Reply to Referee #1

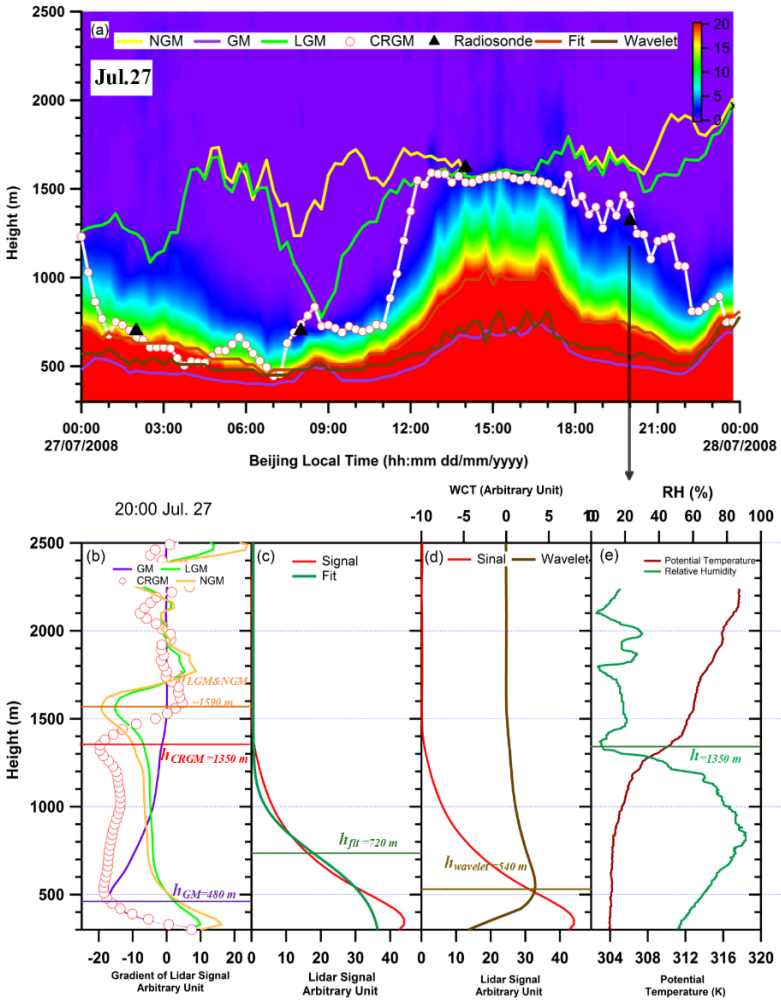
The authors would like to thank the reviewer for their constructive and valuable comments. Accordingly, point-by-point answer is given as follows:

(1) The three algorithms used for comparison are all "Gradient" based methods, the core of which is to find 'local' largest gradient, therefore, their performance is more likely to be impact by data's vertical resolution compare to algorithms focused more on the entire RSCS profile, such as ideal curve fit (Steyn et al. 2000) and wavelet method (Davis et al. 2000). I suggest the authors to do comparisons between CRGM and such algorithms, at least for several cases, to enrich the proof of CRGM's reliability.

Following the reviewer's suggestion, the performance of the CRGM and existing other gradient methods in retrieving the boundary layer height (BLH) has been compared against the ideal curve fit and wavelet methods under five cases including starting, undergoing, ending of a heavy polluted episode, clean atmosphere and multi-layer condition. In ideal curve fit method, entrainment zone thickness (EZT) is set to be 2.77*s* (Steyn et al. 1999) while Haar function with a median dilation of 150 m is employed in wavelet method (Davis et al. 2000). Detailed discussion case by case is given below.

-Undergoing heavy polluted episode (27 July)

As illustrated in ReFig.1-1b, at 20:00 on 27 July (heavy polluted episode), the BLH retrieved by CRGM is 1350 m, in good agreement with radiosonde result (1350 m), against 720m for ideal curve fit and a maximum of 540m for the wavelet covariance transform coefficient (ReFig.1-1c, d). Obviously, ideal curve fit and wavelet methods significantly underestimate the BLH by approximately 630 m and 810 m, respectively, and seem to be unable to fully capture the diurnal cycle of the BLH as showed on ReFig.1-1a. Such underestimate of the BLH from the wavelet method was also previously reported by several studies (Sawyer and Li, 2013; Wang et al., 2012; Su et al., 2017).

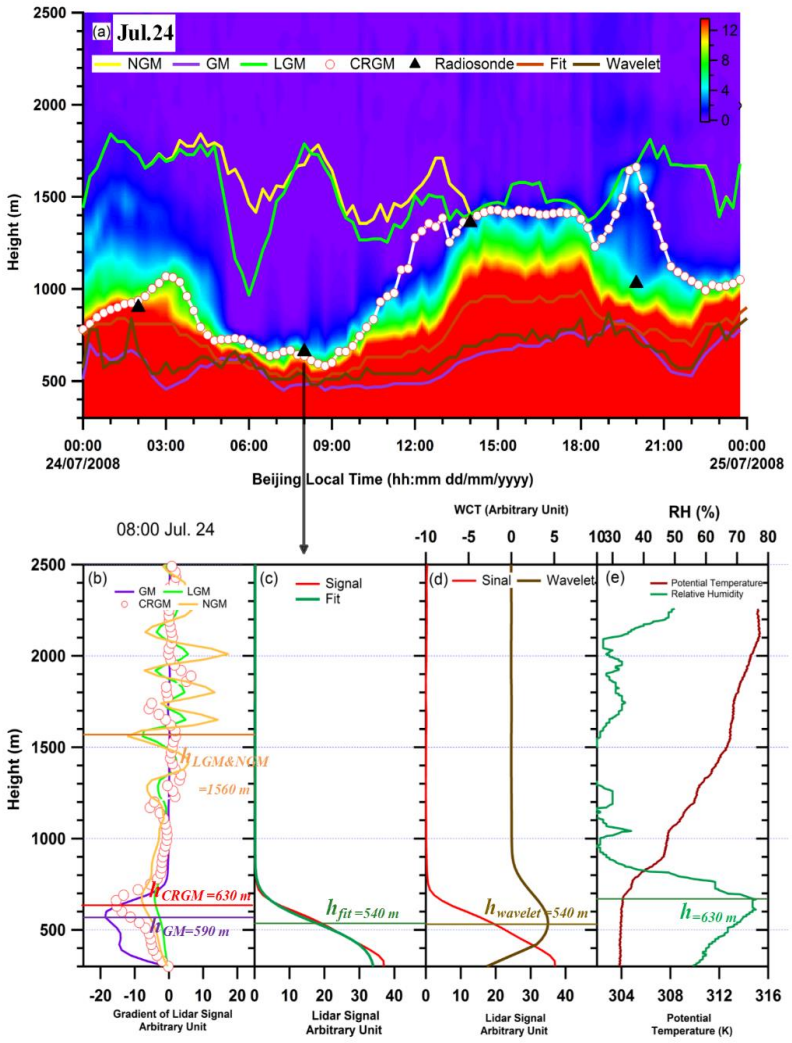


ReFig.1-1: (a) Evolution of the Lidar range-squared -corrected signal (RSCS) at 532 nm on 27Jul. The color bar indicates the intensity of the RSCS; the diurnal BLH retrieved by CRGM, three typical gradient method (LGM, GM and NGM), ideal curve fit and wavelet methods are illustrated as red dotted line, green, purple, yellow, origin and brown lines, respectively; (b) the profile of CRGM and the other three gradient methods and the corresponding retrieval BLH at 20:00 27 July, CRGM is illustrated in red dotted line and LGM, GM and NGM are illustrated as green, purple and yellow lines, receptively; (c) RSCS signal in red line, and ideal fit curve in green line; (d) RSCS signal in red line and wavelet covariance transform (WCT) in brown line; (e) Potential temperature and relative humidity at 20:00 on 27 July.

-Starting heavy polluted episode (24 July)

The diurnal cycle of BLH retrieved by these algorithms in a starting day of heavy polluted episode and typical profiles of the methods are illustrated in ReFig.1-2. The results are similar with those found for undergoing polluted episode. The BLH

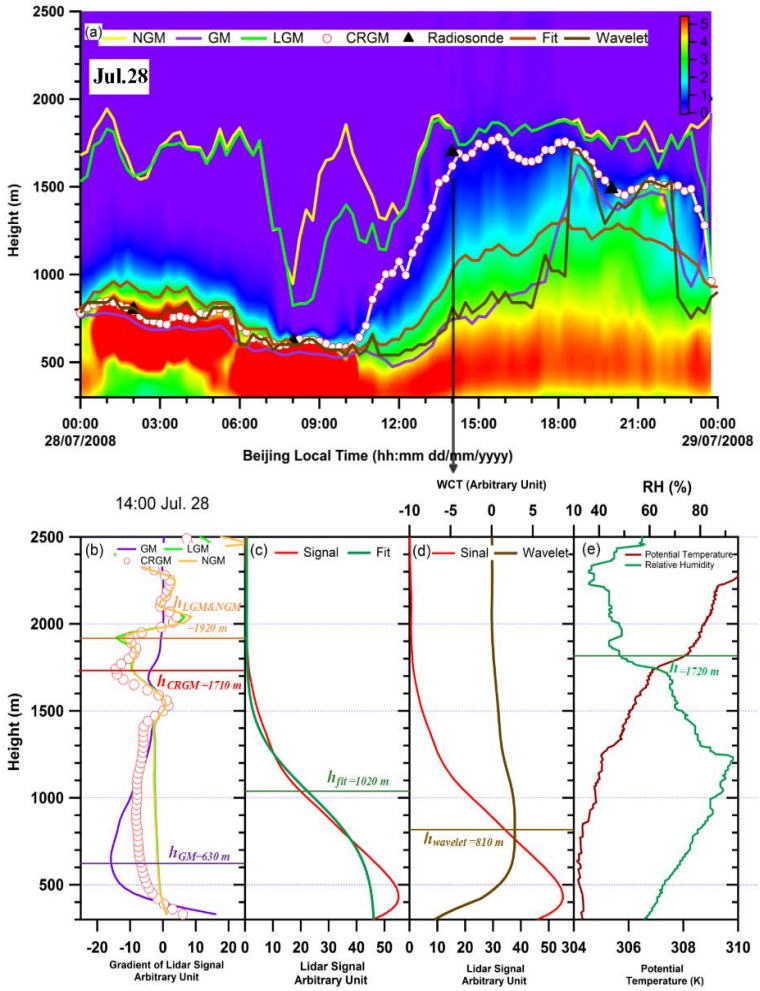
retrievals from ideal fit curve and wavelet methods are significantly underestimation with a maximum of ~600 m and ~ 800 m, in particular in noon and afternoon.



ReFig.1-2: same as ReFig.1-1, but for 24 July.

-Ending heavy polluted episode (28 July)

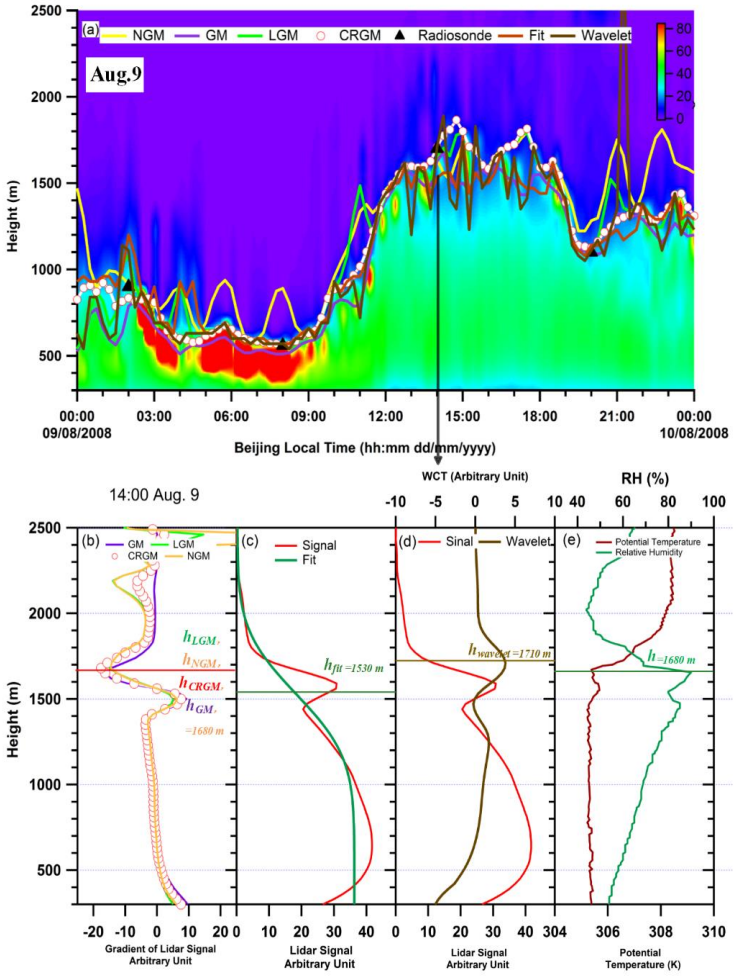
The comparison results are showed on ReFig.1-3 and confirmed persistent underestimation of BLH retrievals from the ideal fit curve and wavelet methods (with a maximum underestimation of ~700 m and ~1000 m respectively), more obvious in moon and afternoon (as found in previous cases).



ReFig.1-3: same as ReFig.1-1, but for 28 July.

-Clean atmosphere day (9 Aug)

In this case, the results show that all retrieval algorithms capture relatively well the overall diurnal cycle of the BLH (ReFig.1-4). However, the retrieval from the ideal curve fit seems to present a slight underestimation of 150m at 14:00 9 Aug (ReFig.1-4c).



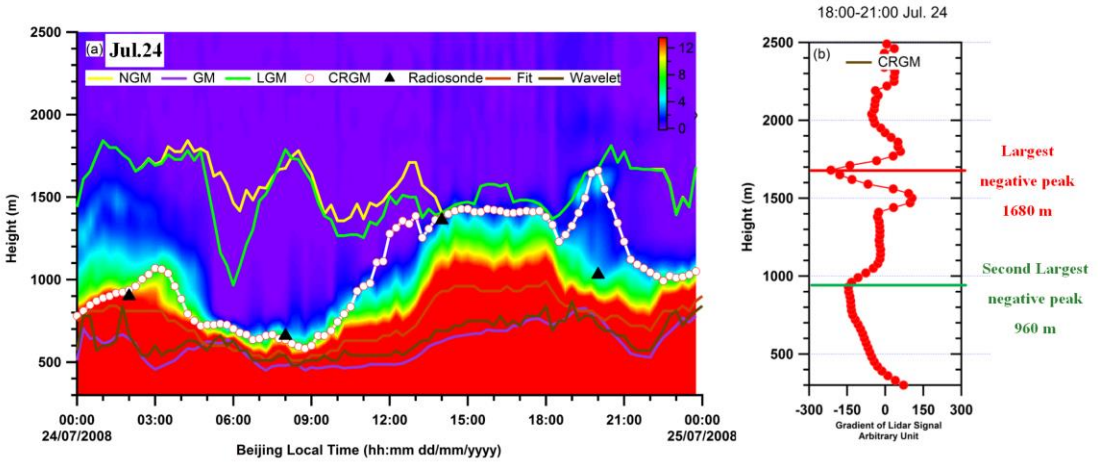
ReFig.1-4: same as ReFig.1-1, but for 9 Aug.

-Multi-layer condition

Proper analysis of algorithm performance under multi-layer condition is given with the reply to comment (3).

(2) ***A case shows that the new algorithm seems fail to distinguish residual layer from boundary layer from about 18:00-21:00 in Jul 24, 2008 (Figure S4). More discussions should be done to analyze the reason of fault retrieval result.***

Yes, exactly! The new algorithm fails to define the boundary layer height for 3 hours on 24 July. By further examining the mean profile of CRGM over 18:00-21:00, as illustrated in ReFig.1-5b, it presents two significant negative peaks (1680 m and 960 m). We believe that the boundary layer height determination might be significantly disturbed by the largest negative value (at 1680m), resulting in a failure of the CRGM.



ReFig.1-5: same as ReFig.1-1, (b) the main gradient profile of CRGM during 18:00-21:00 Jul24.

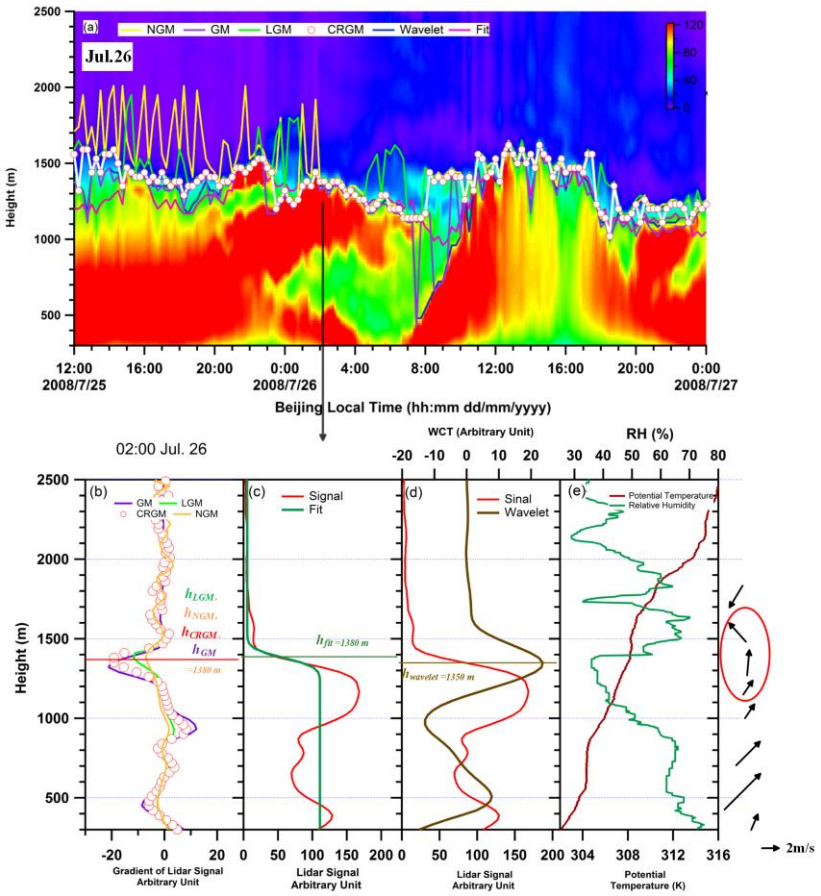
Actually, in our manuscript, the main purpose is to evaluate the performance of the algorithms (advantages and limitations) without further data quality control of the retrieval BLH, thus the BLH results provided in Fig.S4 are independent from moment to moment. However, when carefully considering the data quality checking process (usually applied in such circumstance) consisting of (1) Searching and checking for more accurate negative peak under BLH decreasing condition at sunset (if the BLH is determined by the largest negative peak, higher than that of previous moment), (2) eliminating the fault negative peaks according to BLH climatic condition at near ground level or high altitude (over 3 km in summer, >1 km at night time), (3) comparing with neighboring moment results and ensuring the retrieval results are reliable in time and space. By applying the above quality checking processes, the determination of the BLH without failure is relatively assured (960 m at second largest negative gradient peak in this case, ReFig.1-5b).

(3) This multi-layer aerosol structure is just the difficulty and emphasis for BLH retrieval from Lidar profiles, just like the impact of cloud in BLH retrieval. It is better that the authors show some more cases of retrieval result for the multi-layer aerosol to verify the availability of new algorithm.

In accordance with the reviewer suggestion, we have applied all algorithms to the most complex case for multi-layer aerosol structure in our database, occurring on 26 July (ReFig1-6), in order to provide further discussion on the BLH retrieval

performance and limitation. Meanwhile, some improving suggestions also discussed for complex multi-layer cases.

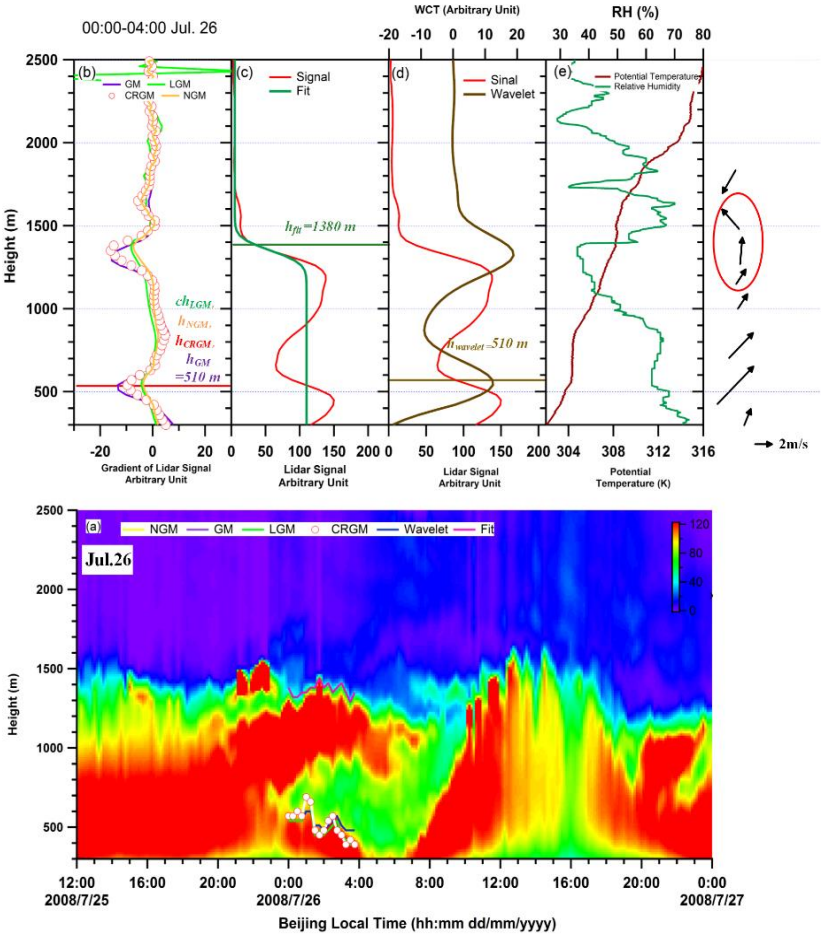
As illustrated in ReFig.1-6a, an obvious two-layer aerosol structure is observed from 00:00-04:00 Jul.26, associated with a low aerosol concentration "hole" extended from ~300m to 1300m. The BLH at noon time (12:00) of the day before (Jul.25) is about 1500m, then gradually decreases to ~1300 m at 20:00 due to the ground cooling in absence of solar short-wave radiation. Significant variations of aerosol signal at upper and lower parts are also observed on 26 July between 00:00 and 04:00. Basically, the formed upper and lower parts seem not to be the natural residual and nocturnal boundary layer, but probably rather triggered by large wind driving force.



ReFig.1-6: same as ReFig.1-1, but for 26 July.

The BLH is defined at 1380m by CRGM, NGM, GM, LGM and the ideal fit curve, and at 1350 m by the wavelet transform method at 02:00 Jul.26 (ReFig.1-6 b, c ,d). However, all algorithms seem to only track the top of the residual layer over the multi-layer sequence (00:00-04:00 Jul. 26) and fail to capture the nocturnal layer top.

Furthermore, analysis of the mean vertical profiles of the algorithms (ReFig.1-6 b, c ,d; ReFig.1-7b,c,d), enables to notice two obviously negative peaks induced by CRGM, NGM, GM, LGM and the wavelet method profiles (a first largest gradient peak noticeable at ~1380 m on the top of residual layer, and a second largest peak occurring at 510 m near the top of nocturnal boundary layer). Even though all algorithms fail to track the BLH variation induced by the first largest peak, the second largest peaks work well for nocturnal boundary layer with CRGM, CRGM, NGM, GM, LGM and the wavelet method when applying the data quality control checking as described in the reply to comment (2). The retrieved nocturnal boundary layer height variation by each algorithm with data control during the multi-layer period is illustrated in ReFig.1-7a.



ReFig.1-7: same as ReFig.1-6, but after data quality control.

(4) P2, L4. "Duo to" should be "Due to".

It has been revised in the manuscript.

(5) *P2, L24. "BHL" should be "BLH".*

It has been revised in the manuscript.

(6) More states about cloud and rain mask method should be done in Section 2.

The cloud and rain detection follows the methods employed by Asian dust and aerosol Lidar observation network (AD-net) in East Asia, which was supported by world meteorological organization (WMO) Global Atmosphere Watch (GAW) program. Rain was detected by color ratio (γ', the ratio of β_{1064}' to β_{532}') to distinguish rainy and clear (no rain) regions, in which, β_{1064}' *and* β_{532}' present the attenuated backscatter coefficient at 1064 nm and 532 nm, respectively. Large droplets have a large γ' value, so once γ' exceeds a threshold (1.1) over a certain vertical internal in the lower atmosphere, the profile is classified as a rain profile. Cloud base height is determined by the vertical gradient of β_{1064}' and the peak value of β_{1064}' between the cloud base and the apparent cloud top. The detailed description of the method is provided by Shimizu et al., (2016).

Davis, K. J., Gamage, N., Hagelberg, C. R., Kiemle, C., Lenschow, D. H., and Sullivan, P. P.: An Objective Method for Deriving Atmospheric Structure from Airborne Lidar Observations, J. Atmos. Oceanic Technol., 17, 1455-1468, 2000.

Sawyer, V., and Li, Z.: Detection, variations and intercomparison of the planetary boundary layer depth from radiosonde, lidar and infrared spectrometer, Atmos. Environ., 79, 518-528, 10.1016/j.atmosenv.2013.07.019, 2013.

Steyn, D. G., Baldi, M., and Hoff, R. M.: The Detection of Mixed Layer Depth and Entrainment Zone Thickness from Lidar Backscatter Profiles, J. Atmos. Oceanic Technol., 16, 953-959, 1999.

Shimizu, A., Nishizawa, T., Jin, Y., Kim, S.-W., Wang, Z., Batdorj, D., and Sugimoto, N.: Evolution of a lidar network for tropospheric aerosol detection in East Asia, OPTICE, 56, 031219-031219, 10.1117/1.OE.56.3.031219, 2016.

Su, T., Li, J., Li, C., Xiang, P., Lau, A. K.-H., Guo, J., Yang, D., and Miao, Y.: An intercomparison of long-term planetary boundary layer heights retrieved from CALIPSO, ground-based lidar and radiosonde measurements over Hong Kong, J. Geophys. Res:Atmos., 10.1002/2016jd025937, 2017.

Wang, Z., Cao, X., Zhang, L., Notholt, J., Zhou, B., Liu, R., and Zhang, B.: Lidar measurement of planetary boundary layer height and comparison with microwave profiling radiometer observation, Atmos. Meas. Tech., 5, 1965-1972, 10.5194/amt-5-1965-2012, 2012.