogram, and therefore the effect of cloud attenuation can be minimised. In addition to the molecular ~~backscatter~~ attenuated volume backscattering coefficient streak, there is a zero-centred streak visible in the histograms. This is caused by noise when the signal is fully attenuated by cloud. Lastly, a zero-centred "cone" of noise is visible in the observed ~~backscatter~~ attenuated volume backscattering coefficient, increasing with the square of range. The size of this cone is particularly large in the case of the CL31 (Fig. 7c), which is most likely the result of its low receiver sensitivity and low power compared to the other instruments. The standard deviation of the cone at the furthest range is used to determine the noise standard deviation used by the cloud detection algorithm (Sect. 5.3).



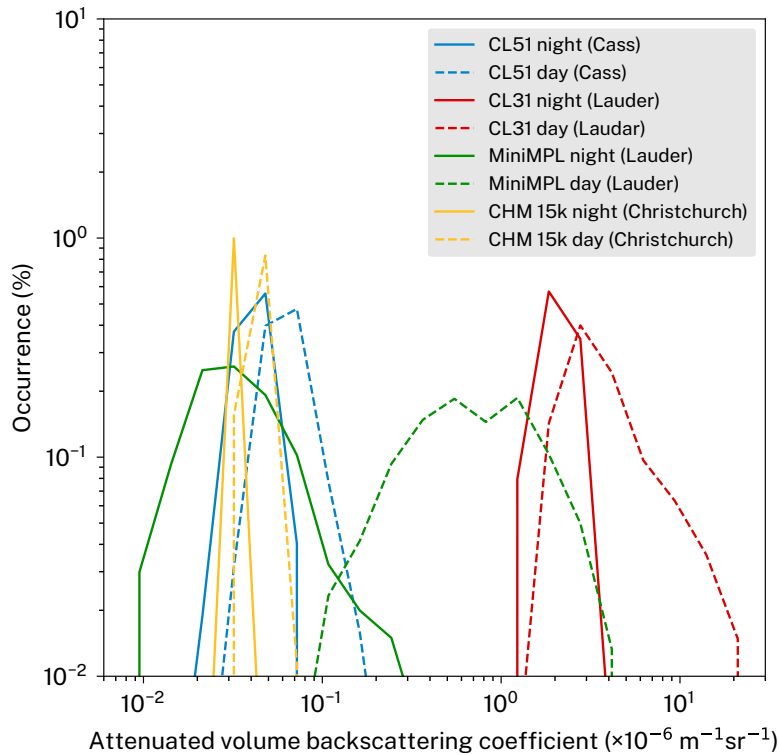
**Figure 7.** Attenuated volume backscattering coefficient histograms as a function of height observed and simulated at three different sites of the case studies calculated from all profiles. The plots show the distribution of attenuated volume backscattering coefficient for values which are on the scale of noise, molecular and aerosol backscattering ( $[-0.5, 0.5]$  for CHM 15k,  $[-1, 1]$  for CL31 and CL51 and  $[-2, 2] \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$  for MiniMPL). The simulated attenuated volume backscattering coefficient is based on the ERA5 atmospheric fields. Visible in the plots is backscattering caused by molecular backscattering (the main "streak"), noise when signal is fully attenuated by cloud (the zero-centred "streak"), and the range-dependent noise (the zero-centred "cone"). The molecular backscattering is marked by a red dashed line on the observed attenuated volume backscattering coefficient plots, the shape of which is taken from the simulated molecular attenuated volume backscattering coefficient for the corresponding instrument and site.





**Figure 8.** The same as Fig. 7 but calculated from clear sky profiles only.

Figure 8 shows the same information as Fig. 7, but for clear sky profiles only. Here, it can be seen that the zero-centred peak caused by the complete attenuation by cloud is no longer present. There is a clear overlap between the centre of the noise cone with the simulated molecular ~~backscatter~~attenuated volume backscattering coefficient; i.e. the noise cone is centred at the observed molecular ~~backscatter~~attenuated volume backscattering coefficient. This is visible with all instruments including CL31 (Fig. 8c), where the overlap between the observed and simulated molecular ~~backscatter~~attenuated volume backscattering coefficient is most clearly visible at about 1 km ASL. Below 1 km ASL, the effect of boundary layer aerosol distorts ~~the molecular backscatter~~molecular attenuated volume backscattering coefficient by an unknown quantity. The clear sky histograms as shown in Fig. ~~7-8~~7-8 may therefore be preferable to the all-sky histograms of Fig. 7 for calibration by fitting ~~the molecular backscatter~~. Note that Figure ~~8i~~8i does not show a good match with the simulated molecular backscatter (Fig. 8j) likely due to miscalibration of the MiniMPL in the lower troposphere, which causes spurious cloud to be detected. The majority of profiles were therefore filtered out as ~~cloudy~~molecular attenuated volume backscattering coefficient. The dead time, after-



**Figure 9.** Attenuated volume backscattering coefficient noise standard deviation histogram calculated for each instrument at sites of the case studies from clear sky profiles over the whole time period. The noise distribution is calculated at the furthest range. Shown is the range-scaled noise distribution at a range of 8 km. "Night" and "day" distributions are calculated separately from nighttime and daytime profiles only.

pulse and overlap MiniMPL calibration supplied by the vendor appears to be deficient and causes range-dependent bias in the backscatter-attenuated volume backscattering coefficient profile.

We now examine the noise in each instrument using the ALCF. Figure 9 shows the distribution of standard deviation of backscatter noise determined at the highest observable range of each instrument and range-scaled to 7-8 km. It can be seen that the CL31 is affected by the greatest amount of noise, peaking at about  $2 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ . This is at the threshold of cloud detection of  $2 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ . Therefore, thin cloud may be obscured by noise at higher ranges with this instrument. The MiniMPL, operating in the visible spectral range, shows a strongly bimodal distribution of backscatter-the attenuated volume backscattering coefficient noise depending on sunlight. During daytime, it peaks at about  $0.7 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ , which is the second highest of the analysed instruments. During nighttime, it peaks at about  $0.02 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ , which is the lowest of the analysed instruments. The CHM 15k histogram-peaks-and CL51 peak between the nighttime and daytime MiniMPL at 0.2 about  $0.05 \times 10^{-6} \text{m}^{-1} \text{sr}^{-1}$ . All of CL31, CL51 observed the second lowest amount of noise after the nighttime MiniMPL at about 0.04 and CHM 15k show a slight reduction of noise during nighttime, presumably because of a small amount of

incoming solar radiation at near IR wavelengths. The difference between the nighttime and daytime ~~backscatter-attenuated volume backscattering coefficient~~ noise in the MiniMPL has been previously analysed by Silber et al. (2018) (Fig. S3) and these results confirm their findings.

## 8 Discussion and conclusions

- 5 We presented the Automatic Lidar and Ceilometer Framework, which combines lidar processing and lidar simulation for the purpose of model evaluation. The lidar simulation is based on the COSP spaceborne lidar simulator by accounting for the different geometry and lidar wavelength. ~~New~~ We calculated new lookup tables for Mie scattering ~~were calculated~~ for a number of ALC wavelengths ~~and~~, developed ice crystal backscattering parametrisation based on temperature and implemented noise removal and cloud detection algorithms ~~were implemented~~. The framework supports the most common ALCs and reanalyses.
- 10 We demonstrated the use of the framework on ALC observations at three different sites in New Zealand, and applied the lidar simulator to three reanalyses and two models. We found that while some reanalyses and models such as the UM and ERA5 show relatively good correspondence with observed cloud, others performed relatively poorly in our time-limited local comparison. All reanalyses and models underestimated the total CF by up to ~~38~~ 34 pp, with underestimation by 20 pp common. In some cases, the observed and simulated ~~backscatter-attenuated volume backscattering coefficient~~ profiles matched relatively
- 15 closely in terms of time and altitude, and a better match was observed with reanalyses with high output temporal resolution such as the UM and ERA5, while reanalyses with low temporal resolution ~~didn't~~ did not allow for reliable direct (non-statistical) comparison of cloud. However, it is clear that more factors than the horizontal and vertical resolution influence the cloud simulation accuracy; especially the cloud, boundary layer and convection schemes employed by the atmospheric model. The reanalysis and model output temporal, horizontal grid resolution and vertical resolution are not always the same as the internal
- 20 resolution of the underlying atmospheric model. Both have an impact on the comparison between the simulated and observed ~~backscatter-attenuated volume backscattering coefficient~~ and cloud. While the output resolution should not have an impact on the long-term statistics, it can be a limiting factor for direct ~~backscatter-attenuated volume backscattering coefficient~~ profile comparison. We demonstrated that the ALCF could be used to identify substantial differences in cloud ~~backscatter-attenuated volume backscattering coefficient~~ which were present in all reanalyses and models. We showed that all the studied instruments
- 25 except for the CL31 are capable of detecting molecular ~~backscatter-backscattering~~ and that this can be used for calibration or for cross-validation of other calibration methods. We found that the nighttime MiniMPL was subject to the least amount of noise of all the instrument examined, followed by the CL51, CHM 15k, daytime MiniMPL and CL31. Noise in the MiniMPL, and to a lesser extent in the other ALCs, was shown to have a bimodal distribution due to day/nighttime. The ALCF can therefore be useful for testing the quality of collected data.
- 30 Currently the framework has several limitations which should be addressed in the future. The water vapour absorption at 910 nm likely affects the instrument calibration of the CL31 and CL51 ceilometers and limits the accuracy of the one-to-one comparison, even though due to the relatively high ~~backscatter-backscattering~~ caused by cloud, the calculated cloud masks are unlikely to be strongly affected. The lidar simulator currently does not simulate backscattering from precipitation.

Observed precipitation is generally detected as "cloud" by the cloud detection algorithm, while the simulated profile contains no ~~backscatter~~ backscattering at the location of precipitation (backscattering and attenuation by rain drops and snow should be implemented in the lidar simulator in the future). If desired, the ~~backscatter~~ attenuated volume backscattering coefficient profiles affected by precipitation can be excluded before the comparison or their fraction determined by visually inspecting

5 the observed ~~backscatter~~ attenuated volume backscattering to assess their possible effect on the statistical results. ~~In addition, backscattering from ice crystals is currently treated in the same way as backscattering from liquid droplets, i.e. the same effective radius of spheres is assumed and the particle concentration is calculated from the model ice mass mixing ratio in the atmosphere layer. In reality, the non-spherical shape of ice crystals, the different typical effective radius and refractive index of ice result in different backscattering from ice crystals versus liquid droplets for the same mass mixing ratio. The ice mass~~

10 ~~mixing ratio reported by models is usually much lower than the liquid mass mixing ratio, therefore the simulated backscatter from cold cloud tends to be much lower than from warm cloud-~~ Aerosol is also not currently implemented in the simulator. Previous studies (Chan et al., 2018) characterised optical parameters of different groups of aerosol, which could be used in a future version of the simulator with models which provide concentration of aerosol in their output. In our case studies aerosol volume backscattering coefficient was less than  $2 \times 10^{-6} \text{ m}^{-1} \text{ sr}^{-1}$  and below 4 km, which could result in worst-case two-way

15 attenuation of about 50% assuming LR of 50 sr. This should not preclude cloud detection due to the large magnitude of typical cloud backscattering. The ALCs also suffer from various measurement deficiencies. Notably incomplete overlap, dead time and afterpulse corrections tend to give sub-optimal results at the near range. It is possible to use semi-automated methods to correct for these deficiencies, such as by calculating the integrated ~~backscatter~~ attenuated volume backscattering coefficient distribution by height of the maximum ~~backscatter~~ backscattering and correcting for the range-dependent bias (Hopkin et al.,

20 2019, Sect. 5.1). This method could be implemented in the framework to enable range-dependent calibration of the observed ~~backscatter~~ attenuated volume backscattering coefficient.

The presented framework streamlines lidar data processing and tasks related to lidar simulation and model comparison. The framework was recently used by Kuma et al. (2020) for Southern Ocean model cloud evaluation in the GA7.1 model and MERRA-2 reanalysis. Considering the existing extensive ALC networks worldwide there is a wealth of global data. We

25 therefore think that ALCs should have a greater role in model evaluation. Satellite observations have long been established in this respect due to their availability, spatial and temporal coverage and their well-developed derived products and tools. ALCs, with their diverse formats and decentralised nature, have so far lacked derived products and tools which would make them more accessible for model evaluation. We hope that this software will enable more model evaluation studies based on ALC observations. Development of lidar data processing is currently hampered by closed development of code. We note that code

30 has very rarely been made available with past ALC studies. Continued improvement of publicly available code for lidar data processing is needed to achieve faster development of ground-based remote sensing and make it more attractive for GCM, NWP model and reanalysis evaluation.

*Code and data availability.* The *ALCF* is open source and available at <https://alcf-lidar.github.io> and as a permanent archive of code and technical documentation on Zenodo at <https://doi.org/10.5281/zenodo.4088217>. The technical documentation is also in the Supplementary information. A tool for converting Vaisala CL31 and CL51 data files to NetCDF *cl2nc* is open source and available at <https://github.com/peterkuma/cl2nc>. A tool for converting MiniMPL raw binary data files to NetCDF *mpl2nc* is open source and available at <https://github.com/peterkuma/mpl2nc>. The observational data used in the case studies are available upon request. The reanalyses data used in the case studies are publicly available online from the respective projects. The *Unified Model* data used in the case studies are available upon request. The Unified Model is proprietary to the UK Met Office and is made available under a licence. For more information, readers are advised to contact the UK Met Office.

*Author contributions.* Peter Kuma wrote the code of the framework, performed the data analysis of the case studies and wrote the text of the manuscript. Adrian McDonald and Olaf Morgenstern provided continuous scientific input on the code development, analysis and text of the manuscript. Richard Querel, Israel Silber and Connor J. Flynn provided calibration of the MiniMPL data and substantial discussion of the theoretical concepts. All authors reviewed the manuscript.

*Competing interests.* The authors declare that they have no competing interests.

*Acknowledgements.* We would like to acknowledge the New Zealand Deep South National Science Challenge project which provided funding for our work; the New Zealand eScience Infrastructure (NeSI) which provided supercomputing resources to run the Unified Model; Vidya Varma, Jonny Williams, Guang Zeng and Wolfgang Hayek for their contribution to setting up a nudged run of the Unified Model; Graeme Plank and Graeme MacDonald who participated on the installation of the Vaisala CL51 at the Cass field station; the COSP project for the code which we used as the basis for the lidar simulator; the AMPS, JRA-55, ERA5, MERRA-2 models and reanalyses which provided public access to their data; the open source libraries numpy (Van Der Walt et al., 2011), scipy (Virtanen et al., 2019), matplotlib (Hunter, 2007), netCDF4 (Rew and Davis, 1990) and Astropy (Price-Whelan et al., 2018) and the Python programming language (Rossum, 1995) which we used in the implementation of our code; the R programming language (R Core Team, 2017), the Natural Earth dataset (<https://www.naturalearthdata.com>) and the Shuttle Radar Topography Mission (SRTM) Version 3 Global 1 arc second digital elevation model (Werner, 2001; NASA JPL, 2013) which we used to produce a map of sites; GitHub which provided free hosting of our code; and the Linux-based (Torvalds, 1997) operating systems Devuan GNU+Linux and Debian GNU/Linux on which we produced this analysis.

~~(a) Scheme showing the operation of the ALCF and (b) the processing commands.~~

~~Map showing the location of sites. Data at three sites in New Zealand were analysed: Cass, Lauder and Christchurch.~~

~~(a) Theoretical distributions of cloud droplet radius based on the log-normal and Gamma distributions parametrised by multiple choices of the effective radius  $r_{\text{eff}}$  and effective standard deviation  $\sigma_{\text{eff}}$ . (b) Lidar ratio (LR) as a function of effective radius calculated for different theoretical cloud droplet size distributions, laser wavelengths and effective standard deviation ratios.~~

Cloud-occurrence-histogram-as-a-function-of-height-above-the-mean-sea-level-observed-at-three-sites-and-simulated-by-the-lidar-simulator-based-on-atmospheric-fields-for-five-reanalyses-and-models. Shown is also the total cloud fraction (CF). The histogram is calculated from the cloud mask as determined by the cloud detection algorithm.

Examples of observed (OBS) and simulated backscatter profiles during 24 hours taken from the three case studies presented here. The observed backscatter was normalised to absolute units and denoised. The vertical levels were interpolated based on the nearest neighbour assumption. The cloud detection threshold of 2 corresponds to the low end of the colour scale, i.e. all visible backscatter in the plots was determined as cloud, unless the difference from the threshold was smaller than three noise standard deviations. The first subcolumn of 10 columns generated by the Subgrid Cloud Overlap Profile Sampler (SCOPS) was selected to make the plots. The red line signifies the station altitude.

10 The same as Fig. 4 but for the Lauder site.

The same as Fig. 4 but for the Christchurch site.

(a) Scatter plot showing the average cloud fraction calculated from all profiles and the average inverse lidar ratio ( $LR^{-1}$ ) calculated from cloudy profiles for the case studies. The site and instrument are coded by the shape of the marker and the observation (legend in the top left corner), reanalysis or model is coded by the colour of the marker (legend in the top right corner). (b) The same as (a), but showing relative differences between the reanalyses and models relative to the corresponding observations.

Backscatter histograms as a function of height observed and simulated at three different sites of the case studies calculated from all profiles. The plots show the distribution of backscatter for values which are on the scale of noise, molecular and aerosol backscatter (-0.5, 0.5 for CHM-15k, -1, 1 for CL31 and CL51 and -2, 2 for MiniMPL). The simulated backscatter is based on the ERA5 atmospheric fields. Visible in the plots is backscatter caused by molecular backscatter (the main "streak"), noise when signal is fully attenuated by cloud (the zero-centred "streak"), and the range-dependent noise (the zero-centred "cone"). The molecular backscatter is marked by a red dashed line on the observed backscatter plots, the shape of which is taken from the simulated molecular backscatter for the corresponding instrument and site.

The same as Fig. 7 but calculated from clear sky profiles only.

25 Backscatter noise standard deviation histogram calculated for each instrument at sites of the case studies from clear sky profiles over the whole time period. The noise distribution is calculated at the furthest range. Shown is the range-scaled noise distribution at a range of 7 km. The MiniMPL nighttime and daytime distributions are calculated from nighttime and daytime profiles only, determined by position of the Sun below and above the horizon, respectively, at the geographical position of the site and time when the profile was observed.

30 Table of ALCs and their technical parameters. Power is calculated as Pulse  $\times$  Pulse Repetition Frequency (PRF). Instrument  $\lambda$  (nm) Laser Rate(s) Res.(m) Depol.Pulse( $\mu$ J) Range(km) PRF (kHz) Overlap(m) Power (mW) CHM-15k-1064 Nd:YAG 2-600 5 no 7-9 15.4 5-7 1000 48 CL31-910 InGaAs 2-120 10 no 1.2 7.7 10 70 12 CL51-910 InGaAs 6-120 10 no 3 15.4 6.5 23020 MiniMPL-532 Nd:YAG 1-900 5-75 yes 3-4 30.0 2.5 20009

35 Reanalyses and models used in the case studies and some of their main properties. The temporal and horizontal grid resolution and vertical levels listed is the resolution of the model output available. The horizontal grid resolution is determined

at 45S. The internal resolution of the model may be different (see Sect. 3.2 for details). The reanalyses and the UM use regular longitude-latitude grids, while the AMPS horizontal grid is regular in the South Pole stereographic projection. Model/Grid Type Time resolution Horizontal grid resolution Vertical levels AMPS/D01 NWP 3 h 0.27 × 0.19 (21 × 21 km) 60 ERA5 Reanalysis 1 h 0.25 × 0.25 (20 × 28 km) 37 JRA-55 Reanalysis 6 h 1.25 × 1.25 (98 × 139 km) 37 MERRA-2 Reanalysis 3 h 0.625 × 0.50 (49 × 56 km) 72 UM (GA7.1)/N96 GCM 20 min. 1.875 × 1.25 (147 × 139 km) 85

Location of sites and instruments. The time periods are inclusive. Site Coordinates Surface altitude (m) Instruments Time period Missing period Days Cass 43.0346°S 171.7594°E 577 CL51 19 Sep–1 Oct 2014 13 Lauder 45.0379°S 169.6831°E 370 MiniMPL, CL31 12–24 Jan 2018 13 Christchurch 43.5225°S 172.5841°E 45 MiniMPL, CHM 15k 17 July–19 August 2019 22–31 July 24

10 Table of physical quantities. Symbol Name Units Expression  $\alpha_s$  ( $\alpha_e$ ) Volume scattering (extinction) coefficient  $\text{m}^{-1}\Omega$  Solid angle  $\text{sr}$   $P_\pi(\theta)$  Scattering phase function at angle  $\theta$   $1 \int_{4\pi} P_\pi(\theta) d\Omega = 4\pi\beta$  Volume backscattering coefficient  $\text{m}^{-1}\text{sr}^{-1}$   $\beta = \alpha_s P_\pi(\pi)/(4\pi)N$  Particle number concentration  $\text{m}^{-3}$   $n(r)$  Number distribution of particle size  $\text{m}^{-4}$   $N = \int_0^\infty n(r) dr$   $Q_s$  ( $Q_e$ ) Scattering (extinction) efficiency of spherical particles  $1 - \alpha_s = Q_s \pi r^2 N$ ,  $\alpha_e = Q_e \pi r^2 N$   $Q_b$  Backscattering efficiency of spherical particles  $\text{sr}^{-1}$   $\beta = Q_b \pi r^2 N$   $S$  Lidar ratio (extinction-to-backscatter ratio)  $\text{sr}$   $S = \alpha_e/\beta k$  Backscatter-to-extinction ratio  $\text{sr}^{-1}$   $k = 1/S r_{\text{eff}}$  Effective radius  $\text{m}$   $r_{\text{eff}} = \int_0^\infty r^3 n(r) dr / \int_0^\infty r^2 n(r) dr$   $\sigma_{\text{eff}}$  Effective standard deviation  $\text{m}$   $\sigma_{\text{eff}} = (\int_0^\infty (r - r_{\text{eff}})^2 r^2 n(r) dr / \int_0^\infty r^2 n(r) dr)^{1/2}$  Boltzmann constant  $k_B \approx 1.38 \times 10^{-23}$   $p$  Atmospheric pressure Pa  $T$  Atmospheric temperature K  $\rho$  Liquid (or ice) density  $\rho_{\text{air}}$  Air density  $q$  Cloud liquid (or ice) mass mixing ratio 1

Table of sensitivity tests of theoretical distribution assumption, effective radius  $r_{\text{eff}}$  and effective standard deviation  $\sigma_{\text{eff}}$  of the cloud droplet/ice crystal size distribution.  $\mu$  and  $\sigma$  are the mean and standard deviation of a normal distribution corresponding to the log-normal distribution, calculated numerically from  $r_{\text{eff}}$  and  $\sigma_{\text{eff}}$ .  $\mu_*$  and  $\sigma_*$  are the actual mean and standard deviation of the distribution (calculated numerically). Distribution  $r_{\text{eff}}$  ( $\mu\text{m}$ )  $\sigma_{\text{eff}}$  ( $\mu\text{m}$ )  $\mu$   $\sigma$   $\mu_*$  ( $\mu\text{m}$ )  $\sigma_*$  ( $\mu\text{m}$ ) log-normal 20 10 2.44 0.47 12.76 6.26 log-normal 20 5 2.84 0.25 17.72 4.43 log-normal 10 5 1.74 0.47 6.40 3.20 Gamma 20 10 9.98 7.00 Gamma 20 5 17.50 4.68 Gamma 10 5 5.00 3.54

Theoretical molecular backscatter calculated at pressure 1000 hPa and temperature 20 and the calibration coefficient, relative to the instrument native units, determined for the instrument based on the molecular backscatter and stratocumulus lidar ratio calibration methods. Instrument Wavelength (nm) Molecular backscatter (–) Calibration coefficient CHM 15k 1064 0.0906 0.34 CL31 910 0.172 1.45 CL51 910 0.172 1.2 MiniMPL 532 1.54 3.75

## References

- G-SDSU, <https://cloud.gsfc.nasa.gov/index.php?section=14>, 2019.
- Baars, H., Kanitz, T., Engelmann, R., Althausen, D., Heese, B., Komppula, M., Preißler, J., Tesche, M., Ansmann, A., Wandinger, U., Lim, J.-H., Ahn, J. Y., Stachlewska, I. S., Amiridis, V., Marinou, E., Seifert, P., Hofer, J., Skupin, A., Schneider, F., Bohlmann, S., Foth, A., Bley, S., Pfüller, A., Giannakaki, E., Lihavainen, H., Viisanen, Y., Hooda, R. K., Pereira, S. N., Bortoli, D., Wagner, F., Mattis, I., Janicka, L., Markowicz, K. M., Achtert, P., Artaxo, P., Pauliquevis, T., Souza, R. A. F., Sharma, V. P., van Zyl, P. G., Beukes, J. P., Sun, J., Rohwer, E. G., Deng, R., Mamouri, R.-E., and Zamorano, F.: An overview of the first decade of PollyNET: an emerging network of automated Raman-polarization lidars for continuous aerosol profiling, *Atmospheric Chemistry and Physics*, 16, 5111–5137, <https://doi.org/10.5194/acp-16-5111-2016>, 2016.
- 5
- 10 Baran, A. J.: A review of the light scattering properties of cirrus, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110, 1239–1260, <https://doi.org/10.1016/j.jqsrt.2009.02.026>, 2009.
- Bastin, S., Chiriaco, M., and Drobinski, P.: Control of radiation and evaporation on temperature variability in a WRF regional climate simulation: comparison with colocated long term ground based observations near Paris, *Climate dynamics*, 51, 985–1003, <https://doi.org/10.1007/s00382-016-2974-1>, 2018.
- 15 Bi, L., Yang, P., Kattawar, G. W., Baum, B. A., Hu, Y. X., Winker, D. M., Brock, R. S., and Lu, J. Q.: Simulation of the color ratio associated with the backscattering of radiation by ice particles at the wavelengths of 0.532 and 1.064  $\mu\text{m}$ , *Journal of Geophysical Research*, 114, <https://doi.org/10.1029/2009jd011759>, 2009.
- Bodas-Salcedo, A., Webb, M., Bony, S., Chepfer, H., Dufresne, J.-L., Klein, S., Zhang, Y., Marchand, R., Haynes, J., Pincus, R., et al.: COSP: Satellite simulation software for model assessment, *Bulletin of the American Meteorological Society*, 92, 1023–1043, <https://doi.org/10.1175/2011BAMS2856.1>, 2011.
- 20 Borovoi, A., Konoshonkin, A., and Kustova, N.: Backscatter ratios for arbitrary oriented hexagonal ice crystals of cirrus clouds, *Opt. Lett.*, 39, 5788–5791, <https://doi.org/10.1364/OL.39.005788>, 2014.
- Bouniol, D., Protat, A., Delanoë, J., Pelon, J., Piriou, J.-M., Bouysse, F., Tompkins, A. M., Wilson, D. R., Morille, Y., Haefelin, M., et al.: Using continuous ground-based radar and lidar measurements for evaluating the representation of clouds in four operational models, *Journal of Applied Meteorology and Climatology*, 49, 1971–1991, <https://doi.org/10.1175/2010JAMC2333.1>, 2010.
- 25 Bréon, F.-M. and Doutriaux-Boucher, M.: A comparison of cloud droplet radii measured from space, *IEEE Transactions on Geoscience and Remote Sensing*, 43, 1796–1805, <https://doi.org/10.1109/TGRS.2005.852838>, 2005.
- Bréon, F.-M. and Colzy, S.: Global distribution of cloud droplet effective radius from POLDER polarization measurements, *Geophysical Research Letters*, 27, 4065–4068, <https://doi.org/10.1029/2000GL011691>, 2000.
- 30 Campbell, J. R., Hlavka, D. L., Welton, E. J., Flynn, C. J., Turner, D. D., Spinhirne, J. D., Scott III, V. S., and Hwang, I.: Full-time, eye-safe cloud and aerosol lidar observation at atmospheric radiation measurement program sites: Instruments and data processing, *Journal of Atmospheric and Oceanic Technology*, 19, 431–442, [https://doi.org/10.1175/1520-0426\(2002\)019<0431:FTESCA>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<0431:FTESCA>2.0.CO;2), 2002.
- Cazorla, A., Casquero-Vera, J. A., Román, R., Guerrero-Rascado, J. L., Toledano, C., Cachorro, V. E., Orza, J. A. G., Cancillo, M. L., Serrano, A., Titos, G., Pandolfi, M., Alastuey, A., Hanrieder, N., and Alados-Arboledas, L.: Near-real-time processing of a ceilometer network assisted with sun-photometer data: monitoring a dust outbreak over the Iberian Peninsula, *Atmospheric Chemistry and Physics*, 17, 11 861–11 876, <https://doi.org/10.5194/acp-17-11861-2017>, 2017.
- 35



- Chan, K. L., Wiegner, M., Flentje, H., Mattis, I., Wagner, F., Gasteiger, J., and Geiß, A.: Evaluation of ECMWF-IFS (version 41R1) operational model forecasts of aerosol transport by using ceilometer network measurements, *Geoscientific Model Development*, 11, 3807–3831, <https://doi.org/10.5194/gmd-11-3807-2018>, 2018.
- Chang, F. and Li, Z.: The effect of droplet size distribution on the determination of cloud droplet effective radius, in: 11th ARM Science Team Meeting, Atlanta, Ga, pp. 19–23, 2001.
- Chepfer, H., Chiriaco, M., Vautard, R., and Spinhirne, J.: Evaluation of MM5 optically thin clouds over Europe in fall using ICESat lidar spaceborne observations, *Monthly weather review*, 135, 2737–2753, <https://doi.org/10.1175/MWR3413.1>, 2007.
- Chepfer, H., Bony, S., Winker, D., Chiriaco, M., Dufresne, J.-L., and Sèze, G.: Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model, *Geophysical Research Letters*, 35, <https://doi.org/10.1029/2008GL034207>, 2008.
- Chiriaco, M., Vautard, R., Chepfer, H., Haefelin, M., Dudhia, J., Wanherdrick, Y., Morille, Y., and Protat, A.: The ability of MM5 to simulate ice clouds: Systematic comparison between simulated and measured fluxes and lidar/radar profiles at the SIRTa atmospheric observatory, *Monthly weather review*, 134, 897–918, <https://doi.org/10.1175/MWR3102.1>, 2006.
- Chiriaco, M., Dupont, J.-C., Bastin, S., Badosa, J., Lopez, J., Haefelin, M., Chepfer, H., and Guzman, R.: ReOBS: a new approach to synthesize long-term multi-variable dataset and application to the SIRTa supersite, *Earth System Science Data*, 10, 919, <https://doi.org/10.5194/essd-10-919-2018>, 2018.
- Costa-Surós, M., Calbó, J., González, J., and Martin-Vide, J.: Behavior of cloud base height from ceilometer measurements, *Atmospheric Research*, 127, 64–76, <https://doi.org/10.1016/j.atmosres.2013.02.005>, 2013.
- Cromwell, E. and Flynn, D.: Lidar Cloud Detection with Fully Convolutional Networks, 2018.
- Cromwell, E. and Flynn, D.: Lidar cloud detection with fully convolutional networks, in: 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 619–627, IEEE, 2019.
- Dee, D. P., Uppala, S. M., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, d. P., et al.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Quarterly Journal of the royal meteorological society*, 137, 553–597, <https://doi.org/10.1002/qj.828>, 2011.
- Diner, D. J., Beckert, J. C., Reilly, T. H., Bruegge, C. J., Conel, J. E., Kahn, R. A., Martonchik, J. V., Ackerman, T. P., Davies, R., Gerstl, S. A., et al.: Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview, *IEEE Transactions on Geoscience and Remote Sensing*, 36, 1072–1087, <https://doi.org/10.1109/36.700992>, 1998.
- Dionisi, D., Barnaba, F., Diémoz, H., Di Liberto, L., and Gobbi, G. P.: A multiwavelength numerical model in support of quantitative retrievals of aerosol properties from automated lidar ceilometers and test applications for AOT and PM 10 estimation, *Atmospheric Measurement Techniques*, 11, 6013, <https://doi.org/10.5194/amt-11-6013-2018>, 2018.
- Ebita, A., Kobayashi, S., Ota, Y., Moriya, M., Kumabe, R., Onogi, K., Harada, Y., Yasui, S., Miyaoka, K., Takahashi, K., et al.: The Japanese 55-year reanalysis “JRA-55”: an interim report, *Sola*, 7, 149–152, <https://doi.org/10.2151/jmsj.2016-015>, 2011.
- ECMWF: Copernicus Climate Change Service (C3S) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate, Copernicus Climate Change Service Climate Data Store (CDS), <https://doi.org/10.24381/cds.bd0915c6>, <https://cds.climate.copernicus.eu/cdsapp#!/home>, 2019.
- Edwards, J. and Slingo, A.: Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model, *Quarterly Journal of the Royal Meteorological Society*, 122, 689–719, <https://doi.org/10.1002/qj.49712253107>, 1996.
- Emeis, S.: Surface-based remote sensing of the atmospheric boundary layer, vol. 40, Springer Science & Business Media, <https://doi.org/10.1007/978-90-481-9340-0>, 2010.

- Emeis, S., Schäfer, K., and Munkel, C.: Observation of the structure of the urban boundary layer with different ceilometers and validation by RASS data, *Meteorologische Zeitschrift*, 18, 149–154, <https://doi.org/10.1127/0941-2948/2009/0365>, 2009.
- Eresmaa, N., Karppinen, A., Joffe, S., Räsänen, J., and Talvitie, H.: Mixing height determination by ceilometer, *Atmospheric Chemistry and Physics*, 6, 1485–1493, <https://doi.org/10.5194/acp-6-1485-2006>, 2006.
- 5 Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geoscientific Model Development*, 9, 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016>, 2016.
- Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., Collins, W. D., Gier, B. K., Hall, A. D., Hoffman, F. M., et al.: Taking climate model evaluation to the next level, *Nature Climate Change*, p. 1, <https://doi.org/10.1038/s41558-018-0355-y>, 2019.
- 10 Flynn, C. J., Mendozaa, A., Zhengb, Y., and Mathurb, S.: Novel polarization-sensitive micropulse lidar measurement technique, *Optics express*, 15, 2785–2790, <https://doi.org/10.1364/OE.15.002785>, 2007.
- Fu, D., Di Girolamo, L., Liang, L., and Zhao, G.: Regional Biases in MODIS Marine Liquid Water Cloud Drop Effective Radius Deduced Through Fusion With MISR, *Journal of Geophysical Research: Atmospheres*, 124, 13 182–13 196, <https://doi.org/10.1029/2019JD031063>, 2019.
- 15 Garnier, A., Pelon, J., Vaughan, M. A., Winker, D. M., Treppe, C. R., and Dubuisson, P.: Lidar multiple scattering factors inferred from CALIPSO lidar and IIR retrievals of semi-transparent cirrus cloud optical depths over oceans, *Atmospheric Measurement Techniques*, 8, 2759–2774, <https://doi.org/10.5194/amt-8-2759-2015>, 2015.
- Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., et al.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), *Journal of Climate*, 30, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>, 2017.
- 20 Geleyn, J. and Hollingsworth, A.: An economical analytical method for the computation of the interaction between scattering and line absorption of radiation, *Contributions to Atmospheric Physics*, 52, 1979.
- Goody, R. M. and Yung, Y. L.: *Atmospheric radiation: theoretical basis*, Oxford University Press, 2 edn., 1995.
- Hansen, A., Ament, F., Grützun, V., and Lammert, A.: Model evaluation by a cloud classification based on multi-sensor observations, *Geoscientific Model Development Discussions*, 2018, 1–20, <https://doi.org/10.5194/gmd-2018-259>, 2018a.
- 25 Hansen, A., Ament, F., Grützun, V., and Lammert, A.: Model evaluation by a cloud classification based on multi-sensor observations, *Geoscientific Model Development Discussions*, 2018, 1–20, <https://doi.org/10.5194/gmd-2018-259>, 2018b.
- Harada, Y., Kamahori, H., Kobayashi, C., Endo, H., Kobayashi, S., Ota, Y., Onoda, H., Onogi, K., Miyaoka, K., and Takahashi, K.: The JRA-55 Reanalysis: Representation of atmospheric circulation and climate variability, *Journal of the Meteorological Society of Japan*. Ser. II, 94, 269–302, <https://doi.org/10.2151/jmsj.2016-015>, 2016.
- 30 Heese, B., Flentje, H., Althausen, D., Ansmann, A., and Frey, S.: Ceilometer lidar comparison: backscatter coefficient retrieval and signal-to-noise ratio determination, *Atmospheric Measurement Techniques*, 3, 1763–1770, <https://doi.org/10.5194/amt-3-1763-2010>, 2010.
- Heymsfield, A. J.: Extinction-ice water content-effective radius algorithms for CALIPSO, *Geophysical Research Letters*, 32, <https://doi.org/10.1029/2005gl022742>, 2005.
- 35 Hines, K. M. and Bromwich, D. H.: Development and testing of Polar Weather Research and Forecasting (WRF) model. Part I: Greenland ice sheet meteorology, *Monthly Weather Review*, 136, 1971–1989, <https://doi.org/10.1175/2007MWR2112.1>, 2008.
- Hogan, R. J.: Fast approximate calculation of multiply scattered lidar returns, *Applied Optics*, 45, 5984–5992, <https://doi.org/10.1364/AO.45.005984>, 2006.

- Hogan, R. J., Jakob, C., and Illingworth, A. J.: Comparison of ECMWF Winter-Season Cloud Fraction with Radar-Derived Values, *Journal of Applied Meteorology*, 40, 513–525, [https://doi.org/10.1175/1520-0450\(2001\)040<0513:COEWSC>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<0513:COEWSC>2.0.CO;2), 2001.
- Hopkin, E.: Use of a calibrated ceilometer network to improve high resolution weather forecasts, Ph.D. thesis, University of Reading, UK, 2018.
- 5 Hopkin, E., Illingworth, A. J., Charlton-Perez, C., Westbrook, C. D., and Ballard, S.: A robust automated technique for operational calibration of ceilometers using the integrated backscatter from totally attenuating liquid clouds, *Atmospheric Measurement Techniques Discussions*, 2019, 1–28, <https://doi.org/10.5194/amt-2018-427>, 2019.
- Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., Folini, D., Ji, D., Klocke, D., Qian, Y., et al.: The art and science of climate model tuning, *Bulletin of the American Meteorological Society*, 98, 589–602, <https://doi.org/10.1175/BAMS-D-15-00135.1>,  
10 2017.
- Hu, Y.: Depolarization ratio–effective lidar ratio relation: Theoretical basis for space lidar cloud phase discrimination, *Geophysical Research Letters*, 34, <https://doi.org/10.1029/2007GL029584>, 2007.
- Hu, Y., Vaughan, M., McClain, C., Behrenfeld, M., Maring, H., Anderson, D., Sun-Mack, S., Flittner, D., Huang, J., Wielicki, B., Minnis, P., Weimer, C., Trepte, C., and Kuehn, R.: Global statistics of liquid water content and effective number concentration of water  
15 clouds over ocean derived from combined CALIPSO and MODIS measurements, *Atmospheric Chemistry and Physics*, 7, 3353–3359, <https://doi.org/10.5194/acp-7-3353-2007>, 2007.
- Hunter, J. D.: Matplotlib: A 2D graphics environment, *Computing in science & engineering*, 9, 90, <https://doi.org/10.1109/MCSE.2007.55>, 2007.
- Illingworth, A., Hogan, R., O’connor, E., Bouniol, D., Brooks, M., Delanoë, J., Donovan, D., Eastment, J., Gaussiat, N., Goddard, J., et al.:  
20 Cloudnet: Continuous evaluation of cloud profiles in seven operational models using ground-based observations, *Bulletin of the American Meteorological Society*, 88, 883–898, <https://doi.org/10.1175/BAMS-88-6-883>, 2007.
- Illingworth, A., Cimini, D., Haeffele, A., Haeffelin, M., Hervo, M., Kotthaus, S., Löhnert, U., Martinet, P., Mattis, I., O’Connor, E., et al.: How can Existing Ground-Based Profiling Instruments Improve European Weather Forecasts?, *Bulletin of the American Meteorological Society*, <https://doi.org/10.1175/BAMS-D-17-0231.1>, 2018.
- 25 Illingworth, A. J., Barker, H., Beljaars, A., Ceccaldi, M., Chepfer, H., Clerbaux, N., Cole, J., Delanoë, J., Domenech, C., Donovan, D. P., et al.: The EarthCARE satellite: The next step forward in global measurements of clouds, aerosols, precipitation, and radiation, *Bulletin of the American Meteorological Society*, 96, 1311–1332, <https://doi.org/10.1175/BAMS-D-12-00227.1>, 2015a.
- Illingworth, A. J., Cimini, D., Gaffard, C., Haeffelin, M., Lehmann, V., Löhnert, U., O’Connor, E. J., and Ruffieux, D.: Exploiting existing ground-based remote sensing networks to improve high-resolution weather forecasts, *Bulletin of the American Meteorological Society*,  
30 96, 2107–2125, <https://doi.org/10.1175/BAMS-D-13-00283.1>, 2015b.
- Jin, Y., Kai, K., Kawai, K., Nagai, T., Sakai, T., Yamazaki, A., Uchiyama, A., Batdorj, D., Sugimoto, N., and Nishizawa, T.: Ceilometer calibration for retrieval of aerosol optical properties, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 153, 49–56, <https://doi.org/10.1016/j.jqsrt.2014.10.009>, 2015.
- Josset, D., Pelon, J., Garnier, A., Hu, Y., Vaughan, M., Zhai, P.-W., Kuehn, R., and Lucker, P.: Cirrus optical depth and lidar ratio retrieval  
35 from combined CALIPSO-CloudSat observations using ocean surface echo, *Journal of Geophysical Research: Atmospheres*, 117, n/a–n/a, <https://doi.org/10.1029/2011jd016959>, 2012.
- Klein, S. A. and Jakob, C.: Validation and sensitivities of frontal clouds simulated by the ECMWF model, *Monthly Weather Review*, 127, 2514–2531, [https://doi.org/10.1175/1520-0493\(1999\)127<2514:VASOFC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<2514:VASOFC>2.0.CO;2), 1999.

- Klein, S. A., Zhang, Y., Zelinka, M. D., Pincus, R., Boyle, J., and Gleckler, P. J.: Are climate model simulations of clouds improving? An evaluation using the ISCCP simulator, *Journal of Geophysical Research: Atmospheres*, 118, 1329–1342, <https://doi.org/10.1002/jgrd.50141>, 2013.
- 5 Klekociuk, A. R., French, W. J. R., Alexander, S. P., Kuma, P., and McDonald, A. J.: The state of the atmosphere in the 2016 southern Kerguelen Axis campaign region, *Deep Sea Research Part II: Topical Studies in Oceanography*, <https://doi.org/10.1016/j.dsr2.2019.02.001>, 2019.
- Knepp, T. N., Szykman, J. J., Long, R., Duvall, R. M., Krug, J., Beaver, M., Cavender, K., Kronmiller, K., Wheeler, M., Delgado, R., Hoff, R., Berkoff, T., Olson, E., Clark, R., Wolfe, D., Van Gilst, D., and Neil, D.: Assessment of mixed-layer height estimation from single-wavelength ceilometer profiles, *Atmospheric Measurement Techniques*, 10, 3963–3983, <https://doi.org/10.5194/amt-10-3963-2017>, 2017.
- 10 Kobayashi, S., Ota, Y., Harada, Y., Ebata, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., et al.: The JRA-55 reanalysis: General specifications and basic characteristics, *Journal of the Meteorological Society of Japan. Ser. II*, 93, 5–48, <https://doi.org/10.2151/jmsj.2015-001>, 2015.
- Kotthaus, S., O’Connor, E., Munkel, C., Charlton-Perez, C., Haeffelin, M., Gabey, A. M., and Grimmond, C. S. B.: Recommendations for processing atmospheric attenuated backscatter profiles from Vaisala CL31 ceilometers, *Atmospheric Measurement Techniques*, 9, 3769–3791, <https://doi.org/10.5194/amt-9-3769-2016>, 2016.
- 15 Kuma, P., McDonald, A., Morgenstern, O., Alexander, S. P., Cassano, J. J., Garrett, S., Halla, J., Hartery, S., Harvey, M. J., Parsons, S., Plank, G., Varma, V., and Williams, J.: Evaluation of Southern Ocean cloud in the HadGEM3 general circulation model and MERRA-2 reanalysis using ship-based observations, *Atmospheric Chemistry and Physics*, 20, 6607–6630, <https://doi.org/10.5194/acp-20-6607-2020>, 2020.
- Lamer, K., Fridlind, A. M., Ackerman, A. S., Kollias, P., Clothiaux, E. E., and Kelley, M.: (GO)2-SIM: a GCM-oriented ground-observation forward-simulator framework for objective evaluation of cloud and precipitation phase, *Geoscientific Model Development (Online)*, 11, <https://doi.org/10.5194/gmd-11-4195-2018>, 2018.
- 20 Lewis, J. R., Campbell, J. R., Welton, E. J., Stewart, S. A., and Haftings, P. C.: Overview of MPLNET version 3 cloud detection, *Journal of Atmospheric and Oceanic Technology*, 33, 2113–2134, <https://doi.org/10.1175/JTECH-D-15-0190.1>, 2016.
- Liou, K.-N.: An introduction to atmospheric radiation, vol. 84, Elsevier, 2 edn., 2002.
- 25 Liu, J., Li, Z., Zheng, Y., and Cribb, M.: Cloud-base distribution and cirrus properties based on micropulse lidar measurements at a site in southeastern China, *Advances in Atmospheric Sciences*, 32, 991–1004, <https://doi.org/10.1007/s00376-014-4176-2>, 2015a.
- Liu, L., Sun, X.-j., Liu, X.-c., Gao, T.-c., and Zhao, S.-j.: Comparison of cloud base height derived from a ground-based infrared cloud measurement and two ceilometers, *Advances in Meteorology*, 2015, <https://doi.org/10.1155/2015/853861>, 2015b.
- Madonna, F., Rosoldi, M., Lolli, S., Amato, F., Vande Hey, J., Dhillon, R., Zheng, Y., Brettle, M., and Pappalardo, G.: Intercomparison of aerosol measurements performed with multi-wavelength Raman lidars, automatic lidars and ceilometers in the framework of INTERACT-II campaign, *Atmospheric Measurement Techniques*, 11, 2459–2475, <https://doi.org/10.5194/amt-11-2459-2018>, 2018.
- 30 Marengo, F., Santacesaria, V., Bais, A. F., Balis, D., di Sarra, A., Papayannis, A., and Zerefos, C.: Optical properties of tropospheric aerosols determined by lidar and spectrophotometric measurement (Photochemical Activity and Solar Ultraviolet Radiation campaign), *Applied Optics*, 36, 6875–6886, <https://doi.org/10.1364/AO.36.006875>, 1997.
- 35 Martucci, G., Milroy, C., and O’Dowd, C. D.: Detection of cloud-base height using Jenoptik CHM15K and Vaisala CL31 ceilometers, *Journal of Atmospheric and Oceanic Technology*, 27, 305–318, <https://doi.org/10.1175/2009JTECHA1326.1>, 2010.

- Masunaga, H., Matsui, T., Tao, W.-k., Hou, A. Y., Kummerow, C. D., Nakajima, T., Bauer, P., Olson, W. S., Sekiguchi, M., and Nakajima, T. Y.: Satellite data simulator unit: A multisensor, multispectral satellite simulator package, *Bulletin of the American Meteorological Society*, 91, 1625–1632, <https://doi.org/10.1175/2010BAMS2809.1>, 2010.
- 5 Mattis, I., Begbie, R., Boyouk, N., Bravo-Aranda, J. A., Brettle, M., Cermak, J., Drouin, M.-A., Geiß, A., Görsdorf, U., Haefele, A., Haeffelin, M., Hervo, M., Komínková, K., Leinweber, R., Müller, G., Münkkel, C., Pattantyús-Ábrahám, M., Pönitz, K., Wagner, F., and Wiegner, M.: The ceilometer inter-comparison campaign CeiLinEx2015, in: *EGU General Assembly Conference Abstracts*, pp. EPSC2016–9687, 2016.
- McGill, M. J., Yorks, J. E., Scott, V. S., Kupchock, A. W., and Selmer, P. A.: The Cloud-Aerosol Transport System (CATS): A technology demonstration on the International Space Station, in: *Lidar Remote Sensing for Environmental Monitoring XV*, vol. 9612, p. 96120A, 10 International Society for Optics and Photonics, 2015.
- Mie, G.: Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen, *Annalen der physik*, 330, 377–445, <https://doi.org/10.1002/andp.19083300302>, 1908.
- Milroy, C., Martucci, G., Lolli, S., Loaec, S., Sauvage, L., Xueref-Remy, I., Lavrič, J. V., Ciais, P., Feist, D. G., Biavati, G., et al.: An assessment of pseudo-operational ground-based light detection and ranging sensors to determine the boundary-layer structure in the 15 coastal atmosphere, *Advances in Meteorology*, 2012, <https://doi.org/10.1155/2012/929080>, 2012.
- Morcrette, C. J., O’Connor, E. J., and Petch, J. C.: Evaluation of two cloud parametrization schemes using ARM and Cloud-Net observations, *Quarterly Journal of the Royal Meteorological Society*, 138, 964–979, <https://doi.org/10.1002/qj.969>, 2012.
- Morille, Y., Haeffelin, M., Drobinski, P., and Pelon, J.: STRAT: An automated algorithm to retrieve the vertical structure of the atmosphere from single-channel lidar data, *Journal of Atmospheric and Oceanic Technology*, 24, 761–775, <https://doi.org/10.1175/JTECH2008.1>, 20 2007.
- Münkkel, C., Eresmaa, N., Räsänen, J., and Karppinen, A.: Retrieval of mixing height and dust concentration with lidar ceilometer, *Boundary-layer meteorology*, 124, 117–128, <https://doi.org/10.1007/s10546-006-9103-3>, 2007.
- Nam, C., Bony, S., Dufresne, J.-L., and Chepfer, H.: The ‘too few, too bright’ tropical low-cloud problem in CMIP5 models, *Geophysical Research Letters*, 39, <https://doi.org/10.1029/2012GL053421>, 2012.
- 25 NASA JPL: NASA Shuttle Radar Topography Mission Global 3 arc second [Data set], NASA EOSDIS Land Processes DAAC, <https://doi.org/https://doi.org/10.5067/MEaSURES/SRTM/SRTMGL3.003>, 2013.
- Noel, V. and Chepfer, H.: A global view of horizontally oriented crystals in ice clouds from Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO), *Journal of Geophysical Research*, 115, <https://doi.org/10.1029/2009jd012365>, 2010.
- O’Connor, E. J., Illingworth, A. J., and Hogan, R. J.: A technique for autocalibration of cloud lidar, *Journal of Atmospheric and Oceanic 30 Technology*, 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.
- Pal, S. R., Steinbrecht, W., and Carswell, A. I.: Automated method for lidar determination of cloud-base height and vertical extent, *Applied optics*, 31, 1488–1494, <https://doi.org/10.1364/AO.31.001488>, 1992.
- Pappalardo, G., Amodeo, A., Apituley, A., Comeron, A., Freudenthaler, V., Linné, H., Ansmann, A., Bösenberg, J., D’Amico, G., Mattis, I., Mona, L., Wandinger, U., Amiridis, V., Alados-Arboledas, L., Nicolae, D., and Wiegner, M.: EARLINET: towards an advanced sustainable 35 European aerosol lidar network, *Atmospheric Measurement Techniques*, 7, 2389–2409, <https://doi.org/10.5194/amt-7-2389-2014>, 2014.
- Parkinson, C. L.: Aqua: An Earth-observing satellite mission to examine water and other climate variables, *IEEE Transactions on Geoscience and Remote Sensing*, 41, 173–183, <https://doi.org/10.1109/TGRS.2002.808319>, 2003.
- Petty, G. W.: *A First Course in Atmospheric Radiation*, Sundog Publishing, 2 edn., 2006.

- Petty, G. W. and Huang, W.: The modified gamma size distribution applied to inhomogeneous and nonspherical particles: Key relationships and conversions, *Journal of the Atmospheric Sciences*, 68, 1460–1473, <https://doi.org/10.1175/2011JAS3645.1>, 2011.
- Powers, J. G., Monaghan, A. J., Cayette, A. M., Bromwich, D. H., Kuo, Y.-H., and Manning, K. W.: Real-Time Mesoscale Modeling Over Antarctica: The Antarctic Mesoscale Prediction System, *Bulletin of the American Meteorological Society*, 84, 1533–1546, <https://doi.org/10.1175/BAMS-84-11-1533>, 2003.
- Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., Conseil, S., Shupe, D. L., Craig, M. W., Dencheva, N., and et al.: The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package, *The Astronomical Journal*, 156, 123, <https://doi.org/10.3847/1538-3881/aabc4f>, 2018.
- R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/>, 2017.
- Rausch, J., Meyer, K., Bennartz, R., and Platnick, S.: Differences in liquid cloud droplet effective radius and number concentration estimates between MODIS collections 5.1 and 6 over global oceans, *Atmospheric Measurement Techniques*, 10, 2105–2116, <https://doi.org/10.5194/amt-10-2105-2017>, 2017.
- Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *Journal of Geophysical Research: Atmospheres*, 108, 4407, <https://doi.org/10.1029/2002JD002670>, 2003.
- Rew, R. and Davis, G.: NetCDF: an interface for scientific data access, *IEEE computer graphics and applications*, 10, 76–82, <https://doi.org/10.1109/38.56302>, 1990.
- Rosoldi, M., Madonna, F., Pappalardo, G., Hey, J. V., and Zheng, Y.: The lesson learnt during interact-I and INTERACT-II actris measurement campaigns, in: *EPJ Web of Conferences*, vol. 176, p. 11002, EDP Sciences, 2018.
- Rossow, W. B. and Schiffer, R. A.: ISCCP cloud data products, *Bulletin of the American Meteorological Society*, [https://doi.org/10.1175/1520-0477\(1991\)072<0002:ICDP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1991)072<0002:ICDP>2.0.CO;2), 1991.
- Rossum, G.: Python reference manual, 1995.
- Schmidt, G. A., Bader, D., Donner, L. J., Elsaesser, G. S., Golaz, J.-C., Hannay, C., Molod, A., Neale, R., and Saha, S.: Practice and philosophy of climate model tuning across six US modeling centers, *Geoscientific Model Development*, 10, 3207–3223, <https://doi.org/10.5194/gmd-10-3207-2017>, 2017.
- Silber, I., Verlinde, J., Eloranta, E. W., Flynn, C. J., and Flynn, D. M.: Polar liquid cloud base detection algorithms for high spectral resolution or micropulse lidar data, *Journal of Geophysical Research: Atmospheres*, 123, 4310–4322, <https://doi.org/10.1029/2017JD027840>, 2018.
- Spinhirne, J. D.: Micro pulse lidar, *IEEE Transactions on Geoscience and Remote Sensing*, 31, 48–55, 1993.
- Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., Illingworth, A. J., O'Connor, E. J., Rossow, W. B., Durden, S. L., et al.: The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation, *Bulletin of the American Meteorological Society*, 83, 1771–1790, <https://doi.org/10.1175/BAMS-83-12-1771>, 2002.
- Stokes, G. M. and Schwartz, S. E.: The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the cloud and radiation test bed, *Bulletin of the American Meteorological Society*, 75, 1201–1222, [https://doi.org/10.1175/1520-0477\(1994\)075<1201:TARMPP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1994)075<1201:TARMPP>2.0.CO;2), 1994.
- Swales, D. J., Pincus, R., and Bodas-Salcedo, A.: The Cloud Feedback Model Intercomparison Project Observational Simulator Package: Version 2, *Geoscientific Model Development*, 11, 77–81, <https://doi.org/10.5194/gmd-11-77-2018>, 2018.

- Telford, P. J., Braesicke, P., Morgenstern, O., and Pyle, J. A.: Technical Note: Description and assessment of a nudged version of the new dynamics Unified Model, *Atmospheric Chemistry and Physics*, 8, 1701–1712, <https://doi.org/10.5194/acp-8-1701-2008>, 2008.
- Torvalds, L.: *Linux: a portable operating system*, Master's thesis, University of Helsinki, 1997.
- 5 Tsaknakis, G., Papayannis, A., Kokkalis, P., Amiridis, V., Kambezidis, H., Mamouri, R., Georgoussis, G., and Avdikos, G.: Inter-comparison of lidar and ceilometer retrievals for aerosol and Planetary Boundary Layer profiling over Athens, Greece, *Atmospheric Measurement Techniques*, 4, 1261–1273, <https://doi.org/10.5194/amt-4-1261-2011>, 2011.
- Van Der Walt, S., Colbert, S. C., and Varoquaux, G.: The NumPy array: a structure for efficient numerical computation, *Computing in Science & Engineering*, 13, 22–30, <https://doi.org/10.1109/MCSE.2011.37>, 2011.
- van Diedenhoven, B.: Remote Sensing of Crystal Shapes in Ice Clouds, in: *Springer Series in Light Scattering*, pp. 197–250, Springer International Publishing, [https://doi.org/10.1007/978-3-319-70808-9\\_5](https://doi.org/10.1007/978-3-319-70808-9_5), 2017.
- 10 Van Tricht, K., Gorodetskaya, I., Lhermitte, S., Turner, D., Schween, J., and Van Lipzig, N.: An improved algorithm for polar cloud-base detection by ceilometer over the ice sheets, *Atmospheric Measurement Techniques*, 7, 1153–1167, <https://doi.org/10.5194/amt-7-1153-2014>, 2014.
- Vaughan, M. A., Liu, Z., McGill, M. J., Hu, Y., and Obland, M. D.: On the spectral dependence of backscatter from cirrus clouds: Assessing CALIOP's 1064 nm calibration assumptions using cloud physics lidar measurements, *Journal of Geophysical Research: Atmospheres*, 115, <https://doi.org/10.1029/2009JD013086>, 2010.
- 15 Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Jarrod Millman, K., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C., Polat, İ., Feng, Y., Moore, E. W., Vand erPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors: SciPy 1.0–Fundamental Algorithms for Scientific Computing in Python, arXiv e-prints, arXiv:1907.10121, 2019.
- 20 Wall, C. J., Hartmann, D. L., and Ma, P.-L.: Instantaneous linkages between clouds and large-scale meteorology over the Southern Ocean in observations and a climate model, *Journal of Climate*, 30, 9455–9474, <https://doi.org/10.1175/JCLI-D-17-0156.1>, 2017.
- Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whittall, M., Williams, K., and Zerroukat, M.: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, *Geoscientific Model Development*, 12, 1909–1963, <https://doi.org/10.5194/gmd-12-1909-2019>, 2019.
- 30 Wang, Z. and Sassen, K.: Cloud type and macrophysical property retrieval using multiple remote sensors, *Journal of Applied Meteorology*, 40, 1665–1682, [https://doi.org/10.1175/1520-0450\(2001\)040<1665:CTAMPR>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<1665:CTAMPR>2.0.CO;2), 2001.
- Warren, E., Charlton-Perez, C., Kotthaus, S., Lean, H., Ballard, S., Hopkin, E., and Grimmond, S.: Evaluation of forward-modelled attenuated backscatter using an urban ceilometer network in London under clear-sky conditions, *Atmospheric Environment*, 191, 532 – 547, <https://doi.org/https://doi.org/10.1016/j.atmosenv.2018.04.045>, 2018.
- 35 Watson-Parris, D., Schutgens, N., Cook, N., Kipling, Z., Kershaw, P., Gryspeerd, E., Lawrence, B., and Stier, P.: Community Intercomparison Suite (CIS) v1. 4.0: a tool for intercomparing models and observations, *Geoscientific Model Development*, 9, 3093–3110, <https://doi.org/10.5194/gmd-9-3093-2016>, 2016.

- Webb, M., Senior, C., Bony, S., and Morcrette, J.-J.: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models, *Climate Dynamics*, 17, 905–922, <https://doi.org/10.1007/s003820100157>, 2001.
- Webb, M. J., Andrews, T., Bodas-Salcedo, A., Bony, S., Bretherton, C. S., Chadwick, R., Chepfer, H., Douville, H., Good, P., Kay, J. E., et al.: The cloud feedback model Intercomparison Project (CFMIP) contribution to CMIP6, *Geoscientific Model Development*, 10, 359–384, <https://doi.org/10.5194/gmd-10-359-2017>, 2017.
- Welton, E. J., Voss, K. J., Gordon, H. R., Maring, H., Smirnov, A., Holben, B., Schmid, B., Livingston, J. M., Russell, P. B., Durkee, P. A., et al.: Ground-based lidar measurements of aerosols during ACE-2: Instrument description, results, and comparisons with other ground-based and airborne measurements, *Tellus B*, 52, 636–651, <https://doi.org/10.1034/j.1600-0889.2000.00025.x>, 2000.
- Welton, E. J., Voss, K. J., Quinn, P. K., Flatau, P. J., Markowicz, K., Campbell, J. R., Spinhirne, J. D., Gordon, H. R., and Johnson, J. E.: Measurements of aerosol vertical profiles and optical properties during INDOEX 1999 using micropulse lidars, *Journal of Geophysical Research: Atmospheres*, 107, INX2–18, <https://doi.org/10.1029/2000JD000038>, 2002.
- Welton, E. J., Campbell, J. R., Berkoff, T. A., Valencia, S., Spinhirne, J. D., Holben, B., Tsay, S.-C., and Schmid, B.: The NASA Micro-Pulse Lidar Network (MPLNET): an overview and recent results, *Opt. Pur. Apl*, 39, 67–74, 2006.
- Werner, M.: Shuttle radar topography mission (SRTM) mission overview, *Frequenz*, 55, 75–79, <https://doi.org/10.1515/FREQ.2001.55.3-4.75>, 2001.
- Wiegner, M. and Gasteiger, J.: Correction of water vapor absorption for aerosol remote sensing with ceilometers, *Atmospheric Measurement Techniques*, pp. 3971–3984, <https://doi.org/10.5194/amt-8-3971-2015>, 2015.
- Wiegner, M. and Geiß, A.: Aerosol profiling with the Jenoptik ceilometer CHM15kx, *Atmos. Meas. Tech*, 5, 1953–1964, <https://doi.org/10.5194/amt-5-1953-2012>, 2012.
- Wiegner, M., Madonna, F., Biniotoglou, I., Forkel, R., Gasteiger, J., Geiß, A., Pappalardo, G., Schäfer, K., and Thomas, W.: What is the benefit of ceilometers for aerosol remote sensing? An answer from EARLINET, *Atmospheric Measurement Techniques*, 7, 1979–1997, <https://doi.org/10.5194/amt-7-1979-2014>, 2014.
- Wiegner, M., Mattis, I., Pattantyús-Ábrahám, M., Bravo-Aranda, J. A., Poltera, Y., Haeffele, A., Hervo, M., Görsdorf, U., Leinweber, R., Gasteiger, J., Haeffelin, M., Wagner, F., Cermak, J., Komínková, K., Brettle, M., Münkkel, C., and Pönitz, K.: Aerosol backscatter profiles from ceilometers: validation of water vapor correction in the framework of CeiLinEx2015, *Atmospheric Measurement Techniques*, 12, 471–490, <https://doi.org/10.5194/amt-12-471-2019>, 2019.
- Williams, D. N., Ananthkrishnan, R., Bernholdt, D., Bharathi, S., Brown, D., Chen, M., Chervenak, A., Cinquini, L., Drach, R., Foster, I., et al.: The Earth System Grid: Enabling access to multimodel climate simulation data, *Bulletin of the American Meteorological Society*, 90, 195–206, <https://doi.org/10.1175/2008BAMS2459.1>, 2009.
- Williams, J., Morgenstern, O., Varma, V., Behrens, E., Hayek, W., Oliver, H., Dean, S., Mullan, B., and Frame, D.: Development of the New Zealand Earth System Model: NZESM, *Weather and Climate*, 36, 25–44, <https://doi.org/10.2307/26779386>, 2016.
- Williams, K. D. and Bodas-Salcedo, A.: A multi-diagnostic approach to cloud evaluation, *Geoscientific Model Development*, 10, 2547, <https://doi.org/10.5194/gmd-10-2547-2017>, 2017.
- Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO mission and CALIOP data processing algorithms, *Journal of Atmospheric and Oceanic Technology*, 26, 2310–2323, <https://doi.org/10.1175/2009JTECHA1281.1>, 2009.
- Wiscombe, W. J.: Mie scattering calculations: Advances in technique and fast, vector-speed computer codes, Tech. rep., National Center for Atmospheric Research Boulder, Colorado, 1979.



- Wiscombe, W. J.: Improved Mie scattering algorithms, *Applied optics*, 19, 1505–1509, 1980.
- Yang, P., Liou, K.-N., Bi, L., Liu, C., Yi, B., and Baum, B. A.: On the radiative properties of ice clouds: Light scattering, remote sensing, and radiation parameterization, *Advances in Atmospheric Sciences*, 32, 32–63, <https://doi.org/10.1007/s00376-014-0011-z>, 2014.
- Yorks, J. E., Hlavka, D. L., Hart, W. D., and McGill, M. J.: Statistics of Cloud Optical Properties from Airborne Lidar Measurements, *Journal of Atmospheric and Oceanic Technology*, 28, 869–883, <https://doi.org/10.1175/2011jtecha1507.1>, 2011.
- Zadra, A., Williams, K., Frassoni, A., Rixen, M., Adames, Á. F., Berner, J., Bouyssel, F., Casati, B., Christensen, H., Ek, M. B., et al.: Systematic Errors in Weather and Climate Models: Nature, Origins, and Ways Forward, *Bulletin of the American Meteorological Society*, 99, ES67–ES70, <https://doi.org/10.1175/BAMS-D-17-0287.1>, 2018.
- Zdunkowski, W., Trautmann, T., and Bott, A.: *Radiation in the atmosphere: a course in theoretical meteorology*, Cambridge University Press, 2007.
- Zhang, Y., Xie, S., Klein, S. A., Marchand, R., Kollias, P., Clothiaux, E. E., Lin, W., Johnson, K., Swales, D., Bodas-Salcedo, A., et al.: The ARM Cloud Radar Simulator for Global Climate Models: Bridging Field Data and Climate Models, *Bulletin of the American Meteorological Society*, 99, 21–26, <https://doi.org/10.1175/BAMS-D-16-0258.1>, 2018.
- Zhang, Z. and Platnick, S.: An assessment of differences between cloud effective particle radius retrievals for marine water clouds from three MODIS spectral bands, *Journal of Geophysical Research: Atmospheres*, 116, <https://doi.org/10.1029/2011JD016216>, 2011.