Responds to Anonymous Referee #2:

The paper entitled "A novel Mie lidar gradient cluster analysis method of nocturnal boundary layer detection during air pollution episodes" presents a novel method based on the cluster analysis of the gradient method to detect the nocturnal boundary layer from lidar signals. The method presented is of interest for the retrieval of the nocturnal boundary layer. However, the language and writing of the paper needs to be greatly improved. At its current state, the paper is quite confusing and difficult to follow. In general, the discussion of your results needs to be improved. My suggestion is to resubmit the paper after the language has been carefully reviewed and the previous comments addressed.

Response:

Thanks a lot for your reviews on our manuscript entitled "A novel Mie lidar gradient cluster analysis method of nocturnal boundary layer detection during air pollution episodes (ID: amt-2020-167). We have revised the manuscript according to the comments, the language has been polished by Elsevier Language Editing Services. Moreover, the comprehensive reference and the discussion of the results have been added. The details are shown as follows.

General comments:

1. Please, use more comprehensive and recent bibliography.

Response:

The recent bibliography of boundary layer detection have been add in the P2 49-63.

The latest developments in the boundary layer height measurement, including classical methodology, graphic methodology, and algorithm with multiple layer structure interference have been added and express the relationship of our studies.

'

'Some graph theory methods, such as the extended Kalman filter (Banks et al., 2014), Pathfinder and PathfinderTURB (de Bruine et al., 2017; Poltera et al., 2017), k-means clustering (Liu et al., 2018; Toledo et al., 2014), and The STRAT-2D algorithm (Haeffelin et al., 2012) have been proposed to yield promising results via an automated method that reduces the incorrect detection of ABLH. However, these techniques strongly depend on the vertical distribution of particle layers (aerosols and clouds) and are unsuitable for use under complicated multilayer conditions (Granados-Muñoz et al., 2012).

The retrieval of BLHs under cloudy conditions is quite challenging. Some researchers have used the threshold of the attenuated scattering ratio (Campbell et al., 2008; Winker and Vaughan, 1994), the ratio of peaks to the base of the range-corrected signal (RCS) (Wang and Sassen, 2001) to locate cloud tops and bases, while others have employed the objective upper limit of the convective condensation level (CCL)(Li et al., 2017), as well as the analysis of signal continuity and the classification of whether the cloud caps the ABLH or is decoupled from the ABL (Dang et al., 2019). The height restriction has significant advantages in removing the influence of clouds. Elevated aerosol layers (EALs) are characteristically similar to the aerosol trapped in ABL, using the threshold of lidar backscatter coefficient can distinguish them (Dubovik et al., 2002;

Hänel et al., 2012; Peng et al., 2017). More instrument and multi-wavelength lidar systems are combined to obtain more accurate results to identified the EALs (Liu et al., 2019; Ortega et al., 2016).

2. State clearly what is the advantage of the proposed method compared to previous studies. From your results, the improvement compared to previous methods is not so evident in some cases.

Response:

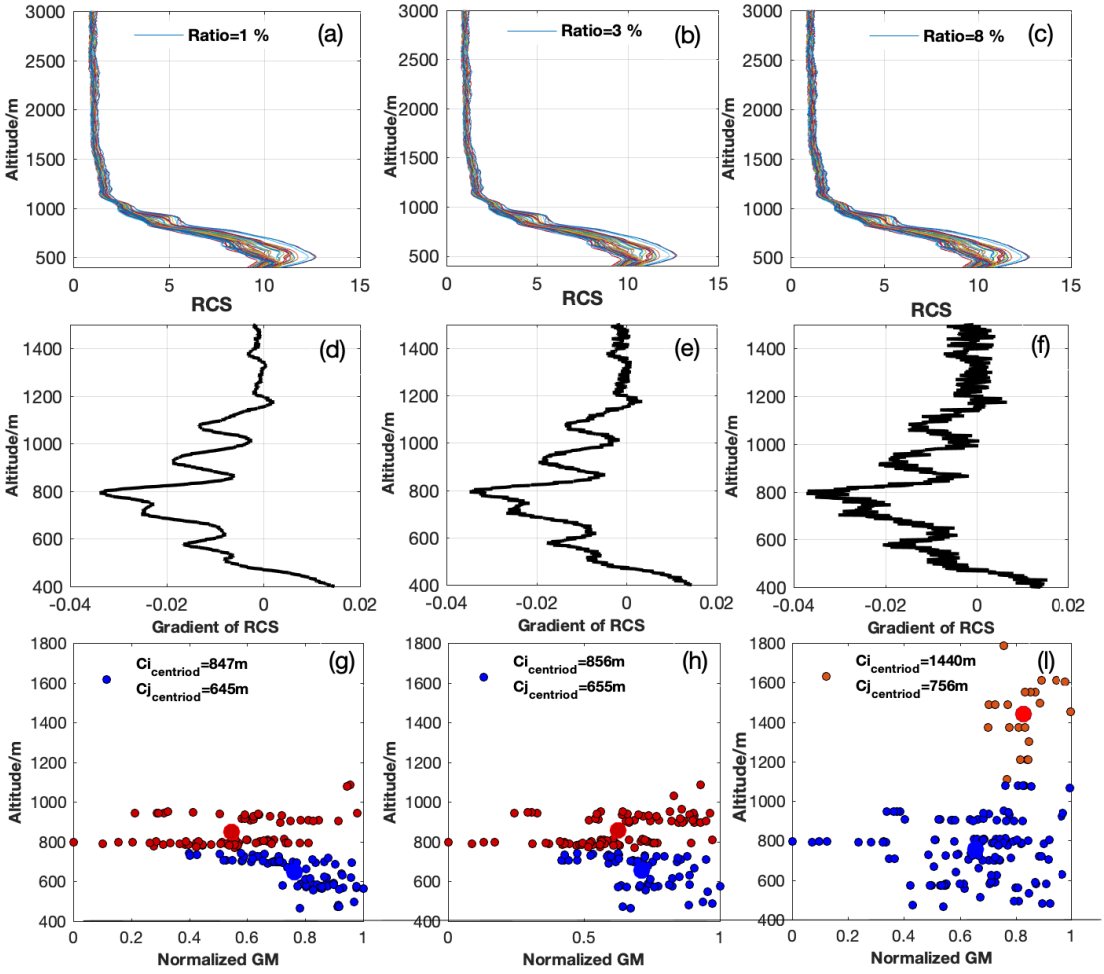
Thank you for your suggestion. We come up with an algorithm based on cluster analysis of gradient method (CA-GM) in the nocturnal boundary layer (NBLH). Compared to the gradient-based methods such as GM and CRGM, it will be more robust in a nosier condition. With the test of the real signal, the CA-GM has better performance through polluted cases (Figure R2-1). Secondly, although wavelet covariance transform (WCT) method is robust in noise affection, it can still be affect by low-level cloud and the aloft aerosol layers. As figure 8-1 shows, the time period on 17:00-18:00 and 21:00-22:00, the CA-GM has significant ability to capture the NBL. Thirdly, according to the Table 2, the best correlation with radiosonde (0.85), the smallest RMSE (203 m) with radiosonde are shown the better performance of the CA-GM in capturing the NBL in polluted cases.

The testing with the real signal are shown below.

Use the $\mathrm{RCS}(z)$ signal, and randomly noised $\mathrm{RCS}^{noised}(z)$ by the expression:

$$\mathrm{RCS}^{noised}(z) = \mathrm{RCS}(z) + [\alpha \times \chi(z)] \quad \text{(R2-1)}$$

Where $\chi(z)$ is the random noise function taking values between 0 and 1, z is the height, and α is a varying parameter as introduced in Eq (R2-1) to produces different levels of noise.



55 **Figure R2-1. The real lidar RCS for the heavily polluted case (17 Dec 2016 20:00-21:00 LST).**
(a-c) three noise level cases, (d-f) with the gradient of RCS, and (g-i) the first weighted k-means clustering.

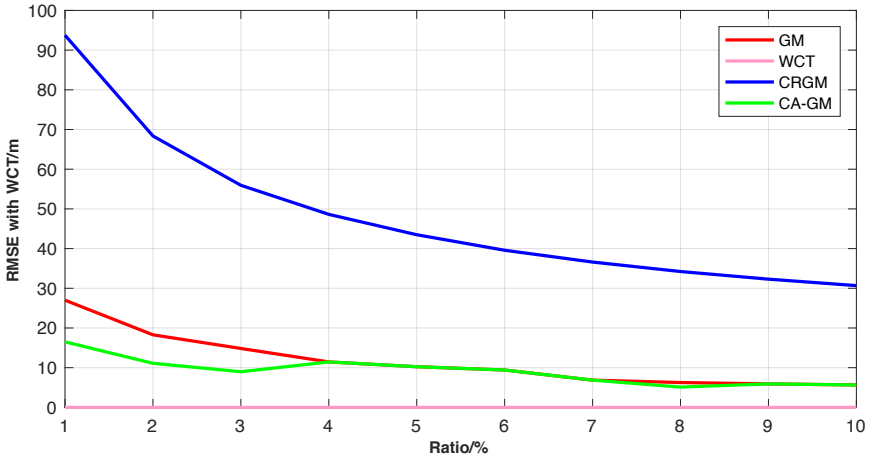


Figure R2-2 RMSE between the WCT and the other three algorithms (GM,CRGM and CA-GM)

As a result of the figure R2-2 shows, the CA-GM has less RMSE than GM at the ratio of 1%-4%. The figure R2-1 (g-h) shows similar groups in different range of noise affection. However, the clustering changes at the results of R2-1(i). Due to the noise distribution of the signal, the centriod of the cluster will get higher and lose the ability to restrict the changes of GM. The difference in the NBL top is found with the noise level 4% case, they are lower than 1% in respect to the estimate for the case with raw signal.

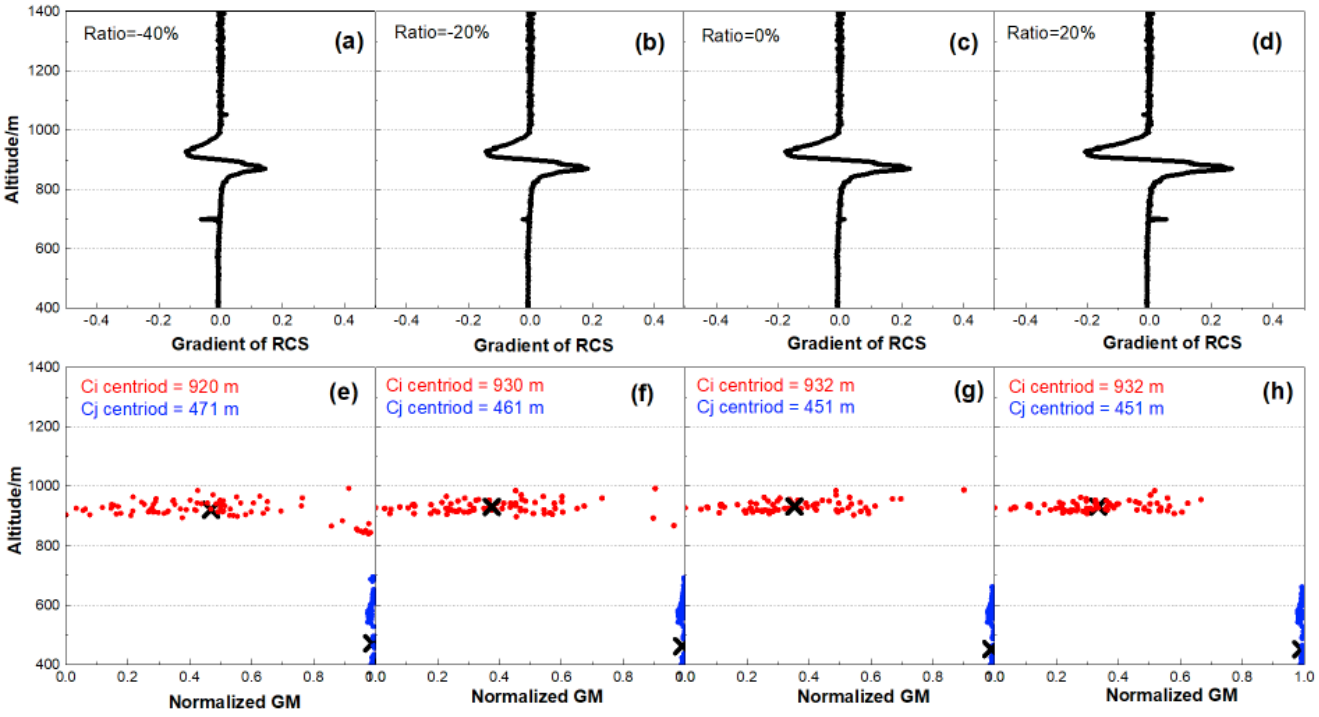


Figure R2-3. The real lidar RCS for the cloud case (5 Jan 2016 00:00-1:00 LST).
(a-d) the different ratio of strength of the cloud layer intensity. (e-h) the first weighted-means clustering

Add the signal of the cloud layer on the raw data, the ratio of the intensity for cloud layer changes from -40% to 40%. As the figure shown the first k-means clustering in figure R2-3(e-h),the intensity of the cloud layer will not influence the CA-GM.

In summary, these results indicate that the degree of estimation of the NBL top by applying CA is weakly affected by the signal noise. In fact, a few NBLH depending on the value of the RCS gradient in a discrete point. CA determines the NBL by taking into account the overall set of observations of a given point, thus decreases the dependence of the method on the RCS values in single moment. The intensity of the CLs changes ±40% and will not affect the cluster of the CA-GM, it can be significant stratified due to the relative significantly signal difference on the backscatter signal. As for the EALs, the strict threshold will defined the EALs accurately. Therefore, the CA-GM approach is able to accurately obtain the NBLH with the effect of noise, cloud layers and elevated aerosol layers.

The superiority of the CA-GM are added at the discussion section.

‘ In relation to the robustness of the CA-GM approach, the effect of the lidar RCS noise in determining the NBLH has been

analyzed. Unlike other gradient-derivate methods, CA-GM results are slightly affected by lidar signal noise. NBL top height as obtained for 'noised' lidar RCS with value of $\alpha < 4\%$ is better than GM results. The intensity of the CLs changes ±40% will not affect with the cluster of the CA-GM in the polluted cases, the significantly starfield structure is related to the relative difference on the backscatter signal. As for the EALs, the strict threshold will defined the EALs accurately.'

3. Indicate clearly what is the method used to retrieve the NBLH from the radiosonde data. In section 4.3 different criteria seems to be used to establish a reference NBLH for the comparison with the radiosonde without being clearly justified.

Respond:

The different criteria description of the radiosonde data was added in the Section 2.1.

'As a result of the complexity of the transition during the morning and early at night, the boundary layer is in a transition between stable and unstable conditions. To determine NBLH from the radiosonde vertical profiles of temperature and humidity, the elevated temperature inversion layer or the height of a significant reduction in moisture is used (Peng et al., 2017). The potential temperature gradient (PTG) should have a good correlation with the relative humidity gradient (RHG), with an allowable error of 100 m (Wang and Wang, 2016). In this study, if the difference between the PTG and RHG is in excess of 100 m, the PTG is considered first, whereas if there is no significant temperature change or the evident changes belong to the cloud or EALs, the result of RHG is referred to as the NBLH.'

4. In general, the discussion of your results needs to be improved.

Responds:

Thank you for your suggestion. The discussion are add two parts of the content, about the uncertainty and the limitation of the algorithm. P17 Line 333-346.

'The uncertainty of the CA-GM is calculated by the real signal. Concerning the robustness of the CA-GM approach, the effect of the lidar RCS noise in determining the NBLH has been analyzed. Unlike other gradient-derivate methods, CA-GM results are slightly affected by lidar signal noise. NBL top height as obtained for 'noised' lidar RCS with value of $\alpha < 4\%$ is better than GM results. The intensity of the CLs changes ±40% will not affect with the cluster of the CA-GM in the polluted cases, the significantly starfield structure is related to the relative difference on the backscatter signal. As for the EALs, the strict threshold will defined the EALs accurately. The limitation of the CA-GM is based on the assumption that the nocturnal boundary layer is stable, hence, we can calculate the distribution of the minima gradients of the RCS in an hour interval to use weighted k-means clustering to work as height restriction to the layers. Secondly, based on the limitation of the lidar system. The lower limit of the BIT-lidar is around 300 m. Too shallow of nocturnal boundary layer height (NBLH) are not be detectable. Thirdly, the method should be used in the high SNR condition, such as night-time and air pollution. '

Thank you so much for your reviewing! We deeply appreciate your recognition of our research work.

Reference

Banks, R. F., Tiana-Alsina, J., María Baldasano, J. and Rocadenbosch, F.: Retrieval of boundary layer height from lidar using extended Kalman filter approach, classic methods, and backtrajectory cluster analysis, edited by A. Comerón, E. I. Kassianov, K. Schäfer, R. H. Picard, K. Stein, and J. D. Gonglewski, p. 92420F, Amsterdam, Netherlands., 2014.

de Bruine, M., Apituley, A., Donovan, D. P., Klein Baltink, H. and de Haij, M. J.: Pathfinder: applying graph theory to consistent tracking of daytime mixed layer height with backscatter lidar, Atmospheric Measurement Techniques, 10(5), 1893–1909, doi:10.5194/amt-10-1893-2017, 2017.

Campbell, J. R., Sassen, K. and Welton, E. J.: Elevated Cloud and Aerosol Layer Retrievals from Micropulse Lidar Signal Profiles, Journal of Atmospheric and Oceanic Technology, 25(5), 685–700, doi:10.1175/2007JTECHA1034.1, 2008.

Dang, R., Yang, Y., Li, H., Hu, X.-M., Wang, Z., Huang, Z., Zhou, T. and Zhang, T.: Atmosphere Boundary Layer Height (ABLH) Determination under Multiple-Layer Conditions Using Micro-Pulse Lidar, Remote Sensing, 11(3), 263, doi:10.3390/rs11030263, 2019.

Dubovik, O., Holben, B., Eck, T. F., Smirnov, A., Kaufman, Y. J., King, M. D., Tanré, D. and Slutsker, I.: Variability of Absorption and Optical Properties of Key Aerosol Types Observed in Worldwide Locations, Journal of the Atmospheric Sciences, 59(3), 590–608, doi:10.1175/1520-0469(2002)059<0590:VOAAOP>2.0.CO;2, 2002.

Granados-Muñoz, M. J., Navas-Guzmán, F., Bravo-Aranda, J. A., Guerrero-Rascado, J. L., Lyamani, H., Fernández-Gálvez, J. and Alados-Arboledas, L.: Automatic determination of the planetary boundary layer height using lidar: One-year analysis over southeastern Spain: DETERMINATION OF THE PBL HEIGHT, Journal of Geophysical Research: Atmospheres, 117(D18), n/a-n/a, doi:10.1029/2012JD017524, 2012.

Haeffelin, M., Angelini, F., Morille, Y., Martucci, G., Frey, S., Gobbi, G. P., Lolli, S., O'Dowd, C. D., Sauvage, L., Xueref-Rémy, I., Wastine, B. and Feist, D. G.: Evaluation of Mixing-Height Retrievals from Automatic Profiling Lidars and Ceilometers in View of Future Integrated Networks in Europe, Boundary-Layer Meteorology, 143(1), 49–75, doi:10.1007/s10546-011-9643-z, 2012.

Hänel, A., Baars, H., Althausen, D., Ansmann, A., Engelmann, R. and Sun, J. Y.: One-year aerosol profiling with EUCAARI Raman lidar at Shangdianzi GAW station: Beijing plume and seasonal variations: ONE-YEAR AEROSOL PROFILING NEAR BEIJING, Journal of Geophysical Research: Atmospheres, 117(D13), n/a-n/a, doi:10.1029/2012JD017577, 2012.

Li, H., Yang, Y., Hu, X.-M., Huang, Z., Wang, G. and Zhang, B.: Application of Convective Condensation Level Limiter in Convective Boundary Layer Height Retrieval Based on Lidar Data, Atmosphere, 8(12), 79, doi:10.3390/atmos8040079, 2017.

Liu, B., Ma, Y., Liu, J., Gong, W., Wang, W. and Zhang, M.: Graphics algorithm for deriving atmospheric boundary layer heights from CALIPSO data, Atmospheric Measurement Techniques, 11(9), 5075–5085, doi:10.5194/amt-11-5075-2018, 2018.

Liu, B., Ma, Y., Gong, W., Zhang, M. and Yang, J.: Improved two-wavelength Lidar algorithm for retrieving atmospheric boundary layer height, Journal of Quantitative Spectroscopy and Radiative Transfer, 224, 55–61, doi:10.1016/j.jqsrt.2018.11.003, 2019.

Ortega, I., Berg, L. K., Ferrare, R. A., Hair, J. W., Hostetler, C. A. and Volkamer, R.: Elevated aerosol layers modify the O2–O2 absorption measured by ground-based MAX-DOAS, Journal of Quantitative Spectroscopy and Radiative Transfer, 176, 34–49, doi:10.1016/j.jqsrt.2016.02.021, 2016.

Peng, J., Grimmond, C. S. B., Fu, X., Chang, Y., Zhang, G., Guo, J., Tang, C., Gao, J., Xu, X. and Tan, J.: Ceilometer-Based Analysis of Shanghai's Boundary Layer Height (under Rain- and Fog-Free Conditions), Journal of Atmospheric and Oceanic Technology, 34(4), 749–764, doi:10.1175/JTECH-D-16-0132.1, 2017.

Poltera, Y., Martucci, G., Collaud Coen, M., Hervo, M., Emmenegger, L., Henne, S., Brunner, D. and Haefele, A.: PathfinderTURB: an automatic boundary layer algorithm. Development, validation and application to study the impact on in situ measurements at the Jungfraujoch, Atmospheric Chemistry and Physics, 17(16), 10051–10070, doi:10.5194/acp-17-10051-2017, 2017.

Toledo, D., Córdoba-Jabonero, C. and Gil-Ojeda, M.: Cluster Analysis: A New Approach Applied to Lidar Measurements for Atmospheric Boundary Layer Height Estimation, Journal of Atmospheric and Oceanic Technology, 31(2), 422–436, doi:10.1175/JTECH-D-12-00253.1, 2014.

Wang, X. and Wang, K.: Homogenized Variability of Radiosonde-Derived Atmospheric Boundary Layer Height over the Global Land Surface from 1973 to 2014, Journal of Climate, 29(19), 6893–6908, doi:10.1175/JCLI-D-15-0766.1, 2016.

Wang, Z. and Sassen, K.: Cloud Type and Macrophysical Property Retrieval Using Multiple Remote Sensors, Journal of Applied Meteorology, 40(10), 1665–1682, doi:10.1175/1520-0450(2001)040<1665:CTAMPR>2.0.CO;2, 2001.

Winker, D. M. and Vaughan, M. A.: Vertical distribution of clouds over Hampton, Virginia observed by lidar under the ECLIPS and FIRE ETO programs, Atmospheric Research, 34(1–4), 117–133, doi:10.1016/0169-8095(94)90084-1, 1994.