https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020

Atmospheric Measurement Techniques Discussions

A novel Mie lidar gradient cluster analysis method of nocturnal boundary layer detection during air pollution episodes

Yingchao Zhang[1], **Su Chen**[1], **Siying Chen**[1], **He Chen**[1], **Pan Guo**[1]

[1] School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China

Correspondence to: Siying Chen (csy@bit.edu.cn)

Abstract. The observation of the nocturnal boundary layer height (NBLH) plays an important role in air pollution and monitoring. Through 39 days heavily polluted observation experiments over Beijing (China) and exhaustive evaluation of gradient method (GM), wavelet covariance transform method (WCT) and cubic roots gradient method (CRGM), a novel algorithm based on cluster analysis of gradient method (CA-GM) of lidar signal is developed to capture the multilayer structure and achieve stability in the nighttime. The CA-GM highlights its performance in comparison with radiosonde data, the best correlation (0.85), the weakest root mean square error (203 m), and the improved 25 % correlation coefficient by the GM. In comparison with long-term experiments with other algorithms, a reasonable parameter selection can distinguish layers with different properties, such as the cloud layer, elevated aerosol layers, and random noise. Consequently, the CA-GM can automatically deal with the uncertainty of the multiple structures and obtain a stable NBLH with a high time resolution, which expected to contribute to air pollution monitoring and climatologies, as well as model verification.

1. Introduction

Air pollution has an important impact on human health, climatic patterns and the ecological environment (Shi et al., 2019). The primary anthropogenic emission source is the particulate matters (PMs) which are the major source of severe haze in Beijing (Ma et al., 2019). Many passive and active remote sensing instruments have been combined to observe the aerosol optical and microphysical properties (Ji et al., 2018b; Wang et al., 2019), the relationship of PMs and meteorology (Li et al., 2019; Zhang et al., 2015), and the aerosol-atmospheric boundary layer height (ABLH) interaction (Dong et al., 2017; Su et al., 2020b). With the development in star and moon photometry, the continues day-to-night detection improved the columned-integrated aerosol properties at night. (Benavent-Oltra et al., 2019; Pérez-Ramírez et al., 2008). Nevertheless, the observations of the variation of nocturnal boundary layer (NBL) are few. Because the complexity of weak wind forcing, significant stratification, and intermittent turbulence (Weil, 2011), resulting in the accumulation of fine particles continuously near-surface. The turbulence mixing process is a companion with the strong physio-chemical effect, which favors the formation of new-particles and induce a worse polluted episode (Hao et al., 2018; Wang et al., 2018). Therefore, acquiring the nocturnal boundary layer height (NBLH) in pollution cases accurately, especially during nighttime, is of great significance for air pollution.

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020


Atmospheric Measurement Techniques Discussions

Multiple approaches have been developed to determine the ABLH on observation, including radiosounding, remote sensing as well as parameterization based on a laboratory experiment (Su et al., 2020a). The lidar uses an aerosol as a tracer for mixing processes with high space and time resolutions. According to the research of Martucci et al (2007), the correlation coefficient is lower under stable conditions due to a complex aerosol structure that increases the difficulty of NBLH retrieval (Emeis and Schäfer, 2006; Martucci et al., 2007). During the nighttime, the NBLH determined by elastic lidar is either the top of the residual layer (RL) or the top of the surface mechanical driven mixing layer (Dang et al., 2019a). In the absence of any external forces, aerosols in the atmosphere get stratified and result in single or multiple layers (elevated or advent aerosols), depending on the location and type of the atmospheric aerosols (Dudeja, 2019). The more complex vertical backscatter signal profile can also form in the specific environmental conditions such as cloud contamination and the local signal noise affection (Dang et al., 2019a; Stull, 1988).

The classical methodologies of lidar-retrieved algorithms are difficult in the identification of multilayer structure in night polluted cases. The gradient-based method, such as first-order gradient method (GM) (Hayden et al., 1997), inflection point method (Menut et al., 1999), logarithm gradient method (Toledo et al., 2017), and cubic root gradient method (CRGM) (Yang et al., 2017), are sensitive to noisy data unless averaging signal or lose some instantaneous useful information. The threshold method is too subjective to set a universal threshold on different weather and terrain (Frioud et al., 2003). The variance method (Hooper and Eloranta, 1986) is easily get infected by lofted aerosol layers and reduce the temporal resolution because of calculating the profile of variance. The haar wavelet covariance transform (WCT) method (Davis et al., 2000) and the idealized backscatter method (Steyn et al., 1999) is more robust to noise, but it still can be infected by the low-level cloud and aloft aerosols (Caicedo et al., 2017). The START-2D (Haeffelin et al., 2012) performs well under multiple-layer situations, but the wide information to the thermodynamic state of atmosphere should be added to the best measure of the transition of the boundary layer. Some advanced methods like extend Kalman filter (Banks et al., 2014) and the PathfinderTURB (Poltera et al., 2017) are proposed due to track the ABLH over time coherently and may jump between different atmospheric layers inadvertently. The layer categorization still remains one of the major sources of uncertainty related to the ABLH retrieval.

Departing from these previous efforts to estimate the ABLH, in this paper, we present a new approach on cluster analysis of gradient method (CA-GM) to overcome multilayer structure and remove the fluctuation of NBLH with raw data resolution. This study proposes a reasonable parameter to reduce the interference of the cloud layer, elevated aerosol layers, and the local noise over the air pollution in the megacity region. The results evaluated by comparing with the nearby radiosonde site and confirming the long-term variations with the traditional methods in different atmospheric layers.

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020


Atmospheric Measurement Techniques Discussions

2. Instruments and datasets

2.1 $PM_{2.5}$ data and Radiosonde

Beijing is located in the North Plain of China and experienced serious intermittent haze pollution from Dec 2016 to Dec 2017. The 39 days lidar and radiosonde data were recorded in that period, and the average concentrations of $PM_{2.5}$ were reached 140 μg/m^3 in the experiment. The dataset for lidar, the averaged PMs and AQI are provided at Sect.1 in the Supplement. In situ $PM_{2.5}$ daily measurements in China are primarily from the official website of the China National Environmental Monitoring Centre (CEMC; http://cnemc.cn/). The radiosonde data were released daily from Nanjiao station (39.80° N,116.47° E) , which located 21 km to the southeast of the Beijing Institute of Technology lidar (BIT-lidar) (39.95° N,116.32° E). The L-band radiosonde provided fine-resolution profile of temperature, pressure, relative humidity, wind speed and direction twice a day at 08:00 and 20:00 local standard time (LST) (Guo et al., 2016). Previous studies (Hennemuth and Lammert, 2006; Seidel et al., 2012) have used the radiosonde as a reference in detecting ABLHs for daily and annual changes in lidar measurement. We resampled the radiosonde data using linear interpolation to achieve a same vertical resolution of lidar and compared to the 1-h average NBLH centred around the radiosonde launch times. As a result of the complexity of the transition during the morning and the early night, we combined relative humidity gradient (RHG) as a secondary reference which should have good consistency with the potential temperature (PTG) method (Seibert et al., 2000).

2.2 the BIT-lidar system

A single-wavelength Raman-Mie lidar is operated in the campus of Beijing institute of Technology, providing aerosol, cloud, ABLHs and temperature measurements. This lidar system has been continuously improved to capture the loading of aerosols (Chen et al., 2014; Ji et al., 2018a). The standardized range-corrected signal (RCS) is subjected to correlation lidar factor correction (correction of electronic noise error, background noise error, overlap factor) and distance correction. The backscatter coefficient can be calculated as the Fernald method (Fernald, 1984) and assumed lidar ratio as 70 sr (Rosati et al., 2016) due to the polluted continental aerosol particles. The detailed parameters of the BIT-lidar are listed at Table 1.

Table 1. Key parameter of the BIT-lidar

Parameter	BIT-lidar
Laser	Nd: YAG
Pulse energy	180 mJ
Repetition	20 Hz
Wavelength	532 nm
Telescope	Newtonian
Telescope diameter	0.4 m

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020

Atmospheric Measurement Techniques Discussions

Mode	coaxial
Temporal resolution	50 s
Vertical resolution	2.5 m

3. The rationale basis and the implement of the novel algorithm

3.1 Weighted k-means cluster

The CA-GM is based on the assumption that the values of turbulence and aerosol concentration are significantly higher in the NBL than in the free atmosphere (FA). Due to the influence of the multilayer structure, the peaks of the minimum gradient of RCS is the potential location of the NBL. And assembling these distinguish peaks with height through the time into groups can be considered as space and time-averaged aerosol concentration. Therefore, it can solve the inadvertently jump between different atmospheric layers. The theoretical schematic of the k-means clustering principle are shown in Figure 1. To form clusters, the Euclidean distance $dis(x_i, x_j)$ between two given signal points x_i and x_j, with coordinates (GM_i, h_i) and (GM_j, h_j), are defined as following Eq. (1).

$$dis(x_i, x_j) = [(GM_i - GM_j)^2 + (h_i - h_j)^2)]^{1/2}\ , \quad (1)$$

Where GM_i (GM_j) is the value of the gradient of RCS, and h_i (h_j) is the height of the peak.

After that, we apply the k-means cluster algorithm to classify the datasets with those notable peaks. The cluster number should be given in prior, and the k-means method build cluster iteratively by moving centroid until the target function sum of squared errors (SSE) towards local minimum (Toledo et al., 2014), the SSE calculated following Eq. (2).

$$SSE = \sum_{i=1}^{k} \sum_{x \in Ci} dis(c_i, x)^2, \quad (2)$$

Where k is a number of clusters, c_i and x is the cluster centroid and all observation in the cluster Ci,respectively. In order to obtain the accurate in compact and well-separated clusters, the criteria in the cluster validation are necessary. Davies-Bouldin index (Davies and Bouldin 1979) is conducted to the cluster validation analysis and defined as Eq. (3).

$$DB = \frac{1}{k} \sum_{i \neq j}^{k} max \left[\frac{S_k(C_i) + S_k(C_j)}{S(c_i, c_j)} \right], \quad (3)$$

Where S_k is the averaged intra-distance between of observations and their cluster centriod, $S(c_i, c_j)$ is the distance between cluster centriod c_i and c_j. The minimum DB index is considered as an optimal cluster classification.

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020

Atmospheric Measurement Techniques Discussions

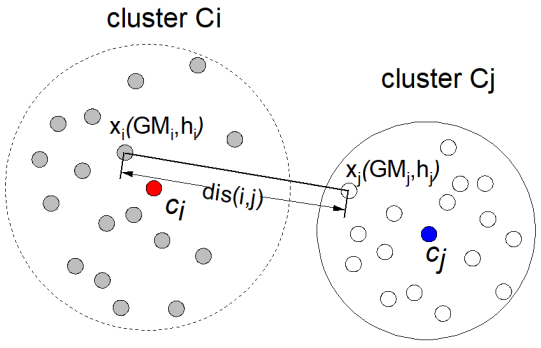


Figure 1. The theoretical schematic of the k means clustering

The standard k-means must be normalized in the case that the variable is quite a difference, data normalization is based on min-max normalization (Virmani et al., 2015). The normalized k-means clustering is "isotropic" in all directions of space and tends to capture a spherical-like shape. Nevertheless, in this paper, we proposed to put weight on height and leaving variances greater along with height. Therefore, the assembling groups of the distribution will tend to be separated along variables with greater variance, which conducive to set the top limiter altitude to classify different atmospheric layers vertically (Figure 2 b). The weighted G is calculated by the difference between maximum altitude h_{max} and the minimum altitude h_{min}, the weighted height h_w is rescaling by the normalized height data h_w as Eq.(4).

$$h_w = G * h_{nor}, \quad (4)$$

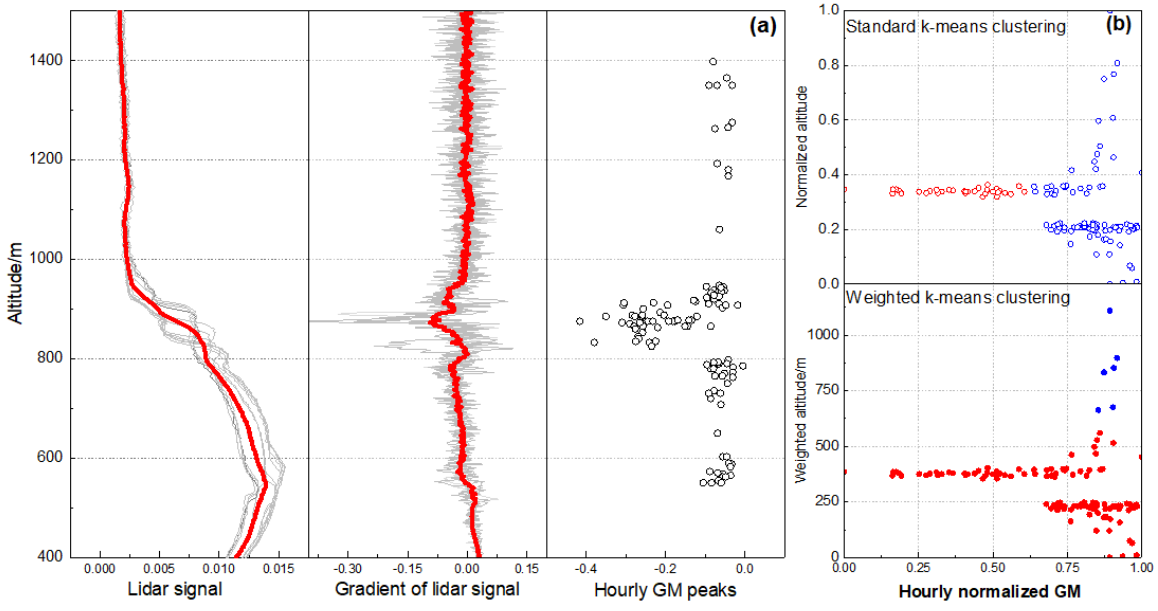


Figure 2. The theoretical schematic of the weighted-k means clustering

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020


Atmospheric Measurement Techniques Discussions

120

3.2 Multiple layers classification

As the presence of the strong gradient signature of the backscatter profile, we perform to collect 3 minima peaks of GM distribution within an hour and classify the different layers on their characteristic (Figure 2 a). The cloud layer (CL) has a lager gradient magnitude of extinction and backscattering coefficient than the aerosol layers (Palm et al., 2012). Additionally, 125 the typical nocturnal clouds are shallow cumulus (Cu), stratocumulus (Sc), and stratus (St) (Kotthaus and Grimmond, 2018). They have shallow vertical dimensions and denser than aerosols at the same altitude, hence, it can be distinguished from the aerosol layer (Wang and Sassen, 2001). Meanwhile, the noise from the GM could be affected by background and electronic noise, it has non-regular distribution and appeared at higher altitude where has lower signal-to-noise ratio. The noise layer is lack of starfield structure but have a similar GM value with the lower height. Thus, we calculate the range of vertical 130 extension of different layers, indicating the cluster significance of noise and other layers. As for elevated aerosols (EALs), the presence of EALs above the NBL represent a difficulty when retrieving the NBL top height, in particular when EALs layers are close to it. Both of the two aerosol layers have a similar characteristic of gradient variance and range of height, we discover by seeking the empirical threshold value of the EALs in backscatter coefficient (Hänel et al., 2012). The typical backscatter threshold for 532 nm wavelength lidar is defined as $\beta_{th} = 1.786 \times 10^{-3}\ \mathrm{km}^{-1} sr^{-1}$, which is calculated by 135 Ångström parameter as 1.2 in the urban-industrial and mixed condition (Dubovik et al., 2002). The gaps between NBL and EALs in the multilayer structure are determined to $D_{th} = 100\ m$ (Peng et al., 2017).

3.3 The implement of CA-GM algorithm

The CA-GM method based on k-means cluster analysis of different types of atmospheric layers is generally used to retrieve multiple layers in polluted cases. The specific ideas are shown in the flowchart in Figure 3, and the specific steps are as 140 follows, detailed results are presented as a case study.

The algorithm is divided into three parts as the pre-processing, the layer attribution classification, and the NBL inspection. The CA-GM algorithm is implemented if the data collection exceeds 30 minutes within an hour period. Firstly, The standardized lidar RCS is applied Savitzky–Golay filter to preliminary denoising. The profile of the backscatter coefficient (β) is calculated, and the reference height (h_{ref}) is limited by the Fernald method as the theoretically height-limiter. The G^* 145 is a dataset of three gradient minima of RCS. Set cluster as two in prior, implement 10 times k-means cluster to seek for the minimum DB index as the optimal grouping Ci and Cj. Second, there is a parameter D_{intra} defined as the minimum distance of inter-cluster, which can measure the cluster stratified significance to classifying the cloud and noise mixed in the G^*. If D_{intra} exceeds the threshold D_{sig}, it can distinguish noise from other layers. The S_{Ci} and S_{Cj} is as a quality control function to noise layer attribution. As for the cloud layer, the vertical extension $R_{h_{Ci}}$ ($R_{h_{Cj}}$) of the cloud is shallower than aerosol 150 layers, therefore, we defined an empirical constant C for this study. In addition, the vertical uniformity parameter Vi (Vj)

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020

Atmospheric Measurement Techniques Discussions

works as quality control tools to features of clouds and other layers as well. If cloud layer or noise exists, the original G^* are removed from the top limiter as h_{cloud} or h_{noise}, respectively. After the elimination of cloud and noise interference, the elevated aerosol layers (EALs) can be judged from the typical aerosol layer β_{th} and the gaps distance of D_{th}. Finally, the newly dataset G^* which has been removed from the different attributed layer goes to the final step with the cluster as Ci and Cj (or Ci' and Cj'). Due to the assumption of NBL distribution, the largest deviation is indicating the location of the NBLH (h_{nbl}).

In summary, through 1-2 times k-means cluster analysis of the vertical-temporal gradient of the GM within an hour, the multilayer NBL structure can be separated according to their physical characteristics of different layers. The CA-GM method is an objective and robust method for judging the attribution of different layers (NBL, EALs, CL) and noise.

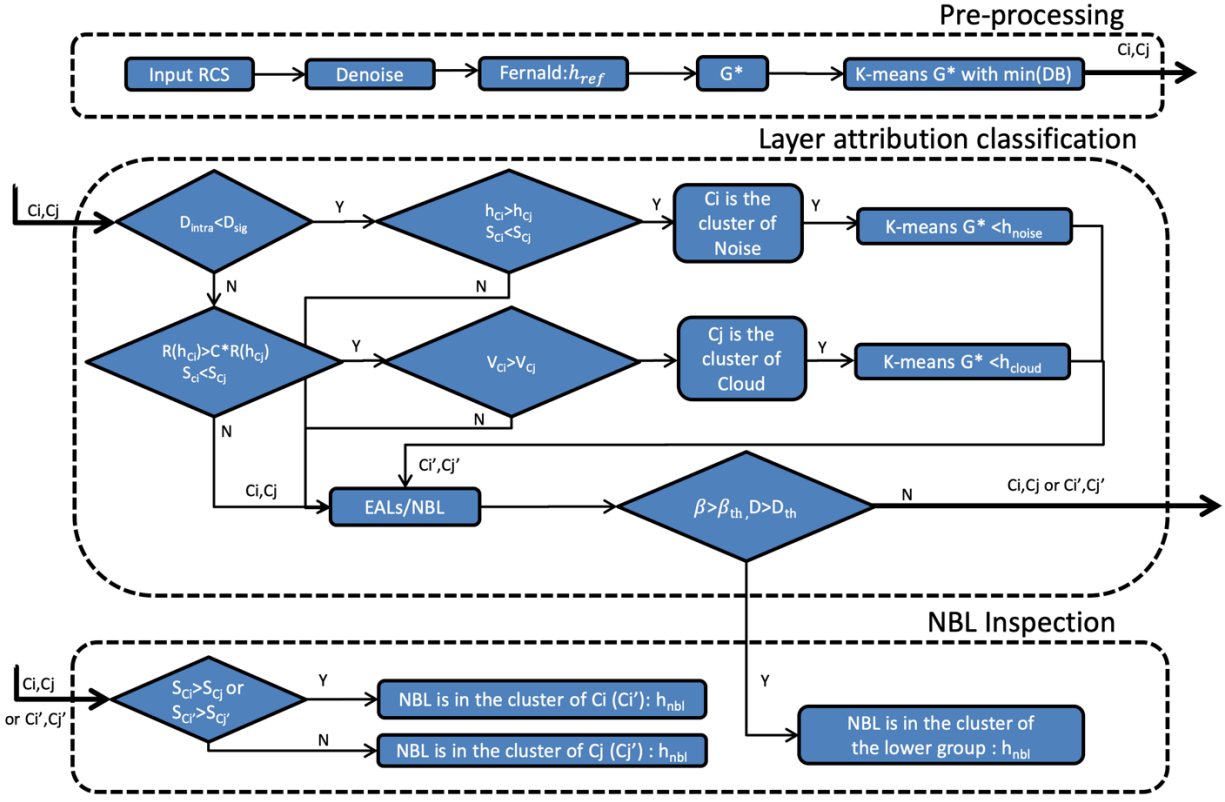


Figure 3. Flowchart of the retrieval method for the CA-GM. $h_{Ci}(h_{Cj})$: the height of centroid of cluster $Ci(Cj)$; D_{intra}: the inter-cluster distance between minimum h_{Ci} and maximum h_{Cj}; $R_{h_{Ci}}$: the intra-cluster range from the minimum h_{Ci} and maximum h_{Ci}; $S_{Ci}(S_{Cj})$: the standard deviation of the GM in the cluster $Ci(Cj)$; $Vi(Vj)$: the vertical uniformity calculated by R_{Ci} / N_{Ci} (R_{Cj} / N_{Cj}), the $N_{Ci}(N_{Cj})$ is the amount of peak in the group $Ci(Cj)$; β_{th}: typical backscatter aerosol layer ($1.786 \times 10^{-3} km^{-1} sr^{-1}$); D_{th}: threshold of distance to defined a gap between multiple aerosol layers (100 m); D_{sig}: empirical threshold as 50 m; C : empirical value as 1.5; h_{nbl}: the final location of nocturnal boundary layer height.

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020

Atmospheric Measurement Techniques Discussions

4. Evaluate and comparative analysis with classical methods

4.1 Evaluate with radiosonde data

The L-band radiosonde provided accurate thermodynamic profiles, the radiosonde-determined NBLHs were used to evaluate the accuracy of the lidar-retrieved NBLHs. Comparing with two-moment radiosonde with all algorithms, it highlighted a good agreement that the correlation coefficients (R) were range from 0.68-0.85. The CA-GM had the highest consistencies among classical methods, with the highest correlation coefficient (0.85), the weakest root mean square error (RMSE) (203 m), the smaller mean bias (28 m), and the minimum mean relative absolute difference (PRD) (17 %) (Table 2).

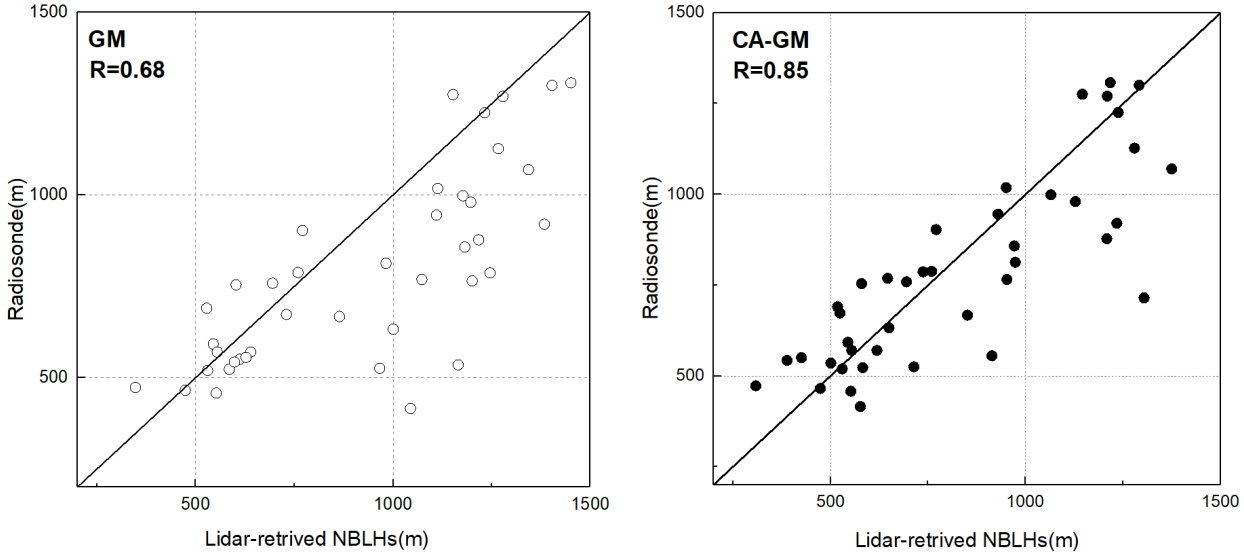


Figure 4. Comparison between the radiosonde-determined and lidar-retrieved measurement of NBLH in gradient method (GM) (hollow circle) and cluster analysis of gradient method (CA-GM) (filled circle). The correlation coefficients is represented by R. The black solid line is the 1:1 line.

The NBLH retrieved by GM and CA-GM (Figure 4) had a good correlation of the radiosonde and the latter method enhanced 25% consistency of the correlation coefficient. With the implementation of CA-GM, the data were concentrated and reduced the RMSE from 292 m to 203 m (Table 2). The means bias of GM is greater than the CA-GM, corresponding the decease of PRD from GM to CA-GM. Additionally, comparing with the WCT and CRGM, the WCT underestimated NBLH about 13 m, and the CRGM shown to overestimate the altitude as 186 m. The RMSE of CA-GM is smaller than WCT and CRGM, which was similar to the result of PRD. Therefore, the CA-GM had a good correlation with radiosonde and had the smallest fluctuation and highest consistency with NBLH retrieval.

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020

Atmospheric Measurement Techniques Discussions

Table 2. Statistic parameters of the lidar-retrieved algorithm compared with radiosonde measurement. Mean bias (MB), correlation coefficient (R), root-means-square deviation (RMSE), and the percent of relative absolute bias different (PRD) are shown below.

NBLH retrieved method	Mean Bias (m)	R	RMSE (m)	PRD (%)
Gradient method (GM)	162	0.68	292	30
Wavelet covariance transform transition method (WCT)	-13	0.80	241	21
Cubic root gradient method (CRGM)	186	0.73	277	32
Cluster analysis with gradient method （CA-GM）	28	0.85	203	17

4.2 Compared with other classical methods

To enrich our analysis, a comparison of CA-GM with GM, WCT, and CRGM in the 39-day nighttime period was applied in the supplement weakness of the rare time resolution of the radiosonde. The results were shown in Figure 5 as follows.

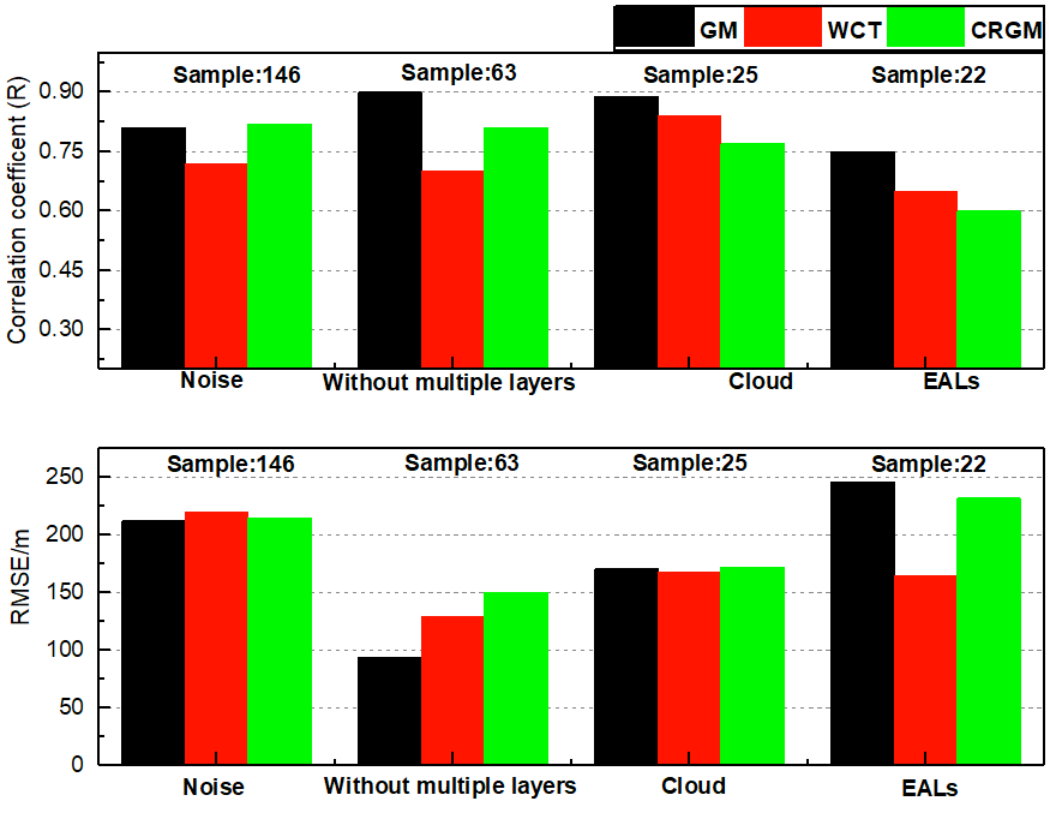


Figure 5. The result for correlation coefficient and RMSE compared with CA-GM method at all conditions (see text for details) using gradient method (GM), wavelet covariance transform transition method (WCT), and cubic root gradient method (CRGM). The sample number is shown on the top of the column, and the condition is located as the x-axis.

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020

Atmospheric Measurement Techniques Discussions

The valid CA-GM data was implemented for a total of 256 hours, the data were analyzed to compare with other retrieval algorithms. Under the condition without the infatuations of multiple layers, the CA-GM had a good correlation as 0.90, 0.70 and 0.82, for GM, WCT, and CRGM, respectively, and the RMSE was the smallest compared with other situations. Consequently, the CA-GM was more similar to the other three methods in the case of without multilayer structure, which proved reasonable of CA-GM with classical boundary layer retrieved method.

Moreover, the extensive results showed that the WCT method was more accurate than GM during nighttime (Caicedo et al., 2017) and less affected by the low signal-to-noise ratio condition (Brooks, 2003). The dilation and the threshold of WCT were selected carefully in this study (Mao et al., 2013), thus, the performance of WCT can ensuring the certify the identification of the cloud layer and noise. Notably, the improvement of the correlation coefficient from 0.70 to 0.84 in cloud contamination and from 0.70 to 0.72 in noise affection, which was conducted the better consistency to remove those attributed layers. Whereas the fluctuations in noise and cloud layers are relatively large, the CA-GM performed outstanding ability for cloud removal the eliminate noise. As for EALs, because of the ambiguous cluster of the EALs and NBL, all the methods had poor correlation coefficient with CA-GM. The observation of EALs is the most challenging part for the multiple layer structure, more active remote sensing instruments (such as multi-wavelength lidar, polarized lidar) and methods are needed to prove the accurate layout of the EALs.

Table 3 shown the criteria parameters in the CA-GM. The cluster significant parameter D_{intra} for noise was 20.75±14.62 m, which is significantly smaller among other conditions. The typical altitude of NBL, EALs, cloud, and noise in severe haze pollution is 590.49±202.84 m, 1024.69±166.36 m, 1252.52±303.28 m, and 1100.66±253.04 m, respectively. The vertical extension of the cloud layer was indeed shallower than the other layer, with the typical extension of 128.6±82.13 m. The backscatter coefficient of EALs is $1.12\pm0.76\times10^{-3}$ $km^{-1}sr^{-1}$, which was an evidence of choosing the suitable empirical value β_{th}. The cloud had the smallest value of the vertical uniformity, which indicated denser peak distribution than other layers.

Table 3. Computed criteria parameters for the layer attribution

Parameter definition	Parameter	NBL	EALs	Cloud	Noise
Cluster signification (m)	D_{intra}	119.84±83.70	103.41±87.41	198.3±86.69	20.75±14.62
Altitude (m)	h_{Ci}, h_{Cj}	590.49±202.84	1024.69±166.36	1252.52±303.28	1100.66±253.04
Vertical extension (m)	$R_{h_{Ci}}$, $R_{h_{Cj}}$	383.77±188.02	317.39±89.59	128.6±82.13	390.14±176.58
Backscatter coefficient ($km^{-1}sr^{-1}$)	β	$6.23\pm5.36\times10^{-3}$	$1.12\pm0.76\times10^{-3}$	$7.77\pm7.42\times10^{-3}$	$6.55\pm8.40\times10^{-4}$
Vertical uniformity	Vi,Vj	5.95±2.19	5.87±2.47	4.63±1.63	7.39±4.21

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020

Atmospheric Measurement Techniques Discussions

4.3 Cases study with multiple layer structure

4.3.1 Effects of cloud contamination

On 23 Dec 2016, there was a cloud layer at 1.3 km above ground level (AGL) between 18:00-23:00 LST (Figure. 6-1), which was presented as a light blue region. Below the cloud base, there was a distinct surface aerosol layer and a strong signal negative gradient, indicating as the WCT method capture. The cloud seriously influences the GM and CRGM determination and capture the upper edge of the cloud. After 21:00 LST, the cloudiness decrease, the lidar can capture the NBL signal. After defining the minimum in the upper-cluster (C_i) as the top limiter altitude, the CA-GM was able to capture the slowly increased NBLH as shown in Figure 6-1. Figure 6-2 was shown the significant two-layer structure distribution hourly on the first k-means cluster distribution. The centroid of two-layers was indicated the approximate location at 839 m and 1428 m (Figure 6-2 b), and the cloud located at the upper layer which had a shallow vertical extension and relative dense distribution. The radiosonde had a good correlation with lidar-retrieved NBLH in Figure 6-3, the PTG was shown the steepest slope at 1.37 km, but correspond the height at the location of cloud. Therefore, we selected NBLH using the RHG method as 0.78 km, which was smaller than CA-GM retrieved height at 20:00 LST.

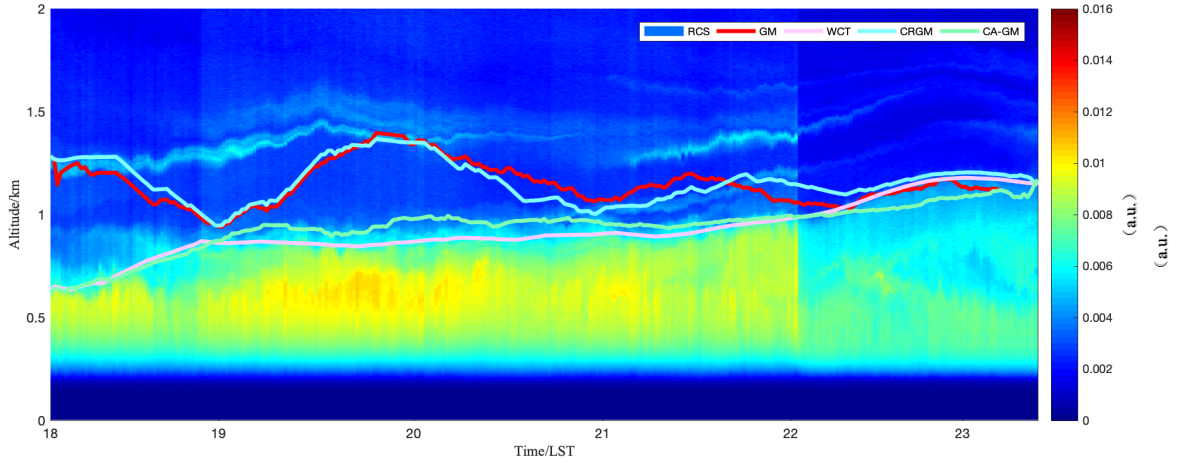


Figure 6-1. The time-height cross-section of range-corrected signal (RCS) with four NBLH retrieved methods on 23 Dec 2016.

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020

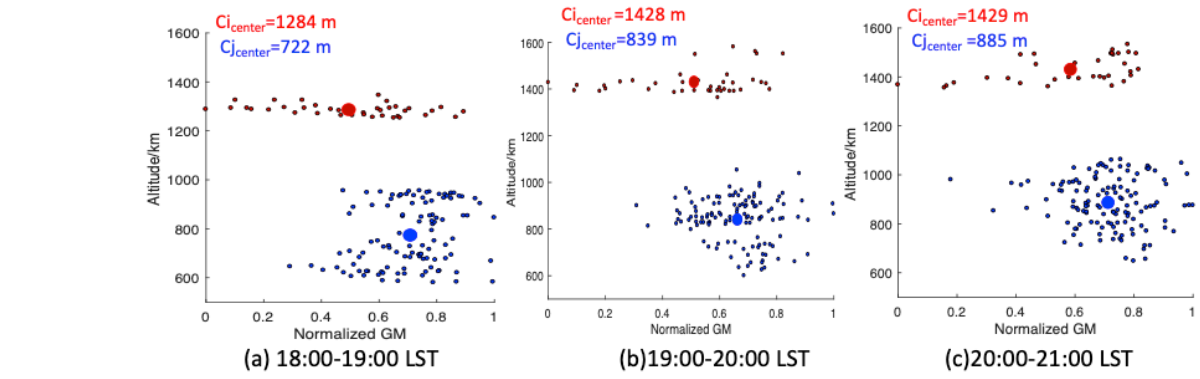


Figure 6-2. The distribution of altitude and normalized gradient method (GM) values in 18:00-21:00 LST. The Fig (a), (b) and (c) are separate hourly.

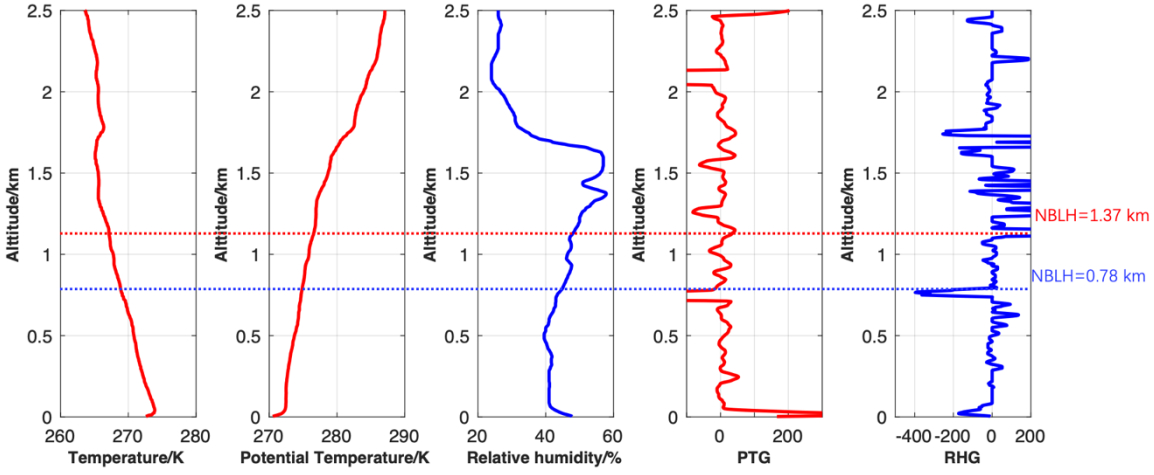
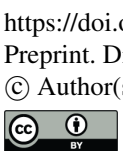


Figure 6-3.Planetary boundary layer height estimates using radiosonde. Profiles include temperature, potential temperature, relative humidity, potential temperature gradient (PTG), and relative humidity gradient (RHG). Estimated NBLH by PTG (red) and RHG (blue) was shown by dashed horizontal lines.

4.3.2 Noise affection

On 6 Apr 2017, the noise distribution was prone to appear when the low-load aerosol for the GM. The gradient-based methods were effected by noise and with a wide range of fluctuations (Figure 7-1). On the contrary, the WCT captured well the edge of the concentration of aerosols. From the distribution of GM with height distribution, the Figure 7-2 was seen the obvious mixing without starfield layer structure. Therefore, the noise was mixed in the upper layer of the centriod at height at 1479 m, 1452 m and 1173 m at 19:00, 20:00 and 21:00, respectively, which should set the top limiter and recalculate the NBLH hourly. The radiosonde data was calculated by the rapid change of PTG method as 0.79 km (Figure 7-3), corresponding to the height retrieval by with CA-GM as 0.74 km.

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020



Atmospheric Measurement Techniques Discussions

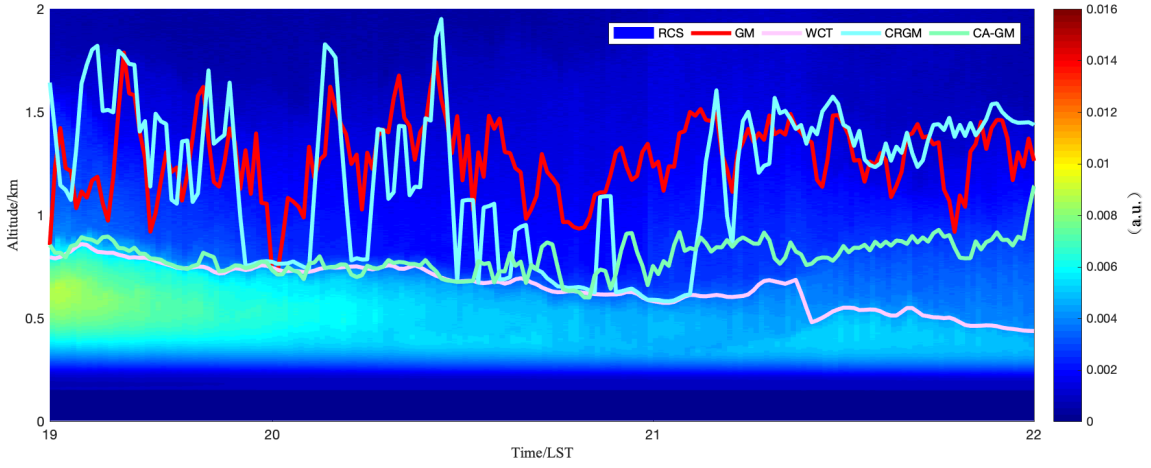
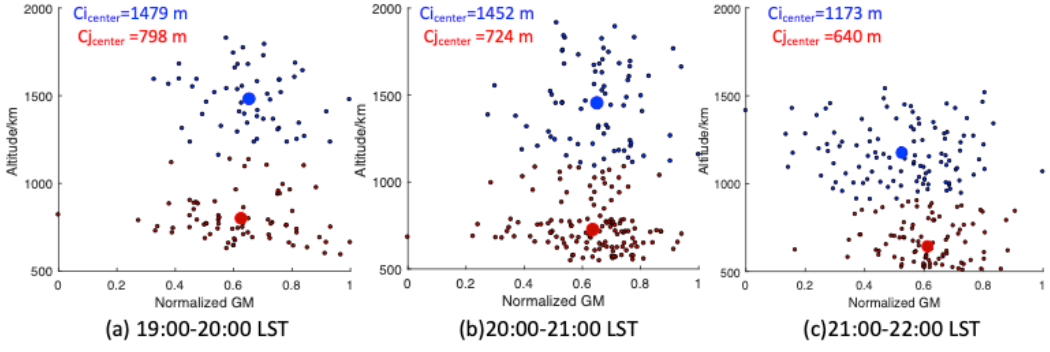


Figure 7-1 . The time-height cross-section of RCS with four NBLH retrieved methods On 6 Apr 2017.



255

Figure 7-2. The distribution of altitude and the normalized gradient method (GM) value in 19:00-22:00 LST

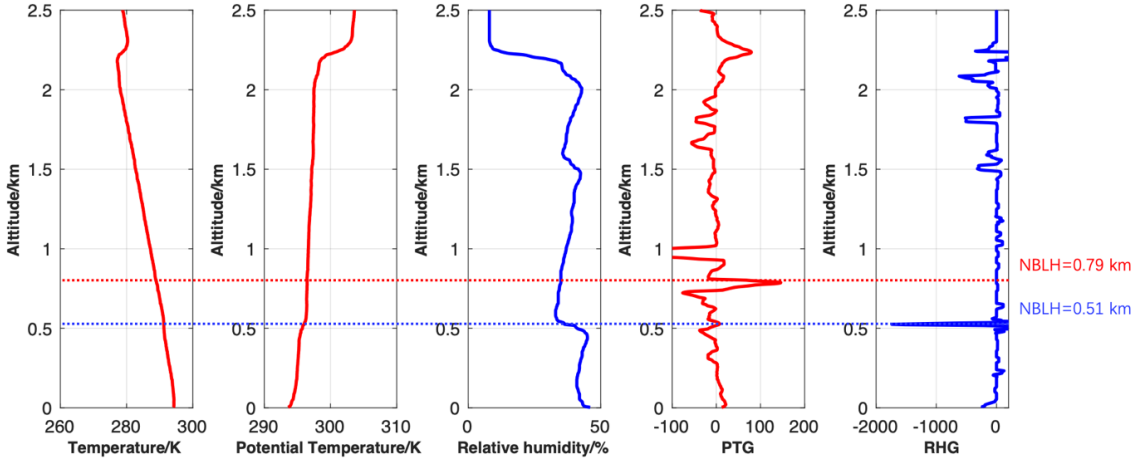


Figure 7-3 .Similar profile to Figure 6-3 on 20:00 LST

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020

Atmospheric Measurement Techniques Discussions

4.3.3 Nocturnal aloft aerosol layer

On 2 Jan 2017, the EALs frequently appeared in the lower troposphere. There was a distinct aerosol layer between 0.7-1.2 km AGL between 17:00-22:00 LST (Figure. 8-1). Without any limitation, the GM, CRGM, and WCT captured the height of EALs when the signal negative gradient at the EALs is stronger than the NBL， corresponding to the aloft aerosol structure from 17:00-22:00 LST. As the Figure 8-2 was shown, the distinct two peaks of the cluster were the two aerosol layers, deviation of the upper was lager at first and two layers gradually almost the same gradient magnitude with the time transition. The centroid of upper-cluster upward is another evidence of the NBLH with a topped EALs. After using the CA-GM method to limit the base of the lofting aerosol layers, the effect of EALs in the polluted causes can be separated successfully. Similarly, in Figure 8-3, the first gradient maxima above the surface inversion layer (SIL) is the NBLH, and both PTG and RHG had a good consistency and the NBLH is at 0.44 km. The other peak with PTG and RHG corresponding to the height of EALs.

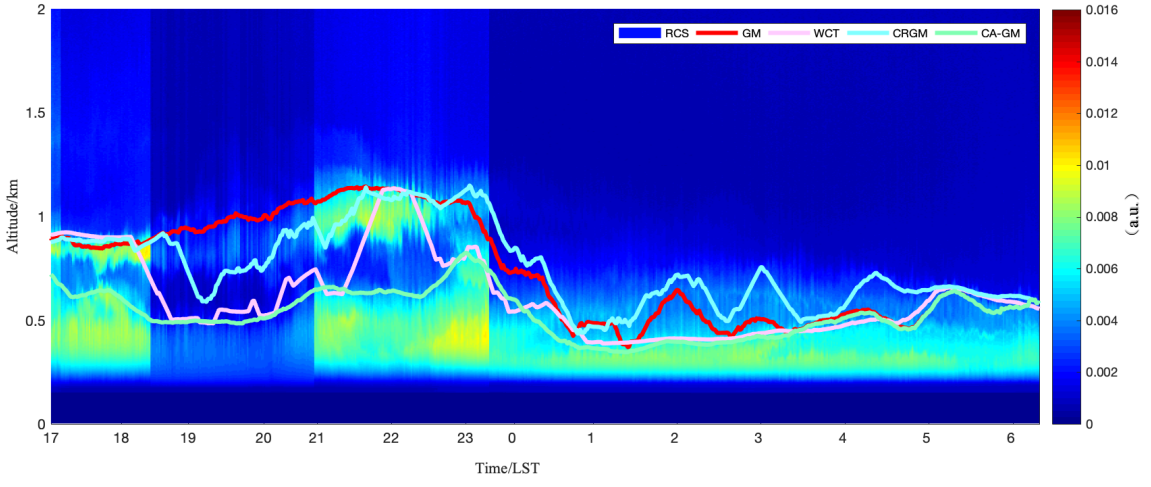
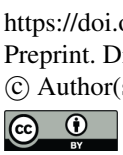


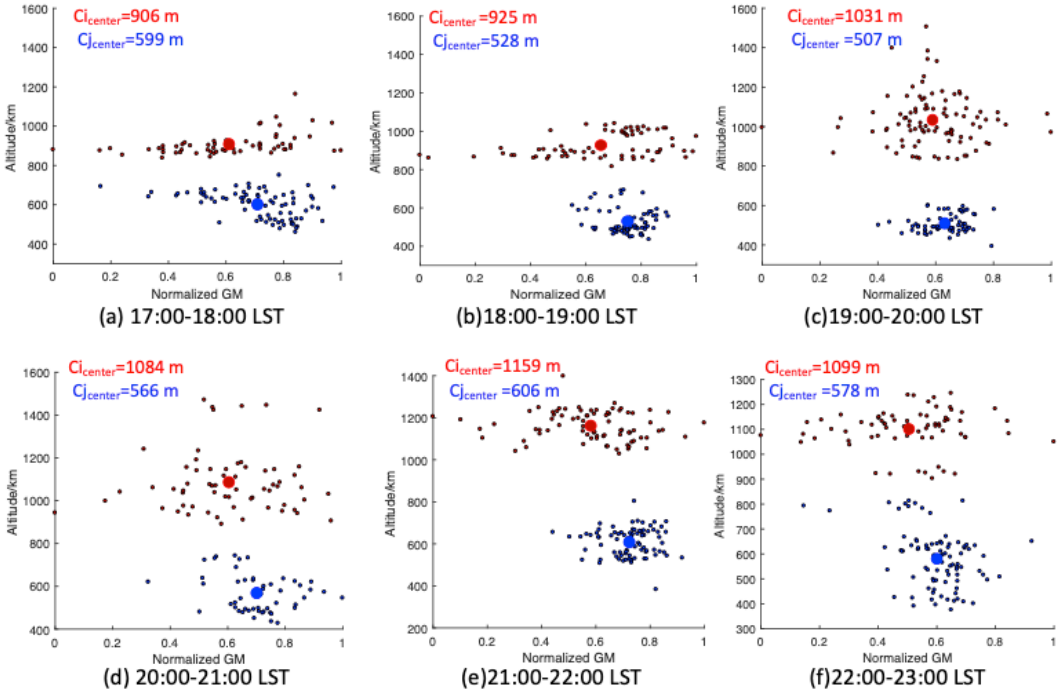
Figure 8-1. The time-height cross-section of RCS with four NBLH retrieved methods.

(The discontinuity of RCS at 18:06-18:07 is as result of detecting electric noise. The discontinuity of RCS at 20:39-20:58, 23:18-23:39 because of the adjustment laser energy and test the signal)

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020



Atmospheric Measurement Techniques Discussions



275 **Figure 8-2.The distribution of altitude and the normalized gradient method (GM) value in 16:00-23:00 LST.**

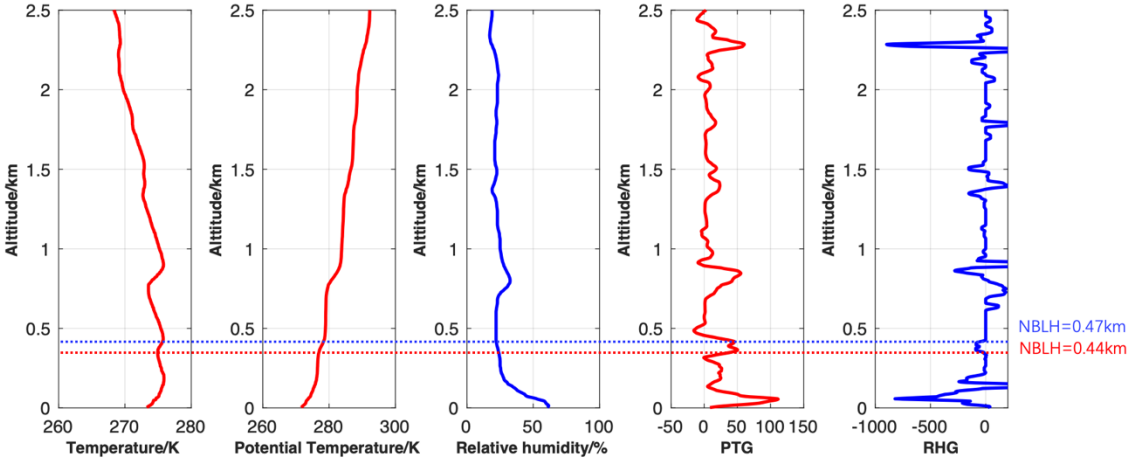


Figure 8-3. Similar profile to Figure 6-3 on 20:00 LST

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020


Atmospheric Measurement Techniques Discussions

Discussion and conclusion

The elastic lidars are superiority instruments with which to determine nighttime boundary layer height with high space and vertical resolution. Multiple layer structures in severely polluted cases impeding buoyancy forces and influence pollutants dispersed and diluted. In this paper, the novel algorithm of CA-GM was automatics developed to capture the multilayer structure and achieve stability in the nighttime with high time resolution. A 39 days heavily polluted observation experiment over Beijing (China) thoroughly evaluate the limitation of the current methods for boundary layer height determination and developed an algorithm suitable to pollution conditions.

Overall, the CA-GM highlights its high performance in comparison with radiosonde data, the best correlation (0.85), and the weakest root mean square error (203 m) and improved 25 % correlation coefficient of the GM. The possible deviations are due to the different definitions of thermodynamic NBLs from radiosondes and the aerosol NBLs. The sounding data are multi-layered as well due to the affection of aerosol layers and cloud layers, the radiosonde-retrieved NBLH should combine the PTG and RHG method to discuss the uncertainty of NBLs in the pollution period. The calculation of three minimum gradients can obverse the potential starfield layer structure, which provides a worst-case scenario for estimating surface concentrations of pollutants released into the NBL. In comparison with long-term performance with other algorithms, a reasonable parameter selection can distinguish different atmospheric layers such as cloud layer, elevated (or advected) aerosol, and random noise. The D_{intra}, R_{h_C}, β,and Vi provide a novel idea of classification multiple layers on their physical characteristic, which is more objective to automatic clustering in a complex condition. The correlation coefficient with CA-GM and WCT was elevated correlation coefficient from 0.7 to 0.84, 0.7 to 0.72 in cloud and noise affection, which proved the ability of CA-GM can ensure the upper edge of low-level cloud and remove the random noise. The EALs are often located at the top of NBLs with a similar characterise of the NBLs, thus, using the empirical threshold on a single wavelength Mie-lidar is a good way to classify EALs in polluted cases. Consequently, the CA-GM can deal with the uncertainty of the multi-layered structure and obtain a stable NBLH with a high time resolution, which expected to contribute to long-term observation of the single wavelength lidar system and micro pulse lidar monitoring in air pollution.

Code/Data availability. Contact csy@bit.edu.cn for data requests.

Author contribution. Y.Z and S.C analyzed the experimental data and co-wrote the paper. H.C and P.G supported the experiment and for maintenance of BIT-lidar. SY.C designed and lead the study. All co-authors discussed the results and commented on the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020

Atmospheric Measurement Techniques Discussions

Acknowledgements. We wish to thank the China Meteorological Administration (CMA) for providing the radiosonde data. This research is supported by the National Natural Science Foundation of China (No. 61505009).

References

Baars, H., Ansmann, A., Engelmann, R. and Althausen, D.: Continuous monitoring of the boundary-layer top with lidar, Atmos. Chem. Phys., 16, 2008.

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.

Benavent-Oltra, J. A., Román, R., Casquero-Vera, J. A., Pérez-Ramírez, D., Lyamani, H., Ortiz-Amezcua, P., Bedoya-Velásquez, A. E., de Arruda Moreira, G., Barreto, Á., Lopatin, A., Fuertes, D., Herrera, M., Torres, B., Dubovik, O., Guerrero-Rascado, J. L., Goloub, P., Olmo-Reyes, F. J. and Alados-Arboledas, L.: Different strategies to retrieve aerosol properties at night-time with the GRASP algorithm, Atmospheric Chemistry and Physics, 19(22), 14149–14171, doi:10.5194/acp-19-14149-2019, 2019.

Brooks, I. M.: Finding boundary layer top: Application of a wavelet covariance transform to lidar backscatter profiles, Journal of Atmospheric and Oceanic Technology, 20(8), 1092–1105, 2003.

Caicedo, V., Rappenglück, B., Lefer, B., Morris, G., Toledo, D. and Delgado, R.: Comparison of aerosol lidar retrieval methods for boundary layer height detection using ceilometer aerosol backscatter data, Atmospheric Measurement Techniques, 10(4), 1609–1622, doi:10.5194/amt-10-1609-2017, 2017.

Chen, H., Chen, S., Zhang, Y., Chen, H., Guo, P. and Chen, B.: Experimental determination of Raman lidar geometric form factor combining Raman and elastic return, Optics Communications, 332, 296–300, 2014.

Dang, R., Yang, Y., Hu, X.-M., Wang, Z. and Zhang, S.: A Review of Techniques for Diagnosing the Atmospheric Boundary Layer Height (ABLH) Using Aerosol Lidar Data, Remote Sensing, 11(13), 1590, doi:10.3390/rs11131590, 2019a.

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, 2019b.

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, Journal of Atmospheric and Oceanic Technology, 17(11), 1455–1468, 2000.

Devara, P. C. S., Maheskumar, R. S., Raj, P. E., Pandithurai, G. and Dani, K. K.: Recent trends in aerosol climatology and air pollution as inferred from multi-year lidar observations over a tropical urban station, International Journal of Climatology: A Journal of the Royal Meteorological Society, 22(4), 435–449, 2002.

Dong, Z., Li, Z., Yu, X., Cribb, M., Li, X. and Dai, J.: Opposite long-term trends in aerosols between low and high altitudes: a testimony to the aerosol–PBL feedback, Atmospheric Chemistry and Physics, 17(12), 7997–8009, doi:10.5194/acp-17-7997-2017, 2017.

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.

Emeis, S. and Schäfer, K.: Remote Sensing Methods to Investigate Boundary-layer Structures relevant to Air Pollution in Cities, Boundary-Layer Meteorology, 121(2), 377–385, doi:10.1007/s10546-006-9068-2, 2006.

Fernald, F. G.: Analysis of atmospheric lidar observations: some comments, Applied Optics, 23(5), 652, doi:10.1364/AO.23.000652, 1984.

Frioud, M., Mitev, V., Matthey, R., Häberli, C., Richner, H., Werner, R. and Vogt, S.: Elevated aerosol stratification above the Rhine Valley under strong anticyclonic conditions, Atmospheric Environment, 37(13), 1785–1797, doi:10.1016/S1352-2310(03)00049-9, 2003.

Garratt, J.: Review: the atmospheric boundary layer, Earth-Science Reviews, 37(1–2), 89–134, doi:10.1016/0012-8252(94)90026-4, 1994.

Guo, J., Miao, Y., Zhang, Y., Liu, H., Li, Z., Zhang, W., He, J., Lou, M., Yan, Y., Bian, L. and Zhai, P.: The climatology of planetary boundary layer height in China derived fromradiosonde and reanalysis data, Atmos. Chem. Phys., 16(20), 13309–13319, doi:10.5194/acp-16-13309-2016, 2016.

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), doi:10.1029/2012JD017577, 2012.

Hao, L., Garmash, O., Ehn, M., Miettinen, P., Massoli, P., Mikkonen, S., Jokinen, T., Roldin, P., Aalto, P. and Yli-Juuti, T.: Combined effects of boundary layer dynamics and atmospheric chemistry on aerosol composition during new particle formation periods, 2018.

Hayden, K. L., Anlauf, K. G., Hoff, R. M., Strapp, J. W., Bottenheim, J. W., Wiebe, H. A., Froude, F. A., Martin, J. B., Steyn, D. G. and McKendry, I. G.: The vertical chemical and meteorological structure of the boundary layer in the Lower Fraser Valley during Pacific '93, Atmospheric Environment, 31(14), 2089–2105, doi:10.1016/S1352-2310(96)00300-7, 1997.

Hennemuth, B. and Lammert, A.: Determination of the Atmospheric Boundary Layer Height from Radiosonde and Lidar Backscatter, Boundary-Layer Meteorology, 120(1), 181–200, doi:10.1007/s10546-005-9035-3, 2006.

Hooper, W. P. and Eloranta, E. W.: Lidar measurements of wind in the planetary boundary layer: the method, accuracy and results from joint measurements with radiosonde and kytoon, Journal of climate and applied meteorology, 25(7), 990–1001, 1986.

Ji, H., Zhang, Y., Chen, S., Chen, H. and Guo, P.: Aerosol characteristics inversion based on the improved lidar ratio profile with the ground-based rotational Raman–Mie lidar, Optics Communications, 416, 54–60, doi:10.1016/j.optcom.2018.02.003, 2018a.

Ji, H., Chen, S., Zhang, Y., Chen, H., Guo, P. and Zhao, P.: Comparison of air quality at different altitudes from multi-platform measurements in Beijing, Atmospheric Chemistry and Physics, 18(14), 10645–10653, doi:10.5194/acp-18-10645-2018, 2018b.

Kotthaus, S. and Grimmond, C. S. B.: Atmospheric boundary-layer characteristics from ceilometer measurements. Part 1: A new method to track mixed layer height and classify clouds, Quarterly Journal of the Royal Meteorological Society, 144(714), 1525–1538, doi:10.1002/qj.3299, 2018.

Li, X., Song, H., Zhai, S., Lu, S., Kong, Y., Xia, H. and Zhao, H.: Particulate matter pollution in Chinese cities: Areal-temporal variations and their relationships with meteorological conditions (2015–2017), Environmental Pollution, 246, 11–18, doi:10.1016/j.envpol.2018.11.103, 2019.

Ma, Wang, Han, Ma, Li, Gong and Chen: Regional Atmospheric Aerosol Pollution Detection Based on LiDAR Remote Sensing, Remote Sensing, 11(20), 2339, doi:10.3390/rs11202339, 2019.

Mao, F., Gong, W., Song, S. and Zhu, Z.: Determination of the boundary layer top from lidar backscatter profiles using a Haar wavelet method over Wuhan, China, Optics & Laser Technology, 49, 343–349, doi:10.1016/j.optlastec.2012.08.017, 2013.

Martucci, G., Matthey, R., Mitev, V. and Richner, H.: Comparison between backscatter lidar and radiosonde measurements of the diurnal and nocturnal stratification in the lower troposphere, Journal of Atmospheric and Oceanic Technology, 24(7), 1231–1244, 2007.

Menut, L., Flamant, C., Pelon, J. and Flamant, P. H.: Urban boundary-layer height determination from lidar measurements over the Paris area, Applied Optics, 38(6), 945, doi:10.1364/AO.38.000945, 1999.

Palm, S. P., Hart, W. D., Hlavka, D. L., Welton, E. J. and Spinhirne, J. D.: The Algorithm Theoretical Basis Document for the GLAS Atmospheric Data Products, , 6, 148, 2012.

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.

Pérez-Ramírez, D., Ruiz, B., Aceituno, J., Olmo, F. J. and Alados-Arboledas, L.: Application of Sun/star photometry to derive the aerosol optical depth, International Journal of Remote Sensing, 29(17–18), 5113–5132, doi:10.1080/01431160802036425, 2008.

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
410 situ measurements at the Jungfraujoch, Atmospheric Chemistry and Physics, 17(16), 10051–10070, doi:10.5194/acp-17-10051-2017, 2017.

Rosati, B., Herrmann, E., Bucci, S., Fierli, F., Cairo, F., Gysel, M., Tillmann, R., Größ, J., Gobbi, G. P. and Liberto, L. D.: Studying the vertical aerosol extinction coefficient by comparing in situ airborne data and elastic backscatter lidar, 2016.

Seibert, P., Beyrich, F., Gryning, S.-E., Joffre, S., Rasmussen, A. and Tercier, P.: Review and intercomparison of operational
415 methods for the determination of the mixing height, Atmospheric environment, 34(7), 1001–1027, 2000.

Seidel, D. J., Zhang, Y., Beljaars, A., Golaz, J.-C., Jacobson, A. R. and Medeiros, B.: Climatology of the planetary boundary layer over the continental United States and Europe: BOUNDARY LAYER CLIMATOLOGY: U.S. AND EUROPE, Journal of Geophysical Research: Atmospheres, 117(D17), n/a-n/a, doi:10.1029/2012JD018143, 2012.

Shi, Y., Hu, F., Fan, G. and Zhang, Z.: Multiple technical observations of the atmospheric boundary layer structure of a red-
420 alert haze episode in Beijing, Atmospheric Measurement Techniques, 12(9), 4887–4901, doi:10.5194/amt-12-4887-2019, 2019.

Steyn, D. G., Baldi, M. and Hoff, R. M.: The Detection of Mixed Layer Depth and Entrainment Zone Thickness from Lidar Backscatter Profiles., Journal of Atmospheric & Oceanic Technology, 16(7), 1999.

Stull, R. B.: An Introduction to Boundary Layer Meteorology, Springer Science & Business Media., 1988.

425 Su, T., Li, Z. and Kahn, R.: A new method to retrieve the diurnal variability of planetary boundary layer height from lidar under different thermodynamic stability conditions, Remote Sensing of Environment, 237, 111519, doi:10.1016/j.rse.2019.111519, 2020a.

Su, T., Li, Z., Li, C., Li, J., Han, W., Shen, C., Tan, W., Wei, J. and Guo, J.: The significant impact of aerosol vertical structure on lower atmosphere stability and its critical role in aerosol–planetary boundary layer (PBL)
430 interactions, Atmospheric Chemistry and Physics, 20(6), 3713–3724, doi:10.5194/acp-20-3713-2020, 2020b.

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.

Toledo, D., Córdoba-Jabonero, C., Adame, J. A., De La Morena, B. and Gil-Ojeda, M.: Estimation of the atmospheric
435 boundary layer height during different atmospheric conditions: a comparison on reliability of several methods applied to lidar measurements, International Journal of Remote Sensing, 38(11), 3203–3218, doi:10.1080/01431161.2017.1292068, 2017.

Virmani, D., Taneja, S. and Malhotra, G.: Normalization based K means Clustering Algorithm, arXiv:1503.00900 [cs] [online] Available from: http://arxiv.org/abs/1503.00900 (Accessed 21 June 2020), 2015.

440 Wang, H., Lu, K., Chen, X., Zhu, Q., Wu, Z., Wu, Y. and Sun, K.: Fast particulate nitrate formation via N 2 O 5 uptake aloft in winter in Beijing, Atmospheric Chemistry and Physics, 18(14), 10483–10495, 2018.

https://doi.org/10.5194/amt-2020-167
Preprint. Discussion started: 22 June 2020


Atmospheric Measurement Techniques Discussions

Wang, H., Li, Z., Lv, Y., Xu, H., Li, K., Li, D., Hou, W., Zheng, F., Wei, Y. and Ge, B.: Observational study of aerosol-induced impact on planetary boundary layer based on lidar and sunphotometer in Beijing, Environmental Pollution, 252, 897–906, doi:10.1016/j.envpol.2019.05.070, 2019.

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.

Weil, J. C.: Stable boundary layer modeling for air quality applications, in Air Pollution Modeling and Its Application XXI, pp. 57–61, Springer., 2011.

Yang, T., Wang, Z., Zhang, W., Gbaguidi, A., Sugimoto, N., Wang, X., Matsui, I. and Sun, Y.: Technical note: Boundary layer height determination from lidar for improving air pollution episode modeling: development of new algorithm and evaluation, Atmospheric Chemistry and Physics, 17(10), 6215–6225, doi:10.5194/acp-17-6215-2017, 2017.

Zhang, L., Wang, T., Lv, M. and Zhang, Q.: On the severe haze in Beijing during January 2013: Unraveling the effects of meteorological anomalies with WRF-Chem, Atmospheric Environment, 104, 11–21, doi:10.1016/j.atmosenv.2015.01.001, 2015.