ABSTRACT

Title of Dissertation:	REMOTE SENSING OF AEROSOL AND THE PLANETARY BOUNDARY LAYER, AND EXPLORING THEIR INTERACTIONS
	Tianning Su, Doctor of Philosophy, 2022
Dissertation directed by:	Prof. Zhanqing Li Department of Atmospheric and Oceanic Science/Earth System Science Interdisciplinary Center

Aerosol-planetary boundary layer (PBL) interaction (API) is an important mechanism affecting the thermodynamics and convection in the lower atmosphere. API plays a critical role in the formation of severe pollution events and the development of convective clouds. Despite the progress made in understanding these processes, their magnitude and significance still have large uncertainties, varying significantly with aerosol distribution, aerosol optical property, and meteorological conditions. This study attempts to develop advanced remote sensing algorithms to retrieve information about the PBL and the aerosols contained within it. These remote sensing techniques are further used to elucidate the mechanisms governing API, enhancing our ability to predict air quality and model convective clouds, as well as understand the impact of aerosols on the climate system.

In particular, we develop algorithms to improve the retrieval accuracy of aerosols and the PBL from satellite sensors and a ground-based lidar. For aerosol remote sensing, we use the deep

neural network (DNN) to construct surface reflectance relationships (SRR) between different wavelengths. We then incorporate the DNN-constrained SRR into a traditional dark-target algorithm to retrieve the aerosol optical depth (AOD) using information from a current-generation geostationary satellite, i.e., Himawari-8, as input. As a result, the performance of AOD retrievals over East Asia is significantly improved. For PBL remote sensing, we explore different techniques for retrieving the PBL height (PBLH) from both a space-borne lidar (i.e., the Cloud-Aerosol Lidar with Orthogonal Polarization) and a ground-based lidar. We further develop a new method that combines lidar-measured aerosol backscatter with a stability-dependent model of PBLH diurnal variation. The new method circumvents or alleviates an inherent limitation of lidar-based PBLH detection when a residual layer of aerosols does not change in phase with the evolving thermodynamics. By separately considering surface-cloud coupling regimes, this method also offers high-quality retrievals of PBLH under cloudy conditions.

Utilizing the enhanced retrievals of PBLH and synergistic measurements, we can also address some scientific questions concerning API, including the influencing factors of API and the role of aerosol vertical distributions. The correlation between the PBLH and the concentration of particulate matter with aerodynamic diameters less than 2.5 microns is generally negative. However, the magnitude, significance, and even the sign of their relationship vary greatly, depending on location and meteorological and aerosol conditions. In particular, API is considerably different under three aerosol vertical structure scenarios (i.e., well-mixed, decreasing and increasing with height). The vertical distribution of aerosol radiative forcing differs dramatically among the three types, with strong heating in the lower, middle, and upper PBL, respectively. Such a discrepancy in aerosol radiative forcing leads to different aerosol effects on atmospheric stability and entrainment processes. Absorbing aerosols are much less effective in

stabilizing the lower atmosphere when aerosols decrease with height than in an inverted structure scenario.

REMOTE SENSING OF AEROSOL AND THE PLANETARY BOUNDARY LAYER, AND EXPLORING THEIR INTERACTIONS

by

Tianning Su

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
2022

Advisory Committee:
Professor Zhanqing Li, Chair
Doctor Ralph Kahn
Professor Russell R. Dickerson
Professor Timothy Canty
Doctor Istvan Laszlo
Professor Michael N. Evans (Dean's representative)
Professor [your Dissertation Director], Chair

© Copyright by
Tianning Su
2022

Acknowledgements

First, I would express my sincere gratitude to my advisor, Professor Zhanqing Li, for his strong support and guidance. Professor Li gave me a broad scientific vision and trained me for doing research, writing papers, and giving presentations. Professor Li also supported me to attend many conferences and workshops which are very beneficial for my graduate research.

Moreover, I wish to thank Doctor Ralph Kahn from NASA. During graduate study, Dr. Kahn co-advised me in two studies. We have thoughtful discussions in the research. He offers excellent guidance to my research works and inspired my passions in research.

I am grateful to Doctor Istvan Laszlo. I worked with him on a study concerning AOD retrieval. He offers detailed guidance in principles, algorithms, and validation of aerosol remote sensing. His supports are very important for the completion of my graduate study. In addition, I greatly appreciate the thesis committee members, Drs. Russell Dickerson, Timothy Canty, and Michael Evans, for their guidance.

I would like to thank the many helpful discussions with Dr. Youtong Zheng. I want to thank Ms. Maureen Cribb for helping edit my papers. I also thank my group members and peers, including Lei Zhang, Jing Wei, Yuwei Zhang, Haipeng Zhang, Siyu Shan, Xiaoai Jin, Wenchao Han, etc.

Finally, I wish to express my love to my parents, Qinghua Su and Fangfang Chu. Their love and encouragement are most important to me. A special thank is given to my wife, Chuanyang Shen, for her faithful love and strongest support.

Table of Contents

Acknowledgements ii
Table of Contents iii
List of Tables v
List of Figures vi
List of Abbreviations xiv
Chapter 1 Introduction 1
1.1 Background behind aerosol-planetary boundary layer interactions 1
1.2 Limitations of satellite AOD retrievals 3
1.3 Limitations of lidar-based PBLH detections 4
1.4 The scope of this study 5
Chapter 2 Observational Datasets 7
2.1 Ground-based observations 7
2.2 Satellite data 11
Chapter 3 Satellite AOD retrievals from an improved algorithm 14
3.1 The relationship of spectral surface reflectances 14
3.2 Dark-target – Deep-learning (DTDL) algorithm 22
3.3 Evaluation of AOD retrievals 25
Chapter 4 Lidar-based PBLH detection methods 33
4.1 PBLH derived from space-borne lidar 33
4.2 PBLH derived from ground-based lidar 40
Chapter 5 Remote sensing of continental clouds coupling with surface 48
5.1 Definition of coupled and decoupled clouds based on thermodynamics 49
5.2 Lidar-based method to identify coupled and decoupled clouds 53
5.3 Retrievals of PBLH and coupling states under cloudy conditions 60
Chapter 6 Investigation of aerosol-PBL interactions 64
6.1 Relationships between PLBH and surface pollutants over China 64
6.2 The role of aerosol vertical structure in API 74

6.3 The association between PBL and air pollution during the COVID-19 lockdown in China 82

Chapter 7 Summary, future work, and perspective 91
7.1 Summary 91

7.2 Limitations of the current study 93

7.3 Future work 94

Appendices 96
Appendix A: Miscellaneous supplementary information for Chapter 3 96

Appendix B: Miscellaneous supplementary information for Chapter 4 100

Appendix C: Miscellaneous supplementary information for Chapter 5 103

Appendix D. Miscellaneous supplementary information for Chapter 6 104

Appendix E. Statement of Originality 109

Bibliography 111

List of Tables

Table 2.1. Spectral and spatial characteristics of different channels of the AHI sensor 12
Table 3.1. Empirical functions between surface reflectances (SR) at 0.47, 0.64, and 2.25 μm. Θ indicates the scattering angle. $NDVI_{SWIR}$ is the shortwave infrared normalized difference vegetation index. 17
Table 3.2. Classification of different regions for independent tests. 28
Table 5.1. List of parameters in the flow chart of DTDS (Figure 5). These parameters are related with three factors, including LCL, PBLH, CBH. The sensitivity of selection of these parameters is presented. The detailed impacts of variations in these parameters on the retrievals of cloud coupling and PBLH are presented in Figure 7 and Figure 8, respectively. 54
Table 6.1. Statistics of the partial correlation coefficients of the relationships between $PM_{2.5}$ and multiple meteorological parameters (PBLH, WS, RH, and rainfall amount) during CLD. Also shown are standardized multiple regression equations of $PM_{2.5}$ onto the meteorological parameters. Asterisks (*) denote the correlations that are statistically significant at the 95% confidence level. PBLH shows significant partial correlations with $PM_{2.5}$ over both Beijing and Northeast China. 87

List of Figures

Figure 1.1: The aerosol-PBL feedback loop for the scenarios without and with absorbing aerosols in a megacity. Black lines show the vertical profile of air temperature, and yellow dashed arrows show the reflection of solar ration by the ground surface, clouds, and s scattering aerosols. Red arrows show the absorption of solar radiation by absorbing aerosols. The blue dash-dotted line indicates the top of PBL. White arrows indicate the vertical ventilation of urban plumes induced by circulations or large eddies induced by the urban heat island effect (From Ding et al., 2016). 2

Figure 2.1: (a) Topography of China. The black rectangles outline the five regions of interest: northeast China (NEC): 40.5–50.2∘ N, 120.1–135∘ E; North China Plain (NCP): 33.8–40.3∘ N, 114.1–120.8∘ E; Pearl River Delta (PRD): 22.2–24∘ N, 111.9–115.4∘ E; and Yangtze River Delta (YRD): 27.9–33.5∘ N, 116.5–122.7∘ E. Locations of (b) environmental stations and (c) meteorological stations. (d) Blue lines indicate CALIOP daytime orbits (in ascending node). Ground-based lidar and Sun photometer are deployed in Beijing (red triangle)..... 10

Figure 2.2: Spatial distribution of terrain height of the study region. Red dots indicate the 16 AERONET sites, and pink dots indicate the 2 SONET sites used for validation. 13

Figure 3.1: (a, b) The linear regression between SR0.47 (surface reflectance at 0.47μm) and SR2.25 (surface reflectance at 2.25μm). (c, d) The linear regression between SR0.64 (surface reflectance at 0.64μm) and SR2.25. In (a, c), the color-shaded dots indicate the corresponding NDVISWIR. In (b, d), the color-shaded dots indicate the corresponding scattering angle. The regression equations and correlation coefficients are given at the top of each panel. The SR data are derived from matched ground AOD and AHI data. 17

Figure 3.2: The deep neural network (DNN) diagram to derive SR0.64. SR0.47 is calculated from SR0.64 based on a simple linear function (Table 3.1). 20

Figure 3.3: (a) Density scatterplots of the comparison between SR0.67 predicted by the empirical function and SR0.67 calculated from radiative transfer equations (3.1-3), the "true" surface reflectance. (b-e) Density scatterplots of 10-cross-validation results for SR0.67 derived from DNN considering different numbers of hidden layers in the DNN model. The correlation coefficients (R), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) are given in each panel. The solid black lines represent the linear regression, and the dashed grey lines represent the positions of 20% error. 22

Figure 3.4: Flowchart of DTDL for retrieving AOD over land. The modules enclosed by the (red) dashed line represent the updates the DTDL algorithm introduced into the NOAA/STAR DT algorithm. 24

Figure 3.5: Diurnal variation of Himawari-8/AHI AOD retrievals derived from (a) the original DT and (b) DTDL. Here, all matched pairs of AHI retrievals and ground measurements are used. The diurnal variation derived from ground-based measurements is shown in red, and the bias between AHI and ground AOD is shown in grey. The shaded areas represent the standard deviations. 26
Figure 3.6: Comparisons between AERONET AOD and AOD derived from the original DT (dark-target) method for (a) medium-quality and (b) high-quality. Comparisons between AERONET AOD and AOD derived from DTDL method for (c) medium-quality and (d) high-quality. The correlation coefficients (R), RMSE, sampling numbers (N), and MAE are given in each panel. 27
Figure 3.7: Comparisons between AERONET AOD and AOD derived from the original DT method over regions of (a) R1, (b) R2, (c) L1, (d) L2. Comparisons between AERONET AOD and AOD derived from DTDL over regions of (e) R1, (f) R2, (g) L1, (h) L2, while the training datasets are obtained over the regions of (e) R2, (f) R1, (g) L2, (h) L1, respectively. This process ensures independence of the validation datasets. The correlation coefficients, RMSE, sampling numbers, and MAE are given in each panel..... 29
Figure 3.8: Absolute biases between AOD derived from ground AOD and AHI, for different (a) SR0.47, (b) SR0.64, (c) NDVISWIR, and (d) Scattering Angle. The original DT (red lines) and DTDL (blue lines) are used respectively. The shaded areas indicate the standard deviations. The grey bars represent the occurrence frequency for these parameters. 31
Figure 4.1: (a) Total Attenuated Backscatter (TAB) plot (log scale) from CALIPSO on 24 October 2007 over Hong Kong. The black lines indicate the derived PBLH using MSD with constraint $\beta > 0$ (solid) and without constrain (dash) respectively. The gray line represents the surface elevation. (b) Vertical profile of TAB (black curve) and the vertical standard deviation (blue curve) of TAB closest to MPL site. (c) and (d) are similar to (a) and (b), except that the lines and dots indicate the derived PBLH using WCT. The curves in different color represent various PBLHs based on the assumption of various dilation a values used in WCT. The midpoints of (a) and (c) are the positions closest to MPL. (e) Time evolution of the NS plot from MPL on 24 October 2007. The black line identifies the PBLH derived from MPL and the red line represents the residual layer. (f) The NS profile of MPL average from 1300 to 1400 LT..... 36
Figure 4.2: The comparison of noontime PBLHs derived from CALIPSO and MPL at all the three stations (a), YL (b), STK (c), and PH (d). The comparison of noontime PBLHs derived from CALIPSO and radiosonde (e). The blue lines and dots are derived from MSD, while the red lines and dots are derived from WCT. 37
Figure 4.3: (a) Monthly mean diurnal variation of PBLH derived from MPL. (b) Monthly mean diurnal variation of PBLH derived from radiosonde.

The squares represent the monthly mean PBLHs derived from two approaches of CALIPSO at approximately 1330 LT. The lower rectangles indicate MSD results, while the upper ones indicate WCT results..... 39
Figure 4.4: (a) The updated correlation of PBLH derived from CALIOP and MPL after eliminating cases when difference between MSD and WCT is above 100 m. (b) The same as (a) but for the comparison between CALIPSO and radiosonde. 40
Figure 4.5: Comparisons between MPL PBLH (derived by the gradient method) and radiosonde PBLH for (a) stable (SBL), (b) neutral (NBL), and (c) convective (CBL) boundary layers during afternoon (1200-1900 LT). Comparisons between MPL PBLH (derived by the wavelet method) and radiosonde PBLH for (d) SBL, (e) NBL, and (f) CBL. 42
Figure 4.6: (a) The overview schematic diagram of retrieving the PBLH by the DTDS. In this diagram, H(i) indicates the retrieved PBLH for time i, and Hp indicates the potential PBLH position. The PBLH is retrieved at 10-minute intervals. Max indicate the maximum variation of PBLH allowed over 10-minutes, and is set as 0.2 km in this study. (b) Schematic diagram of the selection scheme. The selection scheme is a process in the retrieval algorithm that describes how an appropriate local maximum position (LMP) as Hp is selected. 44
Figure 4.7: Daily backscattering profiles for (a) SBL, (b) NBL, and (c) CBL cases. Backscattering is presented as a normalized signal on a log-scale, in arbitrary units. Black lines mark the PBLH retrieved from the DTDS. Black dots indicate the LMPs identified in the backscattering profiles. The blue stars indicate the PBLH as derived from radiosonde..... 45
Figure 4.8: Comparisons between afternoon PBLH derived from MPL and from radiosonde under different thermodynamic stability conditions. In (a, b, c), PBLH is calculated by the DTDS without the constraint of morning radiosonde data. In (d, e, f), PBLH is calculated by the DTDS with a morning radiosonde constraint..... 46
Figure 4.9: Seasonal box-and-whisker plots showing 10th, 25th, 50th, 75th, and 90th percentile values of PBLH growth rate during 0700-1230 LT derived from Radiosonde (black), gradient and wavelet methods (red), and the DTDS from MPL (blue) under SBL, NBL, and CBL conditions. 48
Figure 5.1: Idealized vertical profiles of virtual potential temperature (θv) under the clear-sky, coupled cloud, and decoupled cloud over land. The surface layer, outer layer entrainment zone, and free atmosphere are divided by the blue dash lines. The cloudy layer is marked as the shaded area, and PBLH is marked as the pink point. Red and green zones indicate the coupled and decoupled regime, respectively. Elements (e.g., turbulence, heat fluxes, cloud) in the coupled regime are directly affected by the PBL processes, while these elements are not directly affected by the PBL processes in the decoupled regime. For the coupled cases, the cloud base is below the capping inversion of entrainment

zone. For the decoupled cases, the cloud base is above the capping inversion. 50
Figure 5.2: Virtual potential temperature (θv, red lines) and water vapor mixing ratio (WVMR, blue lines) profiles from radiosonde over the Southern Great Plains site for different cases. The differences in virtual potential temperature between the cloud base and the planetary boundary layer (PBL) top are expressed as Δθv (θvCBH-θvPBLH). The time of each radiosonde launch is marked in each panel. Green regions are cloud layers, and green dashed lines indicate their boundaries. PBL heights are marked as dashed pink lines..... 51
Figure 5.3: The flow chart of the updated DTDS algorithm. In this diagram, H(i) is the retrieved planetary boundary layer height (PBLH) at time i. CBH and CTH represent the base and top heights, respectively, of the lowest cloud at time i. The PBLH part for selecting the suitable local maximum position (LMP) follows Su et al. (2020a), and a detailed scheme for identifying a coupled cloud is added to the DTDS algorithm. LCL stands for lifted condensation level. Five empirical parameters (A1, A2, A3, A4, A5) are set as 0.7, 0.2, 0.15, 1.35, 1.1, respectively..... 56
Figure 5.4: Commission errors and omission errors of coupled cloud identifications (a) for different criteria for the lifted condensation level (LCL) and (b) for different criteria for the planetary boundary layer height (PBLH). "Criteria for LCL" means coupled clouds are identified if CBH-LCL < Criteria for LCL. Similarly, "Criteria for radiosonde PBLH" means coupled clouds are identified if CBH - radiosonde PBLH < Criteria for radiosonde PBLH. The red and blue dashed lines indicate the commission and omission errors, respectively, for the DTDS algorithm. CBH stands for cloud-base height, and radiosonde stands for radiosonde. 58
Figure 5.5: Red lines indicate the mean absolute biases between PBLH derived from lidar and radiosonde for selecting different values of empirical parameters (A1, A2, A3, A4, A5) in the DTDS algorithm. Here, we only analyze the low cloud cases. Blue lines indicate the corresponding correlation coefficients between PBLH derived from lidar and radiosonde. Black dash lines indicate the default values. For each test, one parameter is variable, while other parameters are set as default values. For PBLH retrievals under cloudy conditions, A5 is the critical parameter. 59
Figure 5.6: Daily backscatter profiles: (a) short-lived coupled cloud, (b) developed coupled cloud, (c) daylong coupled cloud, and (d) active coupled cloud. Backscatter is normalized to a range of 0–1 in arbitrary units. Red dots and blue dots indicate cloud-top heights (CTHs) and cloud-base heights (CBHs) of coupled clouds. Grey dots mark CBHs for decoupled clouds. Black lines and green stars mark the planetary boundary layer height (PBLH) retrieved from the DTDS algorithm and from radiosonde soundings, respectively. White dashed lines represent lifted condensation levels (LCLs). 61

Figure 5.7: The relationships between (a) LCL and CBH, (b) CBH and radiosonde-derived PBLH, (c) CTH and radiosonde-derived PBLH for coupled clouds, and (d) DTDS-derived PBLH and radiosonde-derived PBLH. Panels (e-h) are similar to panels (a-d) but for decoupled clouds. Black lines represent the linear regressions. The linear fitting functions, correlation coefficients (R), and sampling numbers (N) are given in each panel. 63

Figure 6.1. Spatial distributions of climatological mean PBLH derived from CALIPSO for (a) March-April-May (MAM), (b) June-July-August (JJA), (c) September-October-November (SON), and (d) December-January-February (DJF) during the period 2006–2017. Spatial distributions of climatological mean of early-afternoon PBLH obtained from MERRA for (e) MAM, (f) JJA, (g) SON, and (h) DJF during the same period. 65

Figure 6.2: The relationship between CALIPSO-derived PBLH and early-afternoon $PM_{2.5}$ over (a) NCP, (b) PRD, (c) YRD, and (d) NEC. The black dots and whiskers represent the average values and standard deviation for each bin. The red dash lines indicate the regular linear regressions, and the black lines represent the inverse fit ($fx = Ax + B$). The detailed fitting functions are given at the top of each panels, along with the Pearson correlation coefficient (red) and the correlation coefficient for the inverse fit (black). Here and in the following analysis, R with asterisks indicates the correlation is statistically significant at the 99% confidence level. The color-shaded dots indicate the normalized sample density. 67

Figure 6.3: Similar to Figure 6.2, but for the relationship between CALIPSO PBLH and early-afternoon $PM_{2.5}$/AOD (unit: $\mu g\ m^{-3}$ per AOD) over four ROIs. Here, the AOD data are obtained from MODIS. 68

Figure 6.4: The relationship between CALIPSO PBLH and $PM_{2.5}$ over the NCP for (a) MAM, (b) JJA, (c) SON, and (d) DJF. (e) General relationship between $PM_{2.5}$ and PBLH aggregated over all seasons, with individual observations for each day plotted as gray dots. The box-and-whisker plots showing 10th, 25th, 50th, 75th, and 90th percentile values of $PM_{2.5}$ for each bin. The green, blue, pink, and red dots present the mean values for MAM, JJA, SON, and DJF, respectively. 69

Figure 6.5: The relationship between CALIPSO PBLH and $PM_{2.5}$ over China for (a) strong wind ($WS>4m\ s^{-1}$) and (b) weak wind ($WS<4m\ s^{-1}$). The aerosol extinction profiles at ~550 nm derived from MPL at Beijing change with different MPL-derived PBLH under (c) strong wind and (d) weak wind conditions. In (c, d), the black dots indicate the location of PBL top..... 70

Figure 6.6: (a) Relationship between wind direction/wind speed and $PM_{2.5}$ over Beijing. The red line divides the northerly wind and southerly wind. (b-c) The relationship between $PM_{2.5}$ and MPL-PBLH/ventilation rate ($VR = WS \times PBLH$, unit: $km*m\ s^{-1}$), for southerly winds over Beijing.

(d-e) The relationship between $PM_{2.5}$ and MPL-PBLH/VR, for northerly winds over Beijing. 71
Figure 6.7: Stratification by terrain elevation. The correlation coefficients (R) and slopes (unit: km*ug m^{-3}) between CALIPSO PBLH and $PM_{2.5}$ for the inverse fit ($fx = Ax + B$) are shown for the (a-b) plains and (c-d) highland areas. Noted the slope in the inverse fit is defined as $-A$. (e) The slopes in the inverse fit (i.e., linear slopes between -1PBLH and $PM_{2.5}$) under different station elevations, with color-shading indicating station mean $PM_{2.5}$ concentration. (f) Box-and-whisker plots showing the 10th, 25th, 50th, 75th, and 90th percentile values of the early-afternoon WS for plain and highland regions. The dots indicate the mean values..... 73
Figure 6.8: Normalized vertical profiles of aerosol extinction coefficient (AEC) for (a) decreasing, (b) well-mixed, and (c) increasing (i.e., inverse) aerosol structures. Red line mark the position of the PBLH, solid blue lines represent the average profiles of corresponding profiles, and error bars represent the standard deviations..... 75
Figure 6.9: The relationship between MPL-derived PBLH and $PM_{2.5}$ for (a) weakly absorbing and (c) absorbing aerosols for the decreasing aerosol structure. The relationship between MPL-derived PBLH and $PM_{2.5}$ for (b) weakly absorbing and (d) absorbing aerosols for the increasing (i.e., inverse) aerosol structure. Black lines represent the inverse fits, and the whiskers indicate the standard deviations. The fitting functions and number of samples are given in each panel, along with the correlation coefficient (R †) for the inverse fit..... 76
Figure 6.10: The averaged diurnal variations in AEC for (a) decreasing, (b) well-mixed, and (c) increasing (i.e., inverse) aerosol structures. Solid black lines indicate the averaged diurnal cycles of MPL-derived PBLH under the different aerosol structures. Dashed black lines represent the mean MPL-derived PBLH diurnal cycles. (d, e, f) The averaged diurnal variations in surface $PM_{2.5}$ under the different aerosol structures. 77
Figure 6.11: The averaged diurnal variations in aerosol radiative forcing in the vertical for (a) decreasing, (b) well-mixed, and (c) increasing (i.e., inverse) structures of aerosol loading. Solid black lines indicate the mean diurnal cycles of MPL-derived PBLH under different aerosol structures. Dashed black lines represent the mean MPL-derived PBLH diurnal cycles. 79
Figure 6.12: (a) The rate of change in buoyancy (dB/dt) in a layer of the lowest atmosphere for decreasing (blue), well-mixed (black), and inverse (red) aerosol structures during noontime. The bottom of the layer is the surface, and the rate of change in buoyancy is subjected to the top of the layer. The shaded areas show the standard deviations of the rate of change in buoyancy. (b) Box-and-whisker plots showing 10th, 25th, 50th, 75th, and 90th percentile values of the rate of change in LAB (buoyancy within lowest 1 km) during noontime. Red dots indicate the

mean values, and blue stars and pink triangles show the means for weakly absorbing (SSA > 0.9) and absorbing (SSA < 0.85) cases. 80
Figure 6.13: Schematic diagrams describing aerosol-PBL interactions (API) when decreasing and inverse aerosol structures are present. The blue dash-dotted line indicates the top of the PBL. Orange curved arrows indicate solar radiation. The background grey arrow sketches the vertical transport of humidity, aerosols, and heat. The background greyscale indicates the pollution level..... 81
Figure 6.14: Percentage differences in (a) NO_2, (b) SO_2, (c) CO, (d) $PM_{2.5}$, (e) wind speed (WS), (f) relative humidity (RH), (g) rainfall amount, and (h) planetary boundary layer height (PBLH) between mean values during the COVID-19 lockdown (26 January to 17 February 2020) and the climatological mean during the same period of the years 2016 to 2019. The PBLH is retrieved during noontime (1100–1500 Beijing Time, or BJT) due to the diurnal transition, while other parameters are averaged during daytime (0800–1900 BJT). 83
Figure 6.15: (a) $PM_{2.5}$ (red line) and NO_2 (blue line) anomalies, and (b) planetary boundary layer height (PBLH, red line), wind speed (WS, blue line), and relative humidity (RH, green line) anomalies over China. In (a) and (b), a nine-day smoothing window was applied to all anomalies. Colored solid lines represent results averaged over northern China (latitudes above 38°N), and colored dashed lines represent results averaged over central and southern China (latitudes below 38°N). Zero lines represent the climatology averaged over the same COVID lockdown (CLD) period of the years 2016 to 2019. Pink areas represent the Lunar New Year (LNY), and grey areas represent the CLD period. 85
Figure 6.16: The relationship between the planetary boundary layer height (PBLH) and $PM_{2.5}$ over (a) Beijing and (b) northeast China during wintertime. The black dots and whiskers represent average values and standard deviations in each bin, respectively. The vertical black dashed line shows the turning point of the PBLH (1 km). The thick red lines and thick blue lines indicate regressions before and after the turning point, respectively. The fitting functions and coefficient correlations of the inverse fittings are given at the top of each panel. Red and blue areas in panels (c) and (d) represent the probability density functions (PDFs) of the PBLH during the COVID lockdown (CLD) and the climatology, respectively. The red and blue dash lines represent the mean PBLH during the CLD and the climatological mean, respectively. 86
Figure 6.17: Time series of 1/PBLH (planetary boundary layer height, red lines) and $PM_{2.5}$ (blue lines) over (a) Beijing and (b) northeast China, with a three-day smoothing window. Seasonal cycles, trends, and climatology have been removed in the time series. Pink areas represent the Lunar New Year (LNY), and grey areas represent the COVID lockdown (CLD) period. The green areas indicate the periods with daily rainfall

above 1 mm. A notable haze event that occurred in Beijing during the CLD is pointed out in (a). 88
Figure 6.18: The red dot and whisker represent the average value and standard deviation of daytime $PM_{2.5}$ during the COVID lockdown (CLD) haze event in Beijing. Based on wintertime data from 2013 to 2019, days with a similar planetary boundary layer height (PBLH), wind speed (WS), relative humidity (RH), and all three are selected (less than 20% difference). The distribution of daytime $PM_{2.5}$ is presented under these similar conditions. The black dot and whisker represent the average value and standard deviation, while the width of the color-shaded areas represents the smoothed distribution of daytime $PM_{2.5}$. The mean $PM_{2.5}$ values for these categories are given at the top of the figure (unit: μg m-3). 89

List of Abbreviations

ACI = aerosol-cloud interactions
AEC=aerosol extinction coefficient
AERONET=Aerosol Robotic Network
AHI=Advanced Himawari Imager
AOD = aerosol optical depth
API = aerosol-PBL interactions
ARI = aerosol-radiation interactions
ARM = Atmospheric Radiation Measurement
CALIOP=Cloud-Aerosol Lidar with Orthogonal Polarization
CALIPSO=Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation
CBH = cloud-base height
CBL=convective boundary layer
CLD=COVID-19 lockdown
CTH= cloud-top height
COVID-19=Coronavirus Disease 2019
Cu = cumulus
CV=cross-validation
DB =Deep Blue
DNN=deep neural network Afternoon satellites constellation (A-train)
DT =Dark Target
DTDL=Dark-target – Deep-learning
LCL=lifted condensation level
LMPs=local maximum positions
LT=local time
LUT=look-up table
MAE=Mean Absolute Error
MAIAC=Multi-Angle Implementation of Atmospheric Correction
MODIS=Moderate Resolution Imaging Spectroradiometer
MPL = micropulse lidar
MSD=maximum standard deviation
NBL=neutral boundary layer
NCP=North China Plain
$NDVI_{SWIR}$=NDVI at SWIR
NEC=northeast China
NOAA=National Oceanic and Atmospheric Administration
NS=normalized signal
PBL = planetary boundary layer
PBLH = planetary boundary-layer height
PDFs=probability density functions
PM2.5=particulate matter with an aerodynamic diameter smaller than 2.5 μm
PRD=Pearl River Delta
PTD=potential temperature difference

RH=relative humidity
RMSE=Root Mean Square Error
ROI=regions of interest
SBDART=Santa Barbara DISORT Atmospheric Radiative Transfer
SBL=stable boundary layer
SGP = Southern Great Plains top-of-the-atmosphere (TOA)
SNR=signal-to-noise ratio
SONET=Sun–Sky Radiometer Observation Network
SR=surface reflectances
SRR=surface reflectances relationships
SSA= single-scattering albedo
STAR=Satellite Applications and Research
SWIR=short-wave infrared
TAB=total attenuated backscatter
VR=ventilation rate
WCT=wavelet covariance transform
WS=wind speed
YRD=Yangtze River Delta

Chapter 1 Introduction

1.1 Background behind aerosol-planetary boundary layer interactions

Aerosols critically impact Earth's climate through aerosol-cloud interactions (ACI) and aerosol-radiation interactions (ARI) and continue to be a major source of uncertainty in the quantification and interpretation of the earth's changing radiation budget (Charlson et al., 1992; Ackerman et al., 2004; Boucher et al., 2013; Li et al., 2011, 2017a; Guo et al., 2017, 2019). Despite the great advances made in measurements and modeling in past decades, it is still a challenge to accurately quantify the impact of aerosols on the climate system due to inadequate understanding of the mechanisms that control the substantial variations in aerosol type, loading, and vertical distribution (Haywood and Boucher, 2000; Carslaw et al., 2013; Li et al., 2016). Aerosols are known to interact with thermodynamic stability through ARI (Atwater, 1971; Bond et al., 2013). Light-absorbing aerosols can stabilize the atmosphere (Ramanathan et al., 2001; Ding et al., 2016; Li et al., 2017b) and also enhance convection and precipitation under certain conditions (Menon et al., 2002).

Thermodynamic stability in the planetary boundary layer (PBL) dictates the PBL development (Stull, 1988; Zhang et al., 2018), dominating the vertical dissipation of surface pollutants to some degree. Previous studies have shown that thermodynamic stability affects aerosol loading on both temporal and vertical scales (Petäjä et al., 2016; Huang et al., 2018; Z. Wang et al., 2018; Su et al., 2018). Aerosols, in turn, have important feedbacks on the stability in the PBL, depending on aerosol properties, especially those of light-absorbing aerosols (e.g., black, organic, and brown carbon). However, due to significant uncertainties in aerosol radiative forcing, it remains a challenge to quantify the impact of aerosols on thermodynamic stability, especially for different aerosol vertical distributions.

The aerosol-stability interaction serves as the cornerstone for aerosol-PBL interactions (API). The physical processes of API are also far from being well understood since the mechanism underlying the interaction between aerosols and stability remains uncertain. Conventionally, increasing the aerosol loading can reduce the surface sensible heat and stabilizes the atmosphere, suppressing PBL growth. A more stable atmosphere and lower PBL height (PBLH) will further increase the surface aerosol loading. There is thus a positive feedback loop in API, reported in

numerous studies (e.g., Wang et al., 2015; Ding et al., 2016; Petäjä et al., 2016; Dong et al., 2017; Huang et al., 2018). Figure 1.1 schematically illustrates the mechanism of this positive feedback loop.

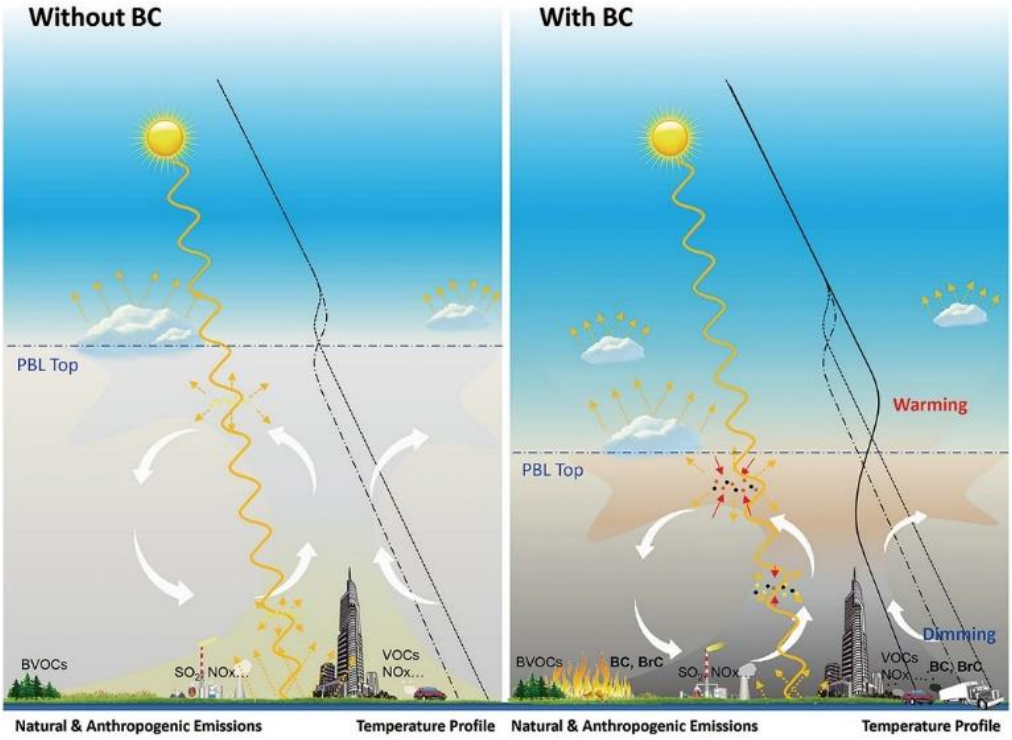


Figure 1.1: *The aerosol-PBL feedback loop for scenarios without and with absorbing aerosols (i.e., black carbon, or BC here) in a megacity. Black lines show the vertical profile of air temperature, and yellow dashed arrows show the reflection of solar radiation by the surface, clouds, and scattering aerosols. Red arrows show the absorption of solar radiation by absorbing aerosols. The blue dash-dotted lines indicate the top of the PBL. White arrows indicate the vertical ventilation of urban plumes induced by circulation or large eddies induced by the urban heat island effect (from Ding et al., 2016).*

In a recent comprehensive review, Li et al. (2017b) present ample evidence of the positive feedback loop in API and characterize determinant factors. However, such a positive feedback loop in API may not exist in all situations and is complicated by factors such as aerosol properties, the underlying surface, and meteorological conditions. Geiß et al. (2017) reported the ambiguous relationship between surface aerosol loading and the PBLH. Wang et al. (2017) classified atmospheric dispersion conditions based on the PBLH and wind speed (WS) and identified

significant surface air pollution changes that varied with dispersion conditions. Miao et al. (2017) investigated the relationship between the summertime PBLH and surface aerosol loading and discussed the impact of synoptic patterns on the development and structure of the PBL. Lou et al. (2019) showed that aerosols have a positive correlation with the PBLH under stable PBL conditions, indicating the importance of thermodynamic conditions in the PBL. However, the majority of studies considered data from only a few stations. Interactions between the PBLH and surface pollutants under different topographic and meteorological conditions thus warrant further investigation.

As absorbing aerosols stabilize the atmosphere below it but may enhance the convection above it, aerosol feedbacks to PBL are especially sensitive to the vertical distribution of aerosol amounts and absorption. Meanwhile, as the key parameter in estimating aerosol radiative effects, vertical variability in aerosol properties are especially complex and challenging to accurately estimate. Although numerical modeling is one viable method used to determine aerosol impacts on atmospheric stability and the PBL (e.g., Wang et al., 2014; Ding et al., 2016; Y. Wang et al., 2018; Zhou et al., 2018), the aerosol vertical distribution in models is generally prescribed and may differ greatly from reality. As a net result, PBL processes are difficult to be accurately estimated by numerical models (Hu et al., 2010; Davies et al., 2007), especially under severe air pollution conditions (Chu et al., 2019). Due to the limitation of model simulations, remote sensing techniques offer great potential for addressing scientific questions concerning API. This study focuses on the estimation of aerosol optical depth (AOD) and the PBLH from passive and active sensors, respectively, and employs them for improving our understanding of their interactions.

1.2 Limitations of satellite AOD retrievals

Several algorithms have been developed to retrieve AOD using data from satellite sensors, with different merits and weaknesses. A widely used algorithm, the Dark Target (DT) algorithm, has been employed in multiple sensors (Kaufman et al., 1997a; Levy et al., 2007a; Jackson et al., 2013), which has good performance over low albedo regions (e.g., dark ocean and dark vegetated surfaces) (Levy et al., 2013). Another popular algorithm, the Deep Blue (DB) or ultraviolet method, can provide retrievals over brighter surfaces with useful accuracy (Hsu et al., 2006, Hsu et al., 2013). As a relatively new method, the Multi-Angle Implementation of Atmospheric Correction

(MAIAC) has been used to retrieve AOD at the native 1-km resolution with high accuracy (Lyapustin et al., 2011, Lyapustin et al., 2018). Since the DT algorithm was mainly developed for relatively dark regions, the DB and MAIAC algorithms offer higher retrieval rates and more accurate retrievals over bright surfaces due to their different strategies (Hsu et al., 2013; Lyapustin et al., 2018). Among these methods for MODIS, the MAIAC-derived AOD has the best performance (Liu et al., 2019; Mhawish et al., 2019).

As the Center for Satellite Applications and Research (STAR), National Oceanic and Atmospheric Administration (NOAA) implemented a DT algorithm to retrieve AOD from Himawari-8/Advanced Himawari Imager (AHI) observations (Laszlo et al., 2008, Laszlo et al., 2018), we primarily focus on improving the DT algorithm, with a focus on the estimation of surface albedo in the visible spectral region where AOD is retrieved. A core part of the DT algorithms is the estimation of spectral surface reflectances (SR) in the visible channels from SR in the short-wave infrared (SWIR) channel. The usual way to accomplish this is to use empirical functions to describe the SR relationships (SRR) between the visible and SWIR channels with consideration of the scattering angle and the NDVI at SWIR ($NDVI_{SWIR}$) (e.g., Jackson et al., 2013; Kaufman et al., 1997b; Levy et al., 2007b). Due to the complexity of the SR relationships, it appears challenging to establish mathematical functions in closed form that accurately characterize the SRR. The empirical regressions for characterizing the SRR usually lead to relationships fraught with large biases, a major source of uncertainties in AOD retrievals over land.

1.3 Limitations of lidar-based PBLH detections

Numerous methods have been proposed to retrieve the PBLH from both space-borne lidar (CALIOP) and ground-based lidar, ranging from visual inspection (Boers et al., 1984), to set a signal threshold (Melfi et al., 1985), wavelet transform (Cohn and Angevine, 2000; Davis et al., 2000; Sawyer and Li 2013), fitting to an idealized profile (Steyn et al., 1999), changes in depolarization ratio (Bravo-Aranda et al., 2017), identifying maximum signal variance (Hooper and Eloranta, 1986; Su et al., 2017a), and first (Flamant et al., 1997), logarithmic (Summa et al., 2013), normalized (He et al., 2006), and cubic root (Yang et al., 2017) gradient methods, as well as the combination of the wavelet technique and image processing (Lewis et al., 2013).

Although these previous studies produced reasonable PBLH retrievals from lidar, some limitations should be noted. For the space-borne lidar, a major difficulty in using CALIOP to derive PBLH is its low signal-to-noise ratio (SNR). In particular, because of the long path length of attenuated backscattered photons, the signals reaching low altitudes can be much noisier compared with ground-based lidar (Winker et al., 2007). Moreover, previous studies utilizing CALIOP data were typically based on the maximum standard deviation (MSD) method (e.g., Jordan et al., 2010; McGrath-Spangler and Denning, 2012), whereas another widely used method, the wavelet covariance transform (WCT), has been largely ignored.

For the ground-based lidar, most PBLH algorithms are designed for convective boundary layer (CBL) conditions (e.g., Poltera et al., 2017; Yang et al., 2013; Chu et al., 2019). The afternoon PBL is typically considered convective, with a peak at ~15:00 local time. In contrast, the stable boundary layer (SBL) generally dominates at night (Nieuwstadt, 1984; Poulos et al., 2002). In addition, SBL and neutral boundary layer (NBL) can form under certain meteorological conditions during the day as well, and can occur even in the afternoon (Medeiros et al., 2005; Poulos et al., 2002; Stull, 1988; Zhang et al., 2018). However, evaluation of these algorithms for SBLs has been limited, and in particular, evaluation of lidar-based PBL detection under the full range of atmospheric stability conditions is still lacking. Moreover, as pointed out in previous studies (Chu et al., 2019; Hageli et al., 2000; Lewis et al., 2013; Su et al., 2017a), it has been a persistent problem to retrieve the PBLH under cloudy conditions, since the large backscatter and step signals from cloud interference would be excluded to avoid interfering with the retrievals.

1.4 The scope of this study

In this study, we first develop algorithms for the remote sensing of aerosols and the PBL. To address the limitations of retrieving AOD from satellite measurements, we propose using deep learning techniques to construct the SRR that accounts for the influences of multiple variables (e.g., $NDVI_{SWIR}$, viewing geometry, and time, among others). The multi-layer artificial neural network, a primitive deep learning technique, has been widely employed (Haykin, 2009; Seide et al., 2011; Cireşan et al., 2012; Liu et al., 2017) for a variety of environmental conditions and geographical locations (e.g., Deng and Yu, 2014; Tong et al., 2019; Ma et al., 2020). A DT algorithm can adopt

deep neural network (DNN)-constrained SRR, with great potential for improving AOD satellite retrievals.

To address the limitations of lidar-based PBLH detection, we develop a new method called the Different Thermo-Dynamic Stabilities (DTDS) algorithm to retrieve the PBL variability from micropulse lidar (MPL) measurements. Thermodynamic stability is represented by the near-surface potential temperature profile in the context of the boundary layer vertical structure. Long-term MPL measurements from 2010 to 2018, acquired at the Department of Energy's Atmospheric Radiation Measurement Southern Great Plains (ARM SGP) site, are used for validation. PBLH retrievals derived from the DTDS algorithm are compared with those obtained from the commonly used gradient and wavelet methods as well as with radiosonde retrievals under different thermodynamic conditions. How well PBLHs are retrieved from MPL measurements under different aerosol loading conditions is also considered. We further develop a cloud scheme to diagnose surface-cloud coupling. The surface-cloud coupling represents the turbulent coupling between surface and cloud. Under the coupled condition, low clouds are driven by surface fluxes through the conduits of the PBL over land. Thus the surface-cloud coupling have important impacts on the cloud and PBL development. By separately considering coupled and decoupled low clouds, we attempt to solve several long-lasting problems with lidar-retrieved PBLHs, namely, that retrievals cannot be generated or large uncertainties are incurred, especially for low clouds.

The two key quantities of our study, AOD and PBLH, critically impact surface particulate matter with aerodynamic diameters less than 2.5 μm ($PM_{2.5}$). Enhanced retrievals of AOD and PBLH can improve $PM_{2.5}$ monitoring and forecasting. By fusing remote sensing datasets and synergistic measurements, we address some important scientific questions concerning API, which itself is important for understanding the evolution of surface pollutants, dictated considerably by the atmospheric thermodynamic state. A radiative transfer model is used to simulate vertical profiles of aerosol radiative forcing that are employed to investigate API, critical for understanding the impacts of aerosols on the climate system.

Chapter 2 Observational Datasets

For this study we compiled a comprehensive dataset of PBL, clouds, radiation, and meteorological datasets from both satellite sensors and ground-based instruments. We made use of abundant field campaign data that provide extensive spatial coverage. Field observations are from two stations located in Beijing-China and the US Department of Energy's Southern Great Plains (SGP), respectively. The observational datasets include aerosol vertical distributions derived from lidar, optical properties derived from sun-photometers, surface fluxes from eddy covariance flux measurements, profiles of meteorological variables from radiosondes, and surface PM2.5 and meteorological parameters. Ample satellite data from several satellite sensors (AHI, MODIS, and CALIOP) are also employed.

2.1 Ground-based observations

2.1.1 Ground-based lidar systems

This study uses two ground-based lidar systems located at the SGP site (36°36′ N, 97°29′ W) and Beijing (39.99° N, 116.31° E), respectively. For the SGP site, the lidar dataset spans 1998 through 2019 with a vertical resolution of 15-30 m and a temporal resolution of 10-30 s (Campbell et al., 2002). The MPL located in Beijing was operated continuously from March 2016 to December 2018, with a temporal resolution of 15 s and a vertical resolution of 15 m. Due to incomplete laser pulses, there is a 150 m near-surface "blind zone". As standard processes, background subtraction, saturation, overlap, after-pulse, and range corrections are applied to the raw MPL data to derive the normalized signal (NS) (Campbell et al. 2002, 2003; Yan et al., 2014; Su et al., 2017b; Welton et al., 2002). There are quality-control flags along with the MPL measurements. Erroneous data are eliminated from our analysis. Lidar backscatter is further used to retrieve the vertical profiles of the aerosol extinction coefficient (AEC) (e.g., Fernald, 1984; Klett, 1985; Liu et al., 2012). As an important parameter in the retrieval process, the extinction-to-backscatter ratio (the so-called lidar ratio) is assumed to be constant within the entire column and is constrained using AERONET-derived AOD at 0.5 μm. Within the blind zone, the AEC is assumed to be equal to the value at the top of this layer. The overall uncertainties from the overlap

function, the lidar ratio, the effects of multiple scattering, and noise fall within the range of 20 %–30 % in the retrieval process (He et al., 2006).

2.1.2 Radiosonde−based observations

We use radiosonde-based observations for characterizing atmospheric thermodynamics and providing the standard estimate of PBLH. The radiosonde offers profiles of water vapor, temperature, pressure, and wind. Over China, radiosondes are launched routinely at 08:00 local time (LT) and 20:00 LT each day and are also launched at 14:00 LT in the summer (June–July–August). Meanwhile, radiosondes are launched frequently and routinely over the SGP site, usually four times per day at 00:30, 6:30, 12:30, and 18:30 LT, that may be doubled during field experiments. Radiosonde vertical resolution varies according to the balloon ascent rate, but data points occur approximately 10 m apart. We resample the radiosonde data using linear interpolation to achieve a vertical resolution of 5-hPa, which is equivalent to 30-70m within the PBL. We use a well-established method developed by Liu and Liang (2010) to retrieve the PBLH based on the gradient profile of potential temperature. Here, we focus only on PBL driven by buoyancy, and hence, low-level jet conditions are excluded, based on radiosonde wind profiles (Liu and Liang, 2010; Miao et al., 2018). Moreover, we also can use the parcel method proposed by Holzworth (1964) to retrieve the diurnal variability of PBLH. Briefly, along with the increase of surface temperature from the morning, this method assumes that the air mass is lifted adiabatically from the surface to higher altitudes and maintains a nearly constant potential temperature. This widely used method continuously tracks PBLHs, based on a morning radiosonde sounding combined with daily meteorological data (Karimian et al., 2016; Zhang et al., 2014; Li et al., 2020).

Furthermore, the PBLH can be classified as stable, neutral, or convective (SBL, NBL, or CBL) based on the thermodynamic stability conditions. PBL types are determined by calculating the near-surface potential temperature difference (PTD) between the lowest fifth and second layers above the surface (i.e., an interval of 5-hPa). Specifically, if the PTD is higher than 0.1 K and the bulk Richardson number in the lowest 0.1 km is positive, the PBL is identified as SBL; if the PTD is between −0.1 K and 0.1 K, the PBL is identified as NBL; other PBL cases are considered CBL (Tokinaga et al., 2006; Zhang et al., 2018).

2.1.3 Ground AOD measurements

In this study, we utilize ground-based AOD and single-scattering albedo (SSA) datasets obtained from two observation networks: the Aerosol Robotic Network (AERONET) and the Chinese Sun–Sky Radiometer Observation Network (SONET). The AERONET ground-based observation network of sun and sky scanning radiometer provides aerosol retrievals under clear-sky conditions. It is widely utilized to evaluate aerosol retrievals derived from satellites (Giles et al., 2019; Holben et al., 1998). With reported uncertainties of 0.01–0.02, AOD retrievals derived from AERONET are more accurate than any satellite retrieval (Eck et al., 1999) and have thus been regarded as "ground-truth." Version 3 Level 1.5 AOD and single-scattering albedos (SSAs) derived from AERONET at multiple wavelengths (i.e., 0.44, 0.5, 0.67, 0.87, and 1.02 μm) are used in our study (https://aeronet.gsfc.nasa.gov/). As the official product, the algorithm of Dubovik and King (2000) was employed to retrieve the SSA from AERONET. The principles of algorithm rely on statistical estimation. By using the spectral radiances made from different viewing angles constrained by a set of assumed aerosol characteristics, the retrieved aerosol properties are estimated as a search for the best fit of all input data by a theoretical model (Dubovik and King, 2000). Multiple scattering will be considered in the radiative transfer model. The ground-based CIMEL radiometer network in China, SONET (http://www.sonet.ac.cn/), uses the same instruments and algorithms as AERONET, with similar uncertainty of 0.01–0.02 (Li et al., 2018).

2.1.4 Cloud products from active sensors

The MPL can be used to detect cloud layers based on signal gradients (Platt et al., 1994). Lidar-based methods accurately determine the cloud-base height (CBH). Still, these methods may miss information about the cloud top due to signal saturation within optically thick clouds (Clothiaux et al., 2000). Under this condition, the cloud radar provides a better estimate of the cloud-top height (CTH). There is also a well-established cloud product over the SGP site, CLDTYPE/ARSCL (https://www.arm.gov/capabilities/vaps/cldtype), which combines information from the lidar/ceilometer and cloud radar to determine the vertical boundaries of clouds (Clothiaux et al., 2000; Flynn et al., 2017).

2.1.5 Surface observational networks of meteorology and PM2.5

As the standard instrument suite, the surface meteorological observation system routinely measures wind speed (WS), wind direction, air temperature, relative humidity, barometric pressure, and precipitation at the surface level. Meanwhile, environmental monitoring stations routinely measure PM2.5. There are surface observational networks for meteorology and PM2.5 over China (Liang et al., 2016; Su et al., 2018). In particular, the topography of China is presented in Figure 2.1a, and pink rectangles outline the four regions of interest (ROI) for the current study: northeast China (NEC), the Yangtze River Delta (YRD), the Pearl River Delta (PRD), and the North China Plain (NCP). In this study, we used ~1600 environmental stations and ~900 meteorological stations over China. The locations of these stations are presented in Figure 2.1b-c.

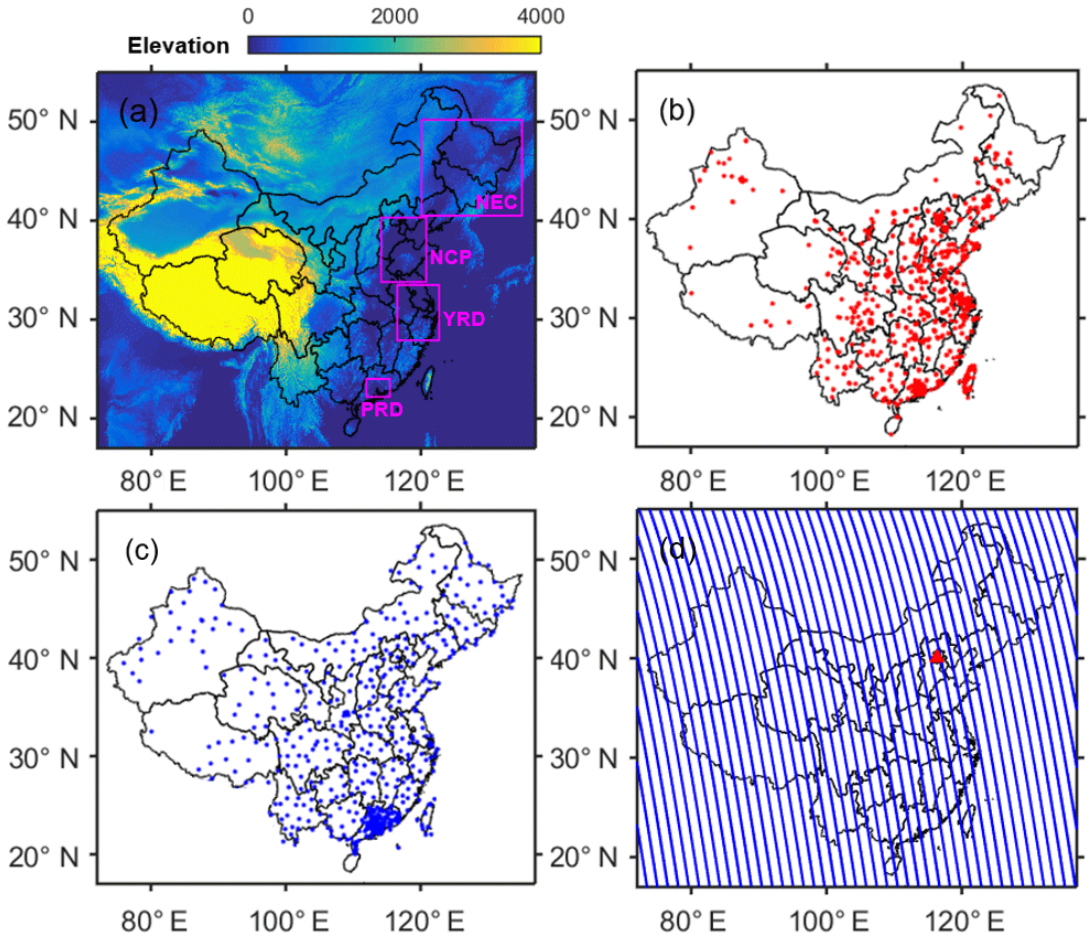


Figure 2.1: *(a) Topography of China. The pink rectangles outline the four regions of interest: northeast China (NEC): 40.5–50.2∘N, 120.1–135∘E; North China Plain (NCP): 33.8–40.3∘N, 114.1–120.8∘E; Pearl River Delta (PRD): 22.2–24∘N, 111.9–115.4∘E; and Yangtze River Delta (YRD): 27.9–33.5∘N, 116.5–122.7∘E. Locations of (b) environmental stations and (c)*

meteorological stations. (d) Blue lines indicate CALIOP daytime orbits (in ascending mode). Ground-based lidar and Sun-photometer are deployed in Beijing (red triangle).

2.2 Satellite data

2.2.1 MODIS aboard Aqua and Terra

NASA's EOS polar-orbiting satellites, Terra and Aqua, are in 705-km Sun-synchronous polar orbits between 82°N and 82°S, with equator crossings at approximately 10:30 and 13:30 local time, respectively. The MODIS instruments on board Terra and Aqua have 2330-km swath widths, and provide daily AOD data with nearly global coverage. In this study, we use the Collection 6 MODIS-Aqua Level-2 AOD products at 550 nm (available at: https://www.nasa.gov/langley, last access: 15 January 2018), which is a widely used parameter to represent the columnar aerosol amount. AOD data are archived with a nominal spatial resolution of 10 km × 10 km, and the data are averaged within a 30 km radius around the environmental stations to match surface PM2.5 data. The MODIS land AOD accuracy is reported to be within ± 0.05+15 % of AERONET AOD (Levy et al., 2010). Note that aerosol loading is significantly different in different regions. To qualitatively account for transported aerosol that is not concentrated in the PBL, we normalize the PM2.5 with MODIS AOD as the ratio between PM2.5 and AOD. A large value of this ratio indicates abundant aerosol concentrated near the surface, while a low value implies relatively more aerosols aloft that may be transported from elsewhere.

2.2.2 CALIOP aboard the CALIPSO

As a part of the Afternoon satellite constellation (A-train) (L'Ecuyer and Jiang, 2010), CALIPSO is in a 705 km Sun-synchronous polar orbit between 82°N and 82°S with a 16-day repeat cycle and with equator crossings at approximately 13:30 and 01:30 local time (Winker et al., 2007; Winker et al., 2009). CALIOP aboard the CALIPSO platform is the first long-term, space-borne lidar optimized for aerosol and cloud profiling, which has a 532 nm channel (parallel and perpendicular polarization) and a 1064 nm channel. CALIOP measures the total attenuated backscatter coefficient (TAB) with a horizontal resolution of 1/3 km and vertical resolution of 30 m in the low and middle troposphere. The attenuated backscatter data (Level 1B) is publicly available online from the Atmospheric Science Data Center at NASA Langley Research Center

(https://www.nasa.gov/langley). As shown in Figure 2.1d, blue lines represent the ground tracks over China for the daytime overpasses of CALIPSO. Because the nighttime heavy surface inversion and residual layers complicate the identification of the PBL, we only focus on daytime observations for this analysis.

2.2.3 AHI aboard Himawari-8

The current generation of geostationary satellite Himawari-8 carrying AHI was launched on 7 October 2014 and became operational on 7 July 2015. AHI is able to image East Asia every 30 seconds while offering full-disk coverage of the Pacific region (longitude: 80°E–160°W, latitude: 60°S–60°N) every 10 minutes from a geostationary orbit over the equator at ~140°E (https://www.data.jma.go.jp/mscweb/en/product/library_data). Table 2.1 summarizes the characteristics of the spectral channels of AHI. The NOAA/STAR over-land AOD algorithm uses five shortwave channels (1, 3, 4, 5, and 6) with central wavelengths of 0.47, 0.64, 0.86, 1.61, and 2.25 μm. The nominal spatial resolution of observations in these channels are, in order, 1, 0.5, 1, 2, and 2 km. To provide consistent inputs, observations from the higher-resolution channels are averaged to the lowest spatial resolution of 2 km, with a temporal resolution of 30 min. We obtained AHI data over a central part of East Asia (95–125°E, 18–30°N) in 2017 (see Figure 2.2). This study used the AOD retrievals from 18 AERONET/SONET sites to validate the satellite AOD product over the study region. We analyze data from 00:00 to 09:00 UTC, which corresponds to the daytime over the region studied.

Table 2.1. *Spectral and spatial characteristics of different channels of the AHI sensor*

Channel ID	Wavelength	Spatial resolution	Channel ID	Wavelength	Spatial resolution
1	0.47 μm	1 km	9	6.9 μm	2 km
2	0.51 μm	1 km	10	7.3 μm	2 km
3	0.64 μm	0.5 km	11	8.6 μm	2 km

4	0.86 µm	1 km	12	9.6 µm	2 km
5	1.61 µm	2 km	13	10.4 µm	2 km
6	2.25 µm	2 km	14	11.2 µm	2 km
7	3.9 µm	2 km	15	12.4 µm	2 km
8	6.2 µm	2km	16	13.3 µm	2km

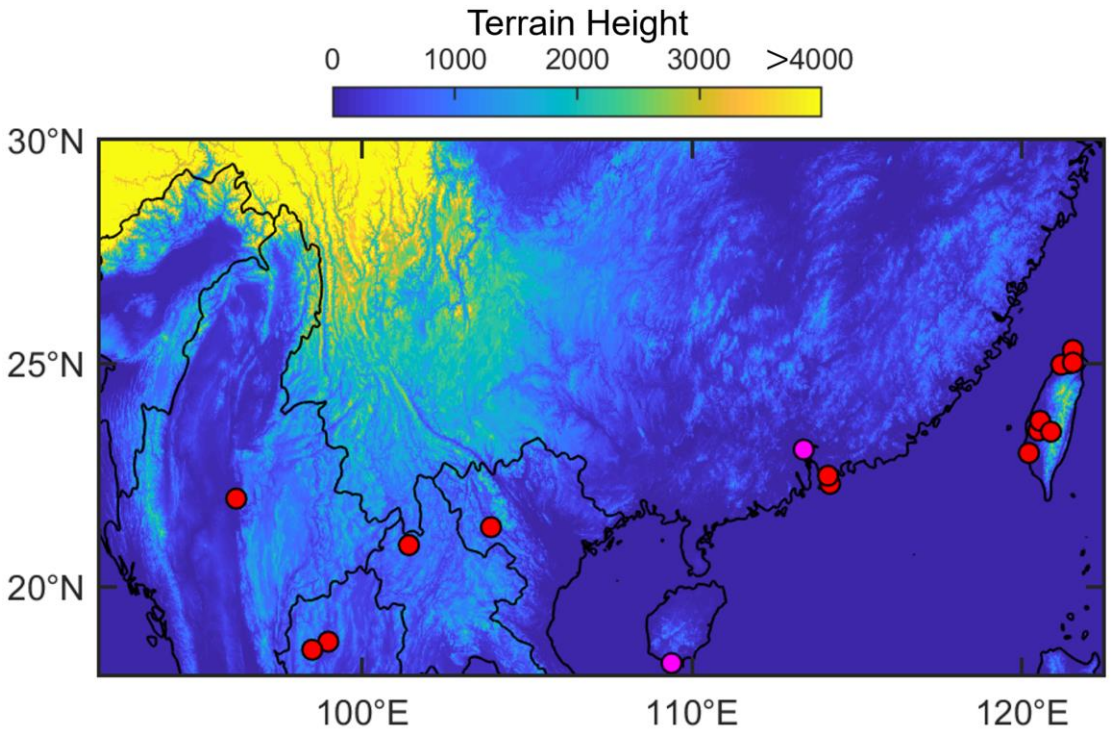


Figure 2.2: *Spatial distribution of terrain height of the study region. Red dots indicate the 16 AERONET sites, and pink dots indicate the 2 SONET sites used for validation.*

Chapter 3 Satellite AOD retrievals from an improved algorithm

In this chapter, we develop a new method that combines the traditional Dark Target (DT) approach used for MODIS and a deep learning technique to improve estimates of AOD retrievals. Specifically, the Deep Neural Network (DNN) to construct the SRR that accounts for the influences of multiple variables ($NDVI_{SWIR}$, viewing geometry, time, etc.). The DNN-constrained Surface Reflectance Relationship (SRR) has been adopted in a DT algorithm to derive the satellite retrievals of AOD. This chapter is largely based on Su et al. (2020b), and its objective is to test the application of a deep learning technique in retrieving AOD from satellites. In Section 3.1, the SRR is constructed by traditional fitting and by the DNN. Section 3.2 presents the methodology of the revised algorithm DTDL. A comprehensive evaluation of the new algorithm is demonstrated in Section 3.3.

3.1 The relationship of spectral surface reflectances

3.1.1 Estimation of SR

In this section, we describe the estimation of spectral SR needed to build the relationships. This is followed by showing "traditional" regression-based relationships and an improved relationship derived by applying a deep neural network. SRs at 0.47, 0.64, and 2.25 μm needed to determine the relationships between them are estimated from the AHI reflectances by accounting for atmospheric effects. Following Vermote et al. (1997a), we use the NOAA/STAR DT AOD algorithm that calculates the surface and atmospheric contributions to the top-of-the-atmosphere (TOA) reflectance. Assuming a Lambertian surface, contributions of the surface (ρ_{surf}) and atmosphere (ρ_{atm}) to the TOA reflectance (ρ_{toa}) are calculated as (Vermote et al., 1997a, 1997b):

$$\rho_{surf} = T^{O_3}T^{og}T^{H_2O}\left[T^{\downarrow}_{R+A}T^{\uparrow}_{R+A}\frac{\rho_{lam}}{1-S_{R+A}\rho_{lam}}\right] \quad (3.1)$$

$$\rho_{atm} = T^{O_3}T^{og}\left[\left(\rho_{R+A}-\rho_R(P_0)\right)T^{\frac{1}{2}H_2O}+\rho_R(P)\right] \quad (3.2)$$

$$\rho_{toa} = \rho_{atm}+\rho_{surf} \quad (3.3)$$

where T^{og} indicates the transmittance from absorption contributed by gases other than water vapor and ozone; T^{O_3} represents the transmittance from ozone absorption; T^{H_2O} and $T^{\frac{1}{2}H_2O}$ are the transmittances from total and half column water vapor absorption, respectively; $T_{R+A}^{\downarrow}/T_{R+A}^{\uparrow}$ indicate the total (i.e., direct plus diffuse) downward/upward atmospheric transmissions; S_{R+A} represents the atmospheric spherical albedo; $\rho_R(P)$ represents the Rayleigh reflectance contributed by molecular scattering at the real surface pressure (P); ρ_{R+A} represents the path reflectance by molecules and aerosols at the standard surface pressure (P_0); and ρ_{lam} is the Lambertian land surface reflectance. When everything needed to calculate ρ_{atm} is known, ρ_{surf} is determined from Eq. (3.3). Equation (3.2) is then solved for the surface reflectance ρ_{lam}. We apply this process at the AERONET/SONET sites where AOD is known from ground-based measurements.

Radiative properties of aerosols (normalized extinction coefficient, phase function, and single-scattering albedo) required to calculate reflectances and transmittances are currently prescribed by one of four candidate aerosol models (generic, urban, smoke, and dust) identified by the NOAA/STAR DT algorithm as part of the AOD retrieval. We recognize that this aerosol model may not be optimal for two reasons. First, the candidate models only represent broad categories of aerosol properties. For a particular retrieval, the actual aerosol properties likely differ from those prescribed by any of the candidate models. Second, since the model selection is influenced by the SRR used in the DT algorithm, uncertainties in SRR contribute to the uncertainty in the model selection, especially when the 550-nm AOD is larger than about 0.5, i.e., when the radiative properties of candidate aerosol models differ the most from each other. For the dataset used in this study, this applies to about 30% of the samples. The approximate threshold of 0.5 applies only for the generic and urban models and for the normalized extinction of the smoke model. Radiative properties of dust and the single-scattering albedo and phase function of smoke are different from those of the generic and urban models, even at optical depths smaller than 0.5. The impact of erroneous model selection is expected to be somewhat mitigated by screening out low-quality retrievals, which, among other factors, are associated with poor agreement between the observed and calculated TOA spectral reflectances; this can happen when the aerosol model picked by the retrieval is likely in error. Even though the calculated SR suffers from multiple sources of uncertainties (e.g., the Lambertian assumption, aerosol optical properties, and aerosol

vertical distribution), it is still considered suitable for retrieving AOD because the same assumptions are made in the AOD retrieval. Thus the SRs calculated from Eqs. (3.1–3) are considered as "ground truth".

Calculations of T^{H_2O} and T^{O_3} require the column amounts of water vapor and ozone, respectively. These parameters, along with other meteorological parameters needed in Eq. (3.1), are obtained from the GFS. Laszlo et al. (2018a) provide details of the quality control process and the calculation of all gaseous transmittances, including T^{og} and Rayleigh reflectance. In constructing the database used for establishing SRR, all available 2-km SR retrievals are averaged within 10 km around the ground sites. As

3.1.2 SRR from regression

The absorption of liquid water and chlorophyll at visible (e.g., blue or red bands) and SWIR wavelengths are related to the state of the vegetation. The relationships of spectral SR are also affected by the vegetation state, as well as seasonality (Kaufman and Remer, 1994; Kobayashi et al., 2007; Remer et al., 2001). The state of vegetation is usually characterized by the top-of-canopy NDVI. However, because this NDVI requires knowledge of AOD, $NDVI_{SWIR}$ is used as a relatively aerosol-independent measure of the vegetation state, in practice. It is defined as (Levy et al., 2007b):

$$NDVI_{SWIR} = (\rho^m_{1.61} - \rho^m_{2.25})/(\rho^m_{1.61} + \rho^m_{2.25}) \qquad (3.4)$$

where $\rho^m_{1.61}$ and $\rho^m_{2.25}$ are the AHI-measured reflectances at 1.61 μm (channel 5) and 2.25 μm (channel 6), respectively.

Figure 3.1 presents the scatterplots of $SR_{0.47}$ (Figure 3.1a-b) and $SR_{0.64}$ (Figure 3.1c-d) as a function of $SR_{2.25}$, with colors representing values of $NDVI_{SWIR}$ (left panels) and scattering angle (right panels). Hereafter, SR_λ represents the SR at wavelength λ (i.e., 0.47, 0.64, and 2.25 μm). Despite the scatter, higher $SR_{2.25}$ is generally associated with lower $NDVI_{SWIR}$ and higher scattering angles. The simplest linear regressions in the form of Y = *m* + *n* X, where *m* and *n* are constants, are also shown. These simple linear regressions cannot fully characterize the SRR between visible and SWIR channels, so it is necessary to add $NDVI_{SWIR}$ and scattering angle into the empirical relationship between SR in multiple channels (Laszlo et al., 2018a; Remer et al., 2013). Table 3.1 presents the SRR obtained this way and used in this study as one of the possible

empirical functions to represent the SRR. The regression coefficients are derived from the least-squares method for multiple parameters (Bühlmann and Van De Geer, 2011).

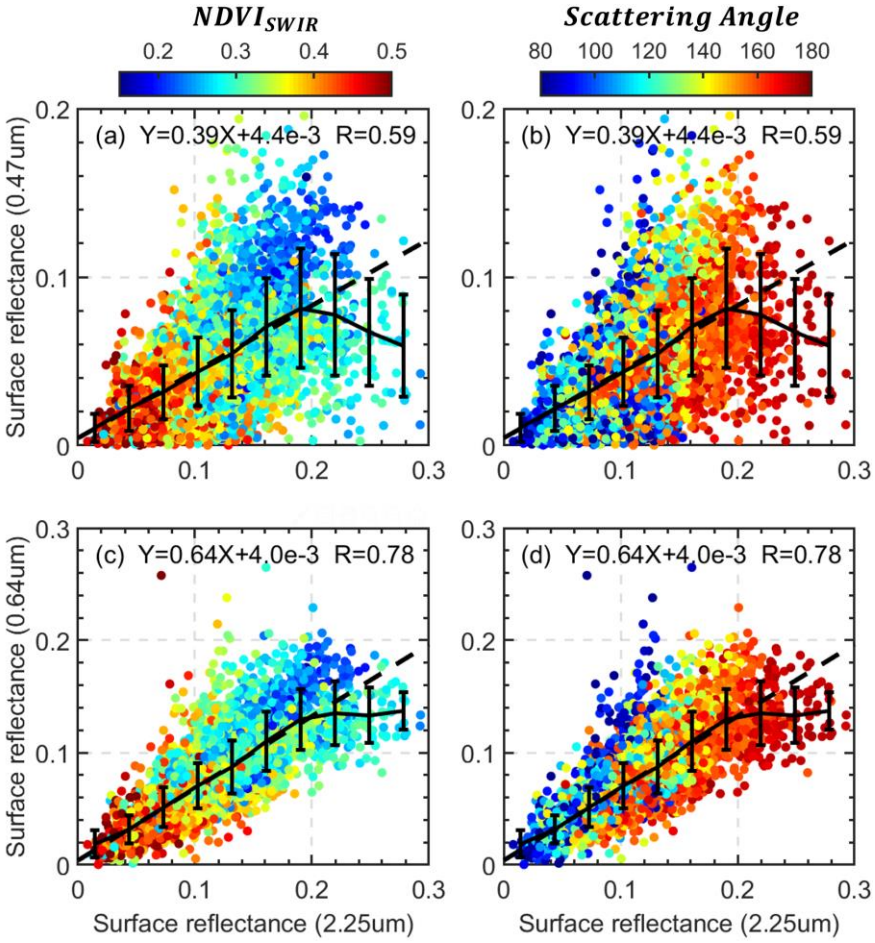


Figure 3.1: *(a, b) The linear regression between $SR_{0.47}$ (surface reflectance at 0.47μm) and $SR_{2.25}$ (surface reflectance at 2.25μm). (c, d) The linear regression between $SR_{0.64}$ (surface reflectance at 0.64μm) and $SR_{2.25}$. In (a, c), the color-shaded dots indicate the corresponding $NDVI_{SWIR}$. In (b, d), the color-shaded dots indicate the corresponding scattering angle. The regression equations and correlation coefficients are given at the top of each panel. The SR data are derived from matched ground AOD and AHI data.*

Table 3.1. *Empirical functions between surface reflectances (*SR*) at 0.47, 0.64, and 2.25 μm. Θ indicates the scattering angle.* $NDVI_{SWIR}$ *is the shortwave infrared normalized difference vegetation index.*

$NDVI_{SWIR}$ range	$SR_{0.64} = m + n \times SR_{2.25}$ $m = a + b \times \Theta;\ n = c + d \times \Theta$ $SR_{0.47} = 0.86 \times SR_{0.64} - 0.02$			
	a	***b***	***c***	***d***
$NDVI_{SWIR}$<0.2	-0.0416	0.00058	0.68	-0.00095
0.2≤ $NDVI_{SWIR}$<0.3	-0.038	0.00060	0.9	-0.0036
0.3≤ $NDVI_{SWIR}$<0.4	-0.026	0.00026	1.21	-0.0045
0.4≤ $NDVI_{SWIR}$<0.5	-0.0147	0.000067	1.32	-0.0048
0.5≤ $NDVI_{SWIR}$	0.0077	-0.00011	0.91	-0.0019

3.1.3 SRR from DNN

We adopted the DNN to derive SR at the visible channel (0.64 μm) from SR at the SWIR channel (2.25 μm). The DNN is designed based on the structure and function of the human nervous system and brain, which contain a large number of interconnected individual elements (i.e., artificial neurons) working in parallel (Cireşan et al., 2012; Sarle, 1994; Seide et al., 2011). Using a sufficient dataset to train the model, a DNN model can provide a predicted output for new input data following the pattern it learned from the training data (Schmidhuber, 2015).

Instead of specific functions, we use the DNN to train the relationship between $SR_{0.64}$ and $SR_{2.25}$ based on measurements within 10 km around the AERONET/SONET sites. Based on the previous discussions, we add additional inputs, including $NDVI_{SWIR}$, scattering angle, TOA reflectances at 0.47/0.64/2.25 μm, seasonality, column amount of ozone, column amount of water vapor, and surface pressure. Meteorological parameters are obtained from the NCEP GFS. The seasonality is classified as winter (December–January–February), spring (March–April–May), summer (June–July–August), and autumn (September–October–November). The four seasons are

denoted as 1, 2, 3, and 4, respectively. The seasonality is then used in the DNN training and prediction. Due to the nonlinearity in the relationships among various parameters, the DNN model is based on the Bayesian regularization approach for tracking various nonlinear functions (Burden and Winkler, 2008). The maximum number of epochs is set as 10,000 to allow sufficient training. As shown in Figure 3.2, the corresponding input data are extracted from a geographical grid. The input data are then interconnected with multiple neurons within the hidden layers. Figure 3.2 also shows the architecture of neurons in the hidden layers. The total number of hidden layers is n+3, where n can be adjusted as needed. The first and second hidden layers have 8 and 4 neurons, respectively, and the last hidden layer has one neuron. Other hidden layers have two neurons. The neuron is a processing element that sums the inputs and weights. The strengths and importance of the connections are represented by neuron biases and weights, which are automatically adjusted in the training process. The predicted $SR_{0.64}$ would then be generated for a given grid based on the training model.

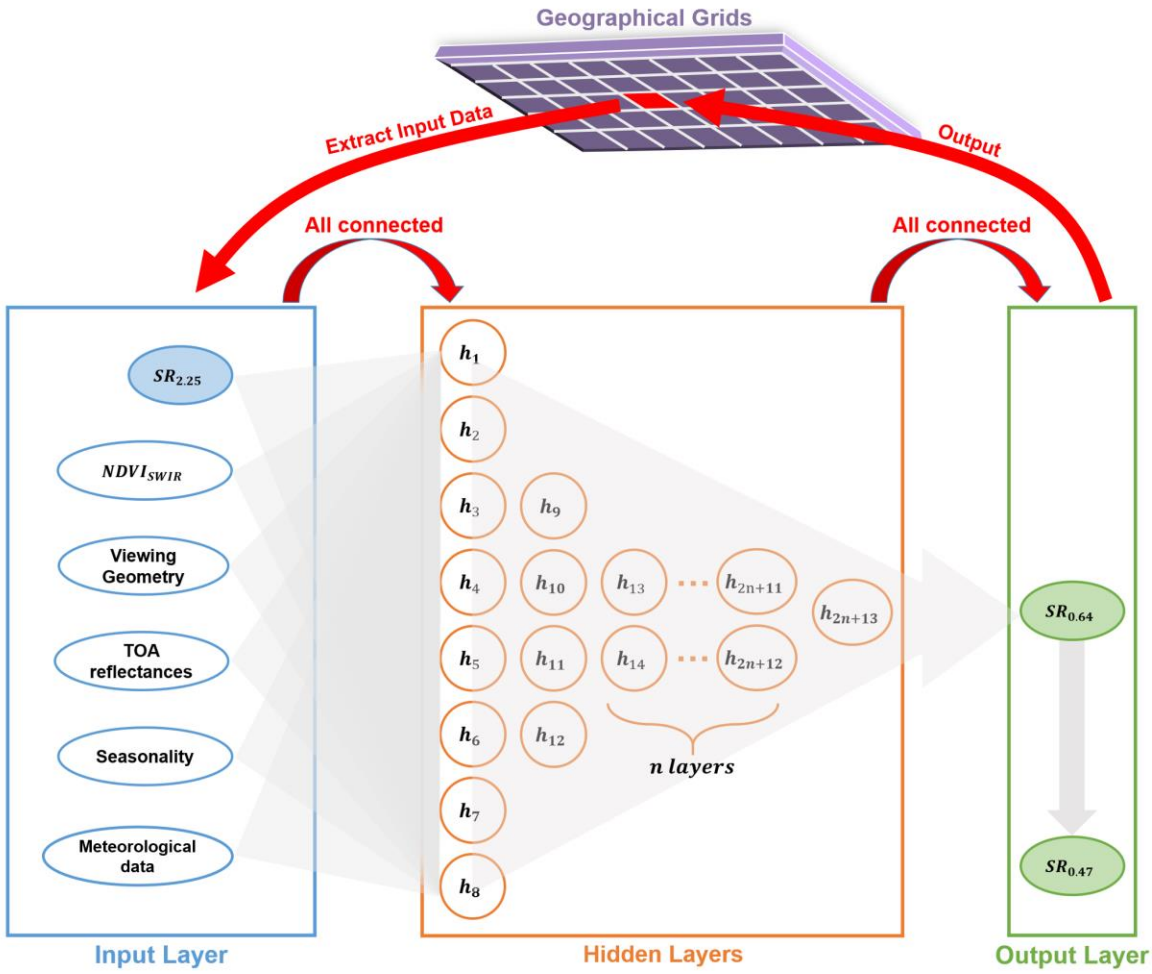


Figure 3.2: *The deep neural network (DNN) diagram to derive* $SR_{0.64}$. $SR_{0.47}$ *is calculated from* $SR_{0.64}$ *based on a simple linear function (Table 3.1).*

3.1.4 The performance of SRR derived from regression and DNN

The performance of the empirical function in the traditional approach is shown in Figure 3.3a that plots the 0.64-μm SR estimated from the empirical function as a function of the "true" 0.64-μm SR obtained from the AHI TOA reflectance and the ground-observed AOD using Eqs. (3.1–3). The plot shows a fair amount of scattering for individual retrievals that could lead to considerable errors in AOD. The empirical function is also likely to underestimate at the high end and overestimate at the low end of the "true" SR values.

For evaluating the performance of the DNN model, we use the sample-based cross-validation (CV) technique (Rodriguez et al., 2010). For the CV process, we randomly choose 70%

of the total samples for training, then utilize the remaining 30% of the dataset for validation. Such a procedure is repeated ten times to avoid biases in the selection. Basically, the data sample size for validation is expanded ten times, while the samples do not overlap with the training data. Figure 3.3b-e present the scatterplots of the 10-CV results of $SR_{0.64}$ for different numbers of hidden layers with the architecture shown in Figure 3.2. Samples of 10-CV are three times the original dataset. The number of neurons is a critical parameter in the DNN but does not significantly affect the DNN performance in this study. Despite a slightly larger bias when the number of hidden layers is three, DNN-derived $SR_{0.64}$ shows consistent performance for different numbers of neurons. With the smallest root-mean-square error (RMSE), we set the total number of the hidden layers to 7 and the total number of neurons to 27. In the architecture constructed in our study, the DNN predicts $SR_{0.64}$ with RMSEs and mean absolute errors (MAEs) that are about a third of those predicted by the empirical function, as shown in Figure 3.3.

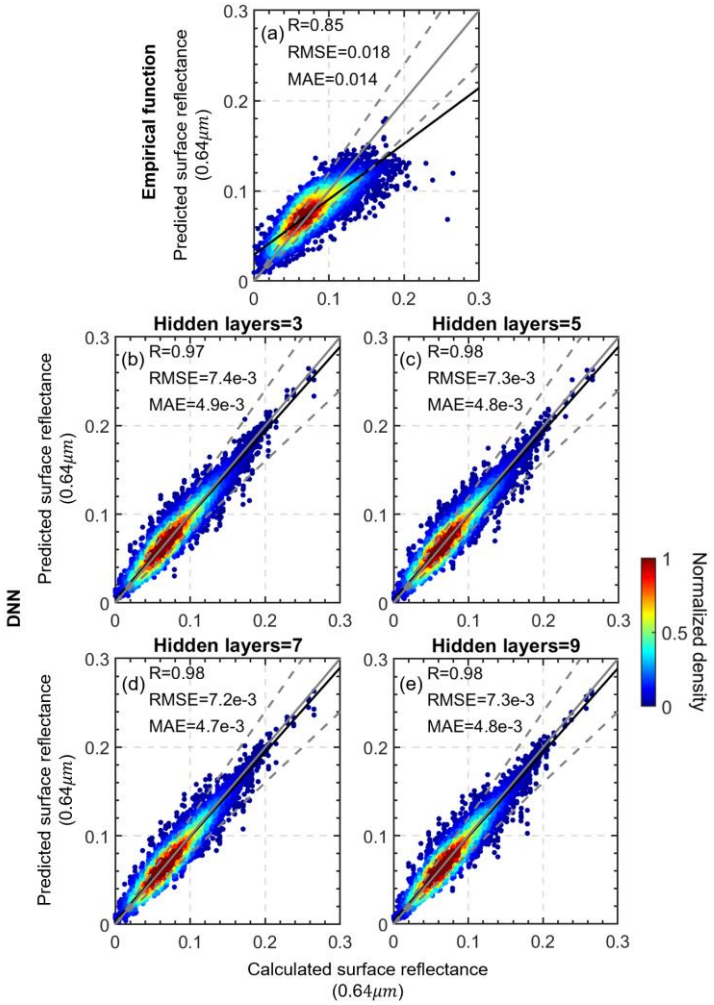


Figure 3.3: *(a) Density scatterplots of the comparison between $SR_{0.67}$ predicted by the empirical function and $SR_{0.67}$ calculated from radiative transfer equations (3.1-3), the "true" surface reflectance. (b-e) Density scatterplots of 10-cross-validation results for $SR_{0.67}$ derived from DNN considering different numbers of hidden layers in the DNN model. The correlation coefficients (R), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) are given in each panel. The solid black lines represent the linear regression, and the dashed grey lines represent the positions of 20% error.*

Following previous studies (e.g., Gupta et al., 2016), we still use an empirical linear relationship to calculate $SR_{0.47}$ from $SR_{0.64}$ (Figure A.1 in *appendix* A). The linear regression predicting $SR_{0.47}$ from $SR_{0.64}$ captures the overall, dominant characteristics of the true $SR_{0.47}$ and results in a relatively good correlation coefficient (0.9). But the residual error of the linear fit is still considerable because of the large scatter in the data. In the current stage, we do not use the DNN to train the relationship between $SR_{0.47}$ from $SR_{0.64}$ because we found that the DNN may predict unphysical values of $SR_{0.47}$ based on preliminary analyses (e.g., smaller than 0 or larger than 0.3, the maximum value expected for the vegetated surface at this wavelengths). The unphysical values may indicate the DNN does not fully capture the SSR at visible channels under some conditions. However, we do not rule out the possibility that the DNN may improve the performance of retrieved $SR_{0.47}$ with appropriate inputs and architecture. Based on the linear fitting of SR, we use a simple regression ($SR_{0.47} = 0.86 \times SR_{0.64} - 0.02$) to derive $SR_{0.47}$ from the DNN-derived $SR_{0.64}$. To avoid negative values, $SR_{0.47}$ is set to 0.01 when $SR_{0.64}$ is less than 0.025.

3.2 Dark-target – Deep-learning (DTDL) algorithm

3.2.1 The algorithm combining the DNN and the DT method

We implemented the DNN scheme for the SRR into the NOAA/STAR DT algorithm for Himawari-8/AHI. This DT algorithm follows the approach applied to the GOES-R Advanced Baseline Imager, which is described in more detail by Laszlo et al. (2018a). For this DT algorithm, the land aerosol models are adopted from previous studies for MODIS (Levy et al., 2007a; Remer et al., 2006). There are notable uncertainties in the AOD retrievals associated with assumed aerosol

models, which may lead to uncertainties of more than 10% in the retrievals (Jeong et al., 2005; Mielonen et al., 2011; Tirelli et al., 2015; Wu et al., 2016). This is very difficult to resolve as we cannot actually obtain or retrieve the real aerosol model from monodirectional, multispectral intensity measurements alone. Moreover, different types of aerosols are mixed in the real atmosphere and show large vertical variations. Such complex information may not be conveyed by some common aerosol models (Li et al., 2020). Unlike oceans, the surface properties of land demonstrate large variability and great complexity, making modeling the spectral SR difficult. The bias in SR further serves as a major source of uncertainties in the aerosol retrieval over land. The SRR trained by the DNN is expected to mitigate the problem.

The retrieval algorithm can be described as searching for the pair of AOD and SR that leads to calculated TOA reflectances that best fit the observed TOA reflectances at 0.47, 0.64, and 2.25 μm. Instead of directly using a radiative transfer model (i.e., 6SV v1.1), a look-up table (LUT) of necessary terms for several candidate aerosol models, AOD values, and geometries was generated, based on the 6S vector radiative transfer model (Vermote et al., 1997a) for efficient calculations. Figure 3.4 outlines the procedure of simultaneous AOD retrievals and SR. For each aerosol model, an iterative procedure is presented by looping over the AOD values in the LUT in ascending order. For a specific step i in the loop, a value of $SR_{2.25}$ is calculated from the AOD and the TOA reflectance in the SWIR channel. $SR_{0.64}$ is then directly estimated through the DNN-trained SRR described in Section 3.1.3. $SR_{0.47}$ is calculated from the DNN-derived $SR_{0.64}$, based on linear regression. The TOA reflectance at 0.47-μm ($\rho_{0.47,i}$) is derived from $\tau_{550,i}$ and $SR_{0.47}$. The loop over the AOD values is terminated when $\rho_{0.47,i}$ converges to the observation ($\rho_{0.47}^{obs}$). For the current aerosol model, the AOD retrieval is further estimated based on linear interpolation. Furthermore, the corresponding TOA reflectance at 0.64 μm is calculated from the current AOD retrieval and $SR_{0.64}$. The associated residual is calculated as the squared difference between the observed TOA reflectance and the calculated TOA reflectance at 0.64 μm. By looping the aerosol model, the AOD retrieval is selected as the solution with the smallest residual.

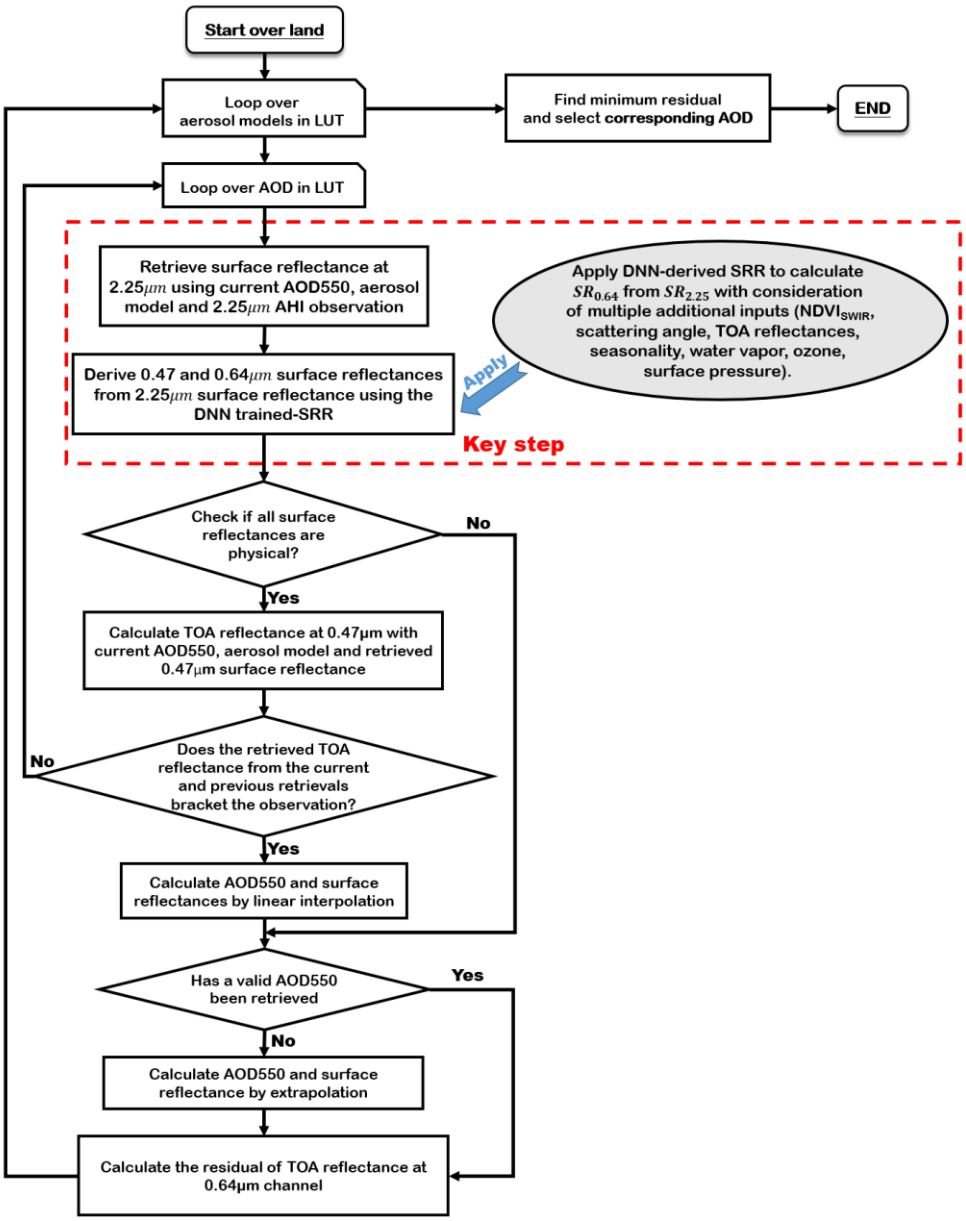


Figure 3.4: *Flowchart of DTDL for retrieving AOD over land. The modules enclosed by the (red) dashed line represent the updates the DTDL algorithm introduced into the NOAA/STAR DT algorithm.*

3.2.2 Quality control

In the NOAA/STAR DT algorithm, several external masks and internal tests are applied to screen out unsuitable pixels and control the quality of AOD retrievals. Details of the procedures

are provided by Laszlo et al. (2018a), but here, we summarize the major ones for completeness. First, pixels with $\rho_{0.47}^{obs} > 0.4$ are deemed as cloudy or too bright for the aerosol retrieval and are excluded. An internal snow test and ephemeral water test are then applied, following previous studies (Jackson et al., 2013; Walton et al., 1998). An internal inhomogeneity test is also performed to further screen pixels. Detailed information on these tests is provided in the publicly available algorithm theoretical basis documents referenced by Laszlo et al. (2018a). Based on the retrieval residual, external masks, and internal tests, the AOD retrievals are classified into four quality levels (i.e., no retrieval, low quality, medium quality, and high quality). Laszlo et al. (2018a) summarizes the criteria for determining the quality level of AOD retrievals. Cases with no retrieval and low quality are excluded from our analyses.

3.3 Evaluation of AOD retrievals

3..3.1 Diurnal variations in AODs

We also compare the diurnal variation in AODs retrieved from Himawari-8/AHI measurements and those derived from ground measurements (Figure 3.5). These diurnal variations are averaged from data for the year 2017. AHI retrievals matched with ground measurements from 18 sites are used to avoid biases due to sampling differences caused by quality control. The ensemble of all matched data is then used in the analyses of the AOD diurnal variations. In general, aerosol loading based on ground data slightly increases between 00:00–02:00 UTC, reaching a maximum value around 02:00 UTC, and gradually decreasing thereafter. The AHI AOD retrievals derived from the original DT method have a very different diurnal pattern. DT-derived AODs generally decrease until about 02:00 UTC, reaching a maximum around 06:00 UTC. As a result, the largest bias exists around 02:00 UTC when the ground AOD reaches its maximum, and the DT-derived AOD is much lower. However, the mean bias of DT is not always larger than the DTDL bias. For example, it is slightly smaller than that of the DTDL at 05:00 UTC.

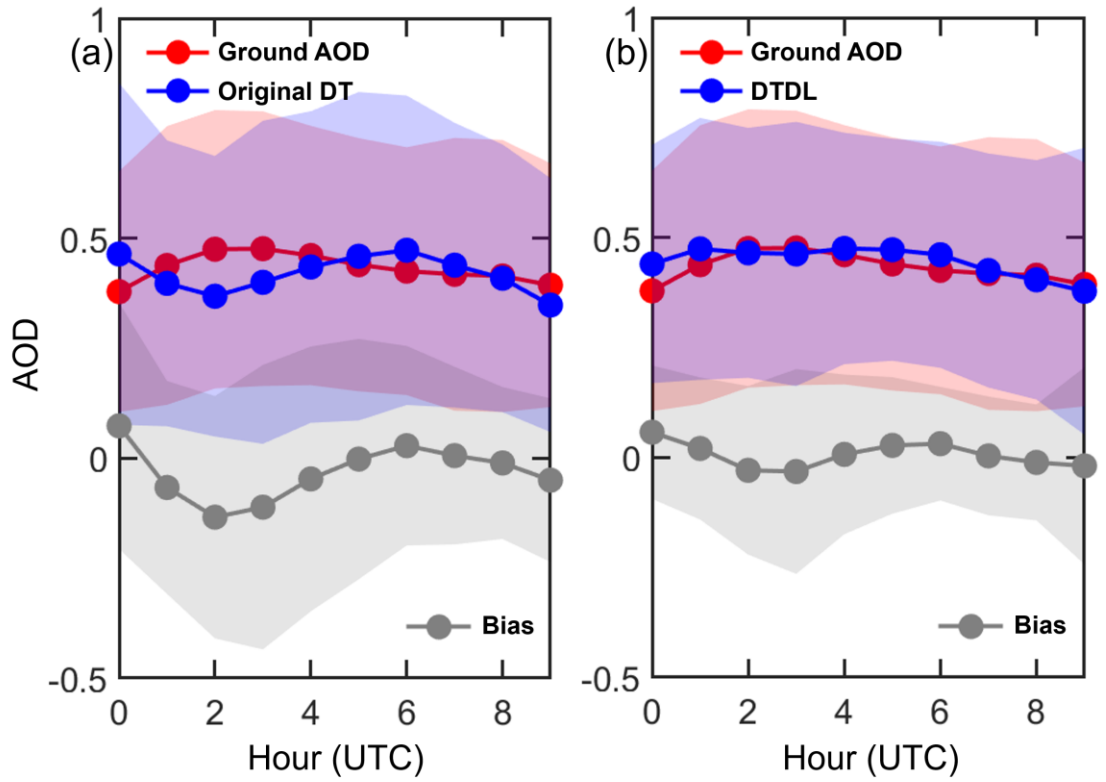


Figure 3.5: *Diurnal variation of Himawari-8/AHI AOD retrievals derived from (a) the original DT and (b) DTDL. Here, all matched pairs of AHI retrievals and ground measurements are used. The diurnal variation derived from ground-based measurements is shown in red, and the bias between AHI and ground AOD is shown in grey. The shaded areas represent the standard deviations.*

The diurnal biases for the original DT and the DTDL algorithms during different seasons are presented in Figure A.2. The considerable mean bias of the DT retrievals around 02:00 UTC is likely caused by the large bias during spring (Figure A.2). After implementing the DTDL method, the systematic biases during spring diminish considerably. The AHI AOD retrievals from the DTDL algorithm show higher consistency with the ground AOD measurements in all seasons. Some features in AOD biases are shared by the original DT and DTDL algorithms, such as the overestimation in AOD retrievals at 00:00 UTC. Such systematic biases need to be considered when using AHI AOD retrievals in this region.

3.3.2. Validation of AOD retrievals against ground measurements

Figure 3.6a-b show comparisons between ground-measured AOD and AOD derived from the original DT algorithm for medium-quality and high-quality data. Figure 3.6c-d show comparisons between ground-measured AOD and AOD derived from the DTDL algorithm. Note that 70% of ground AOD measurements used in model training are excluded from the analysis. The AODs retrieved from the DTDL algorithm are generally improved for both medium-quality and high-quality categories. With the new DTDL algorithm, DNN-based SRR, the uncertainties (represented by the MAE) of the AHI retrievals are reduced by about 30%. Such a significant reduction in noise considerably improves the quality of the AOD product.

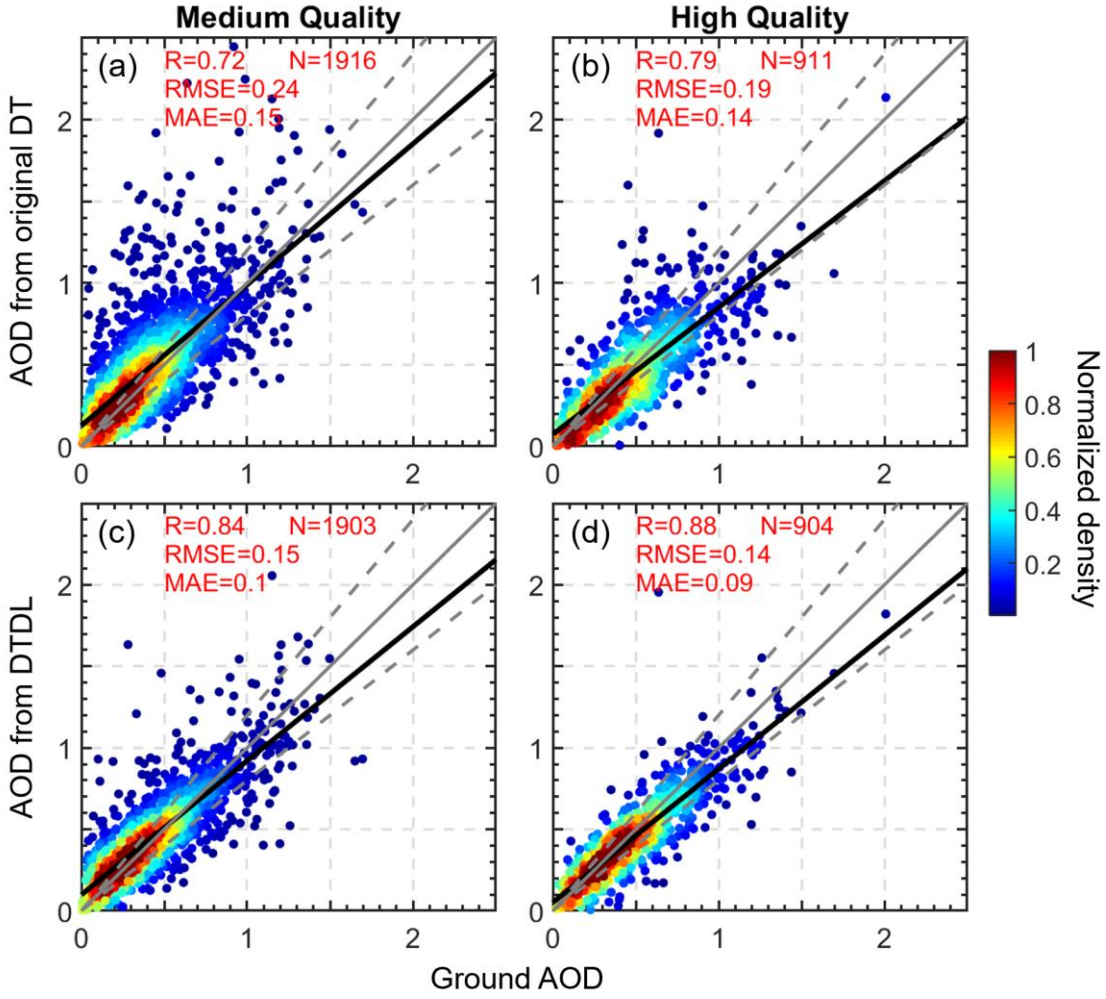


Figure 3.6: *Comparisons between AERONET AOD and AOD derived from the original DT (dark-target) method for (a) medium-quality and (b) high-quality. Comparisons between AERONET AOD and AOD derived from the DTDL method for (c) medium-quality and (d) high-quality. The correlation coefficients (R), RMSE, sampling numbers (N), and MAE are given in each panel.*

The DTDL method is able to produce better AOD retrievals over a region even though the training takes place in another region. Additional independent tests were performed to demonstrate this. The study region was divided into four regions: R1 and R2, defined by latitude, and L1 and L2, defined by longitude (Table 3.2). We use the dataset from R1 to train the model and validate the results over R2. Such a test is repeated for each region. Figure 3.7a-d shows comparisons between ground-measured AOD and DT-derived AOD for the four regions. Figure 3.7e-h shows comparisons between ground-measured AOD and DTDL-derived AOD for R1, R2, L1, and L2, while the training datasets are obtained over regions R2, R1, L2, and L1, respectively. This process ensures the independence of training and validation datasets. Medium-quality and high-quality retrievals are jointly used here. For all four regions, the DTDL-derived AOD shows notable improvement in all three metrics used to measure quality (i.e., correlation coefficient, RMSE, and MAE). The uncertainties of AOD retrievals by DTDL are reduced by 15–25%. The same procedure is performed for high-quality cases, producing similar results (Figure A.3). Thus, the DTDL algorithm not only performs better in regions with training data but also consistently performs better than the traditional DT algorithm over untrained regions.

Table 3.2. *Classification of different regions for independent tests.*

Region code	Classification	Number of sites
R1	Latitude > 23	9
R2	Latitude < 23	9
L1	Longitude > 110	12
L2	Longitude < 110	6

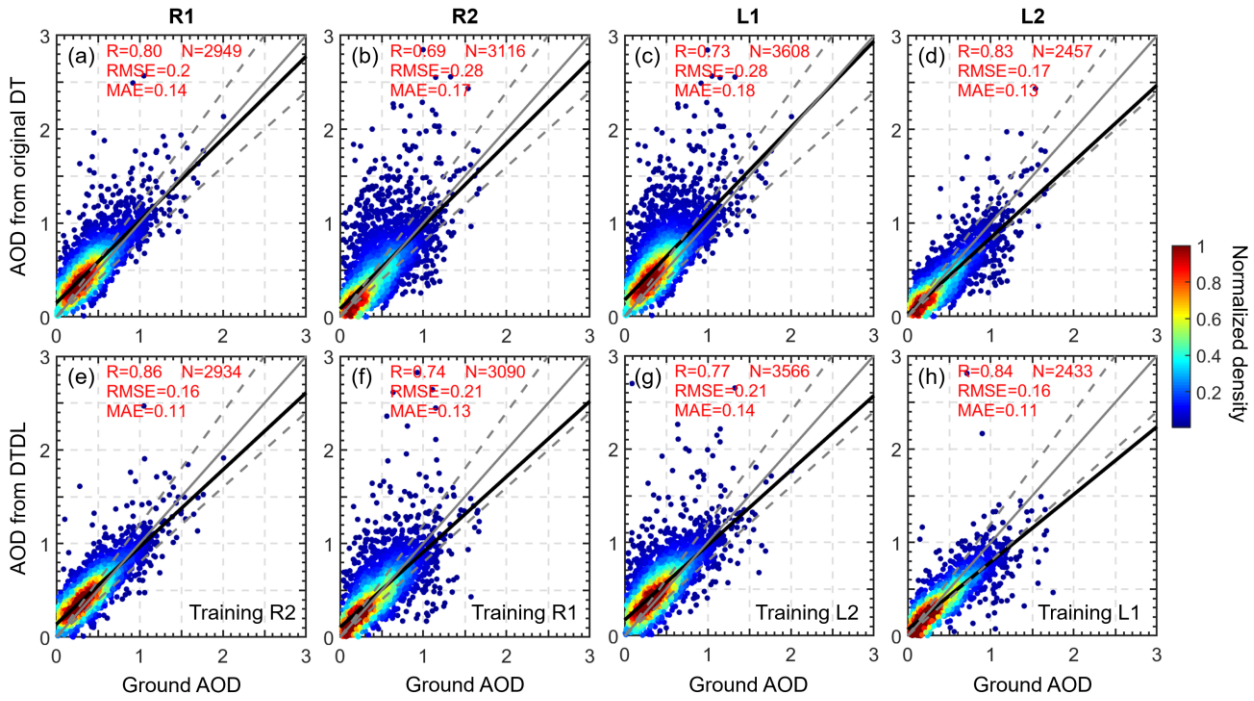


Figure 3.7: *Comparisons between AERONET AOD and AOD derived from the original DT method over regions of (a) R1, (b) R2, (c) L1, (d) L2. Comparisons between AERONET AOD and AOD derived from DTDL over regions of (e) R1, (f) R2, (g) L1, (h) L2, while the training datasets are obtained over the regions of (e) R2, (f) R1, (g) L2, (h) L1, respectively. Medium-quality and high-quality retrievals are jointly used here. This process ensures the independence of the validation datasets. The correlation coefficients, RMSE, sampling numbers, and MAE are given in each panel.*

Regarding the original DT method, the empirical relationship leads to a large bias in estimated SR and can further contribute to the notable uncertainties in AOD retrievals. By implementing a DNN model, multiple contributing factors for the SRR, such as meteorological data and seasonality, can be taken into consideration. The DNN model appears to be a good way to characterize the complex and nonlinear SRR between visible and SWIR channels. The accuracy of AOD retrievals considerably benefits from the improved SRR derived from the DNN. Many factors contribute to the biases in AOD retrievals, such as surface reflectance, cloud contamination, aerosol model, and aerosol vertical distribution, among others. In this study, we have focused on

refining the surface reflectance assumption, but we also bear in mind that other factors can contribute to the error in AOD retrievals.

3.3.3 Uncertainty related to various factors

Since the quality of satellite AOD retrievals depends on the underlying surface and viewing geometry, we calculate the mean absolute differences between the AOD retrievals from ground measurements and AHI under various conditions. Figure 3.8 shows the average absolute bias of AOD for different values of $SR_{0.47}$, $SR_{0.64}$, $NDVI_{SWIR}$, and scattering angle for medium-quality and high-quality retrievals. The mean absolute biases for the original DT and DTDL algorithms are 0.15 and 0.1, respectively. However, the absolute biases in DT-derived AOD can be very large for scenes with a high albedo and low NDVI. This is because the biophysical relationship is valid for dark and dense vegetation but less valid for arid and bright surfaces characterized by low NDVI and high albedo. Note that the frequency of occurrence of high albedo in the region studied is low. Thus, even relatively few outliers can considerably affect the average calculated from AOD over regions with a high surface albedo.

Compared with the original DT algorithm, the absolute bias of the DTDL algorithm is systematically reduced under various conditions. In Figure 3.8, the shaded parts represent standard deviations that are generally larger for the original DT algorithm than for the DTDL algorithm. The traditional DT method cannot deal well with relatively bright surfaces or less vegetated surfaces. Although the DTDL algorithm is also significantly affected by these factors, the revised method can considerably mitigate this issue. The pattern of change in biases with these variables is similar to that shown in Figure 3.8 for the combined medium- and high-quality retrievals, but the magnitude of the biases is smaller. The DTDL algorithm also produces smaller absolute biases in AOD retrievals at all 18 sites (Figure A.4).

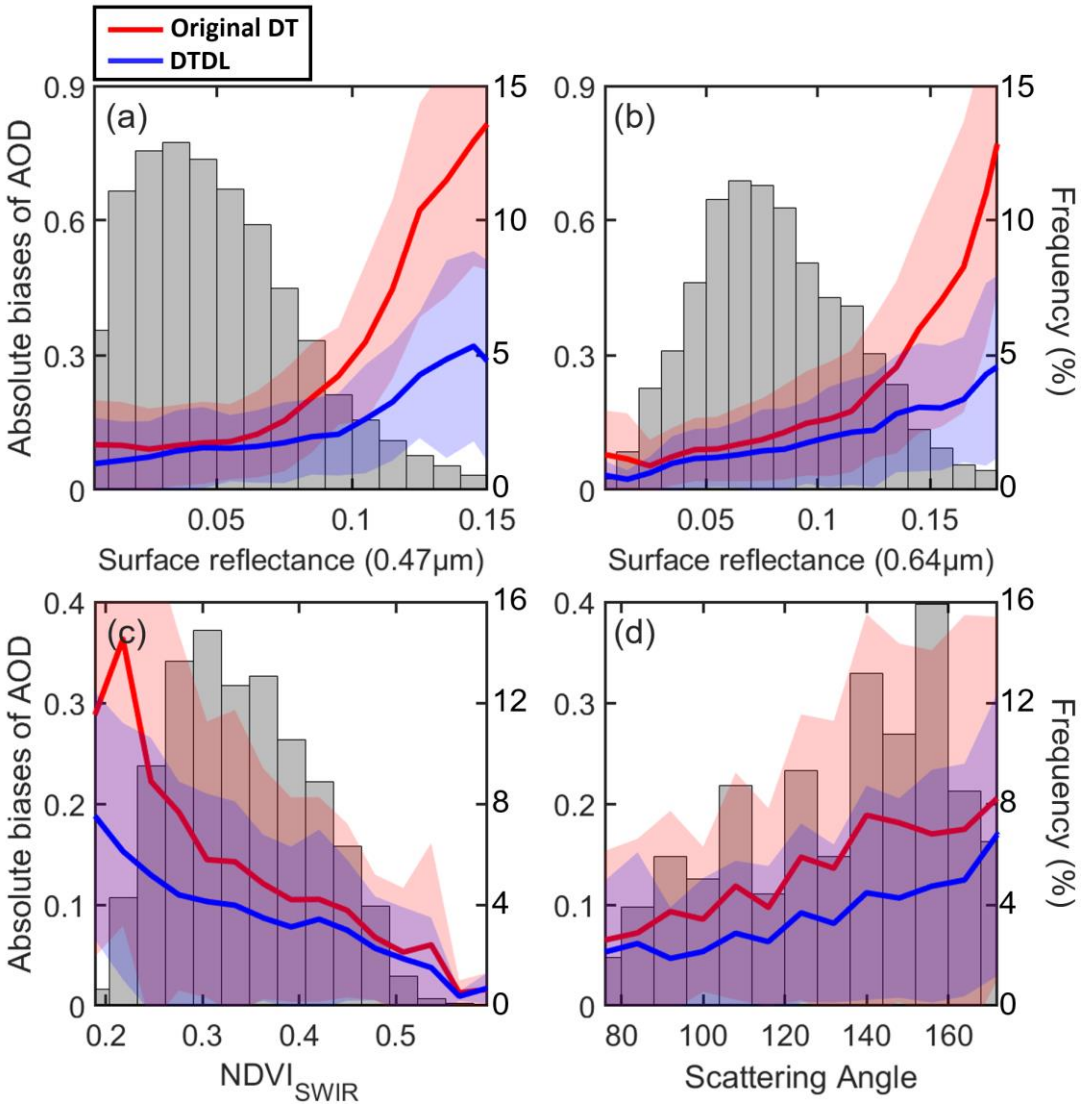


Figure 3.8: *Absolute biases between AOD derived from ground AOD and AHI, for different (a) $SR_{0.47}$, (b) $SR_{0.64}$, (c) $NDVI_{SWIR}$, and (d) Scattering Angle. The original DT (red lines) and DTDL (blue lines) are used respectively. The shaded areas indicate the standard deviations. The grey bars represent the occurrence frequency for these parameters.*

The biases in AOD retrievals are considerably reduced after implementing the DNN-derived SRR in the DTDL algorithm. However, since the visualization of a deep learning model is generally insufficient, interpretability and casualty have been widely recognized as major weaknesses of the DNN (Runge et al., 2015; Zhang et al., 2018), which may be addressed in the future (Montavon et al., 2017; Reichstein et al., 2019). Currently, we only compare the original

DT and DTDL algorithms in this study since there is no public Himawari-8 AOD product from the other two major methods (DB and MAIAC).

Chapter 4 Lidar-based PBLH detection methods

Following the satellite retrievals of AOD, we turn to the lidar-based PBLH detection methods. This chapter explores different techniques for retrieving PBLH from both space-borne lidar (CALIOP) and ground-based lidar. For CALIOP, two techniques are examined: MSD (maximum standard deviation) and WCT (wavelet covariance transform), and their hybrid, which shows better performance (Section 4.1). For ground-based lidar, we developed a new method to retrieve PBL variability under different thermodynamic stabilities (Section 4.2). PBLH retrievals derived from the new method are compared with the commonly used gradient and wavelet methods, and are explicitly compared against radiosonde results under various environmental conditions. We also developed a method to diagnose the surface-cloud coupling, which is used to retrieve the PBLH under cloudy conditions (Section 4.3).

4.1 PBLH derived from space-borne lidar

This section attempts to explore the CALIPSO-derived PBLH using both MSD and WCT techniques. Several refinements have been made to the WCT method to better account for cloud screening, surface detection, noisy signals, etc. Based on Su et al. (2017a), this section presents a comprehensive assessment of the results using 7 years of PBLHs derived from CALIPSO, MPL, and radiosonde.

4.1.1 MSD and WCT methods for CALIPSO

In this study, we utilized the MSD method developed by McGrath-Spangler and Denning (2012), and adopt the threshold method initially proposed by Platt et al. (1994) to identify clouds using the threshold value indicated by Okamoto et al. (2007). Although under some circumstances the top height of boundary layer clouds can be regarded as the PBLH, such estimation requires the correct identification of cloud type which cannot be accurately determined by the lidar. Therefore, the cloudy conditions have been subsequently excluded from section 4.1. We also made several additional modifications and refinements. First of all, the restricted altitude has been modified to the range of 0.3 to 2.5 km according to the characteristics of PBLH over Hong Kong. Moreover, because CALIPSO views from above, when penetrating through an aerosol layer, the signals often become attenuated when it reaches the bottom of the layer. This creates a decline in TAB below

the top of the aerosol layer which may also result in a local maximum of TAB standard deviation. However, this local maximum is not created by the natural jump of aerosol concentration from the PBL to the free atmosphere, but rather by the viewing geometry of the instrument. To address this issue, we define a function:

$$\beta(\mathrm{i}) = \max\{f(\mathrm{i}+2), f(\mathrm{i}+1)\} - \min\{f(\mathrm{i}), f(\mathrm{i}-1)\} \qquad (4.1)$$

where $f(\mathrm{i}+2),\ f(\mathrm{i}+1),\ f(\mathrm{i}),\ f(\mathrm{i}-1)$ are four adjacent altitude bins of 532 nm TAB, with the altitude decreasing with increasing bin number i. We add a constraint of $\beta > 0$ for PBLH identification, which can eliminate the local maximum of standard deviation caused by signal attenuation at low altitudes, and ensures that the location of the PBLH lies at the top of the aerosol layer.

To better illustrate this treatment, Figure 4.1a shows an example of PBLH results derived from CALIPSO using the MSD technique. The TAB from CALIPSO on 24 October 2007 over Hong Kong is plotted with a horizontal smoothing of 3 km and a vertical sliding window of 90 m to improve the SNR. The solid black line indicates the derived PBLH using the MSD method with constraint $\beta > 0$ and the dashed line shows that without constraint. A horizontal smoothing of 7 km is applied to the retrieved PBLHs in order to minimize the influence of outliers and to increase its spatial consistency. Figure 4.1b presents the vertical profiles of TAB (black) and the vertical standard deviation (blue) of TAB close to the MPL site (YL) averaged over 20 km. We can see that there are two local maxima in the standard deviation profile. The lower one is caused by the signal attenuation by the aerosols above, which is effectively eliminated by the constraint of $\beta > 0$. Hereafter, the method using the MSD technique with constraint is simply referred to as "MSD".

In addition, we also introduce the WCT technique to derive PBLH from the TAB profiles of CALIPSO. We use a Haar wavelet whose function h is defined as (Gamage and Hagelberg, 1993):

$$\mathrm{h}\left(\frac{z-b}{a}\right) = \begin{cases} +1: & b - \frac{a}{2} \le z \le b \\ -1: & b < z \le b + \frac{a}{2} \\ 0: & elsewhere \end{cases} \qquad (4.2)$$

where z is altitude, b is called the "translation" of the function where the function is centered, and a is called the "dilation" of the function. Then, the covariance transform of the Haar function $W_f(a, b)$ is defined as:

$$W_f(a,b) = \frac{1}{a}\int_{z_b}^{z_t} f(z)\mathrm{h}\left(\frac{z-b}{a}\right)dz \qquad (4.3)$$

where $f(z)$ is the backscatter signal of the lidar,z_b is the lower limit and z_t is the higher limit of the profile. The value b at which $W_f(a,b)$ reaches the local maximum with a coherent scale of a is usually considered as the PBLH. The selection of dilation a is the key in this process. Brook (2003) proposed that the best selection of dilation should equal the transition zone range of the ideal profile. In previous studies, the selection of dilation a generally ranges from 100 m to 1000 m for ground-based lidar (e.g., Davis et al., 2000; Brooks, 2003; Compton et al., 2013). For choosing an appropriate dilation a, we perform a series of sensitivity tests. Figure 4.1c presents the derived PBLHs using WCT techniques with different values of the dilation roughly within the range of those used in previous studies. Figure 4.1d shows the corresponding results for a spatially averaged profile which is the same as Figure 4.1b. The selected PBLH retrieval range, surface noise check, and cloud screening are the same as the MSD method. The specific $W_f(a,b)$ function corresponding to different values of the dilation a can be found in Figure B.1. The selected PBLH retrieval range, surface noise check, and cloud screening are the same as the MSD method. It is seen from Figure 4.1d that when the dilation a is chosen between 600 m to 900 m, the PBLH results show no substantial difference and all correctly identify the top of the aerosol layer. Therefore, we consider this range appropriate. Note that this range is larger than the typical value used for MPL (Brook, 2003) since larger dilation is required to smooth out the noisier CALIPSO signals. We choose the dilation a as 690 m, and make sure that there do not exist obvious unrealistic conditions through visual inspection. The method using WCT will be simply referred to as "WCT" in the following text.

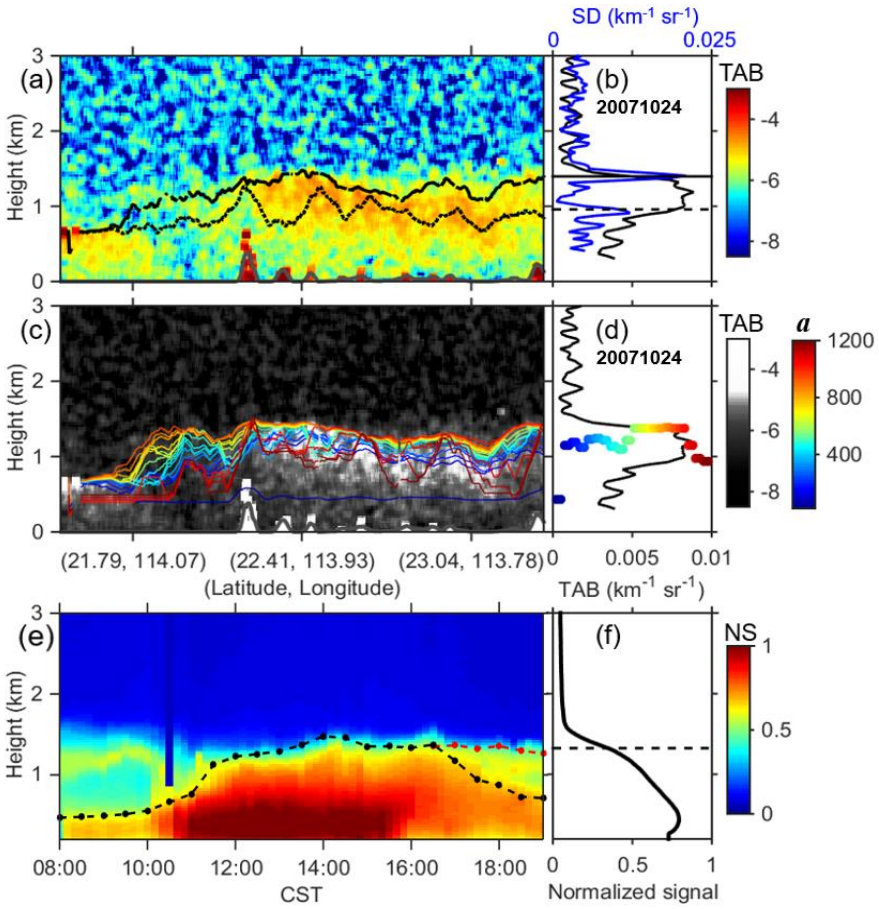


Figure 4.1: *(a) Total Attenuated Backscatter (TAB) plot (log scale) from CALIPSO on 24 October 2007 over Hong Kong. The black lines indicate the derived PBLH using MSD with constraint* $\beta >$ 0 *(solid) and without constrain (dash) respectively. The gray line represents the surface elevation. (b) Vertical profile of TAB (black curve) and the vertical standard deviation (blue curve) of TAB closest to MPL site. (c) and (d) are similar to (a) and (b), except that the lines and dots indicate the derived PBLH using WCT. The curves in different colors represent various PBLHs based on the assumption of various values of the dilation used in WCT. The midpoints of (a) and (c) are the positions closest to MPL. (e) Time evolution of the NS plot from MPL on 24 October 2007. The black line identifies the PBLH derived from MPL and the red line represents the residual layer. (f) The NS profile of MPL average from 1300 to 1400 LT.*

4.1.2 Comparison between PBLHs derived from CALIPSO and ground-based data

In this section, we compare and evaluate in detail CALIPSO-derived PBLH against radiosonde and MPL results, as well as between the MSD and WCT methods. Figure 4.2 shows

the scatter plots of CALIPSO-derived PBLH by the two approaches against those obtained from the MPL at YL(b), STK(c), PH(d), and all these three stations(a). The PBLHs derived from MPL and radiosonde are again averaged between 1300 to 1400 LT. Figure 4.2 indicates that CALIPSO derived PBLH using MSD and WCT both agree well with MPL. The correlation coefficients are mostly above 0.6, and reach the highest value (~0.7) at YL. The RMSE between WCT and MPL is lower than that for MSD, which well indicates that the WCT technique can be quite suitable for retrieving PBLH from CALIPSO data and even outperforms MSD. Compared with MPL versus radiosonde, CALIPSO PBLHs bear larger differences from radiosonde results (Figure 4.2e), with both lower correlation and higher RMSE. This phenomenon is in part due to the different standards used to estimate PBLH for these two types of measurements: both CALIPSO and MPL retrievals are based on the aerosol concentration, whereas radiosonde retrieval is based on thermodynamic conditions.

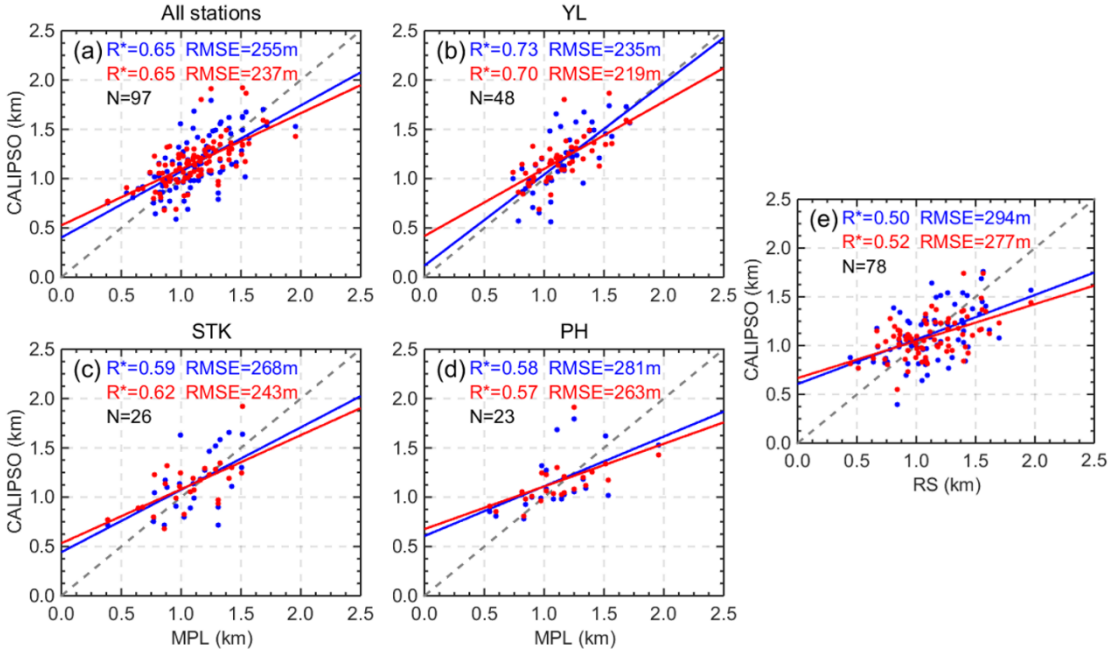


Figure 4.2: *The comparison of noontime PBLHs derived from CALIPSO and MPL at all the three stations (a), YL (b), STK (c), and PH (d). The comparison of noontime PBLHs derived from CALIPSO and radiosonde (e). The blue lines and dots are derived from MSD, while the red lines and dots are derived from WCT.*

The climatology of PBLH, especially its seasonal variability, is also very important in climate-related studies. Figure 4.3 further shows the monthly averaged diurnal cycle of PBLH derived from MPL and radiosonde, with CALIPSO results overlaid at approximately 1330 LT. The MPL-derived and radiosonde-derived PBLHs show similar characteristics of the diurnal cycle, i.e., it grows rapidly from 0800 LT, reaches a maximum around 1400 LT, maintains a relatively high value until 1600 LT, and then starts to drop. Overall, the PBLH over Hong Kong exhibits a clear seasonal cycle, with a maximum in autumn and a minimum in winter, consistent with a previous study by Yang et al. (2013). Radiosonde and MPL derived PBLHs agree well in terms of seasonality, only with a slightly larger range of variability for radiosonde. CALIPSO also captures general PBLH seasonal variability for both approaches. We note that the average value of CALIPSO-derived PBLH agrees well with MPL and radiosonde in the fall and winter, whereas larger differences are found in spring and summer (Figure B.2). The daily maximum PBLH typically increases from January, achieving a maximum of around 1.2 km in October, and then rapidly declining to 0.93 km in December. Figure 4.3 also confirms that the largest disagreements in monthly averaged results between CALIPSO and MPL/RS are in the spring and summer. In particular, an overestimate as large as 0.2 km is found for March and April while an underestimate of similar magnitude is observed during June and July.

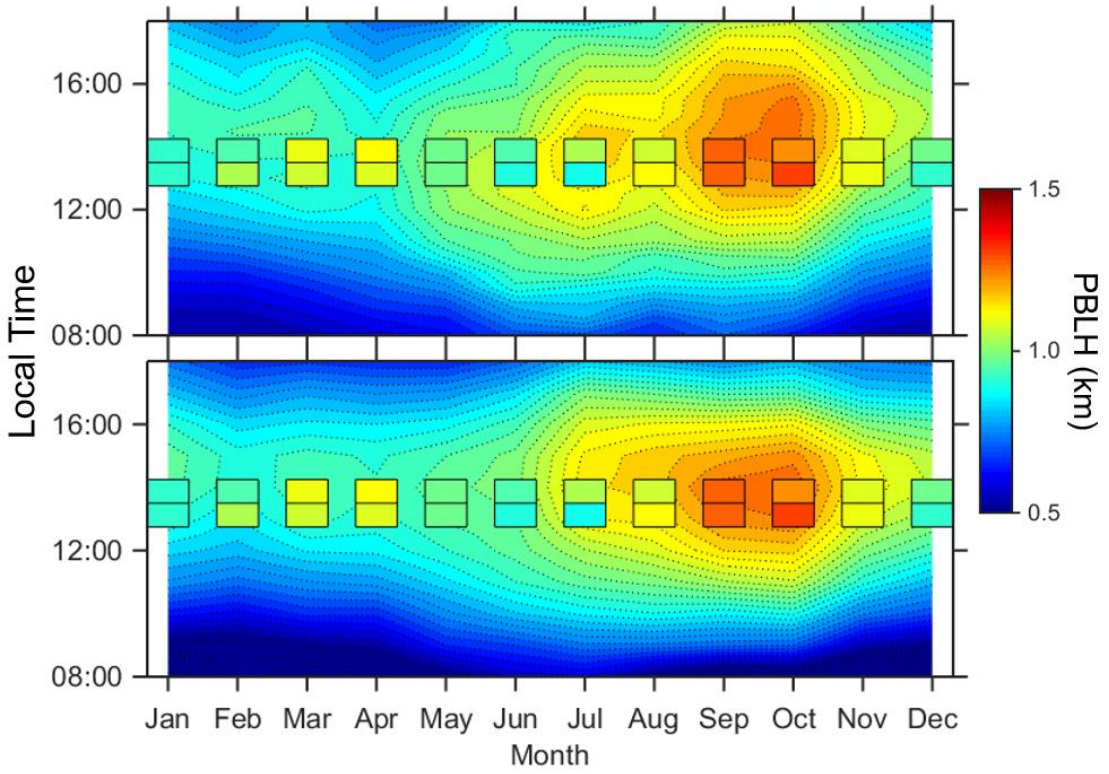


Figure 4.3: *(a) Monthly mean diurnal variation of PBLH derived from MPL. (b) Monthly mean diurnal variation of PBLH derived from the radiosonde. The squares represent the monthly mean PBLHs derived from two approaches of CALIPSO at approximately 1330 LT. The lower rectangles indicate MSD results, while the upper ones indicate WCT results.*

4.1.3 A hybrid approach of WCT and MSD

Comparing these two methods, the implementation of MSD is comparatively simple, making it a widely used method in retrieving PBLHs from CALIPSO data (e.g., Jordan et al., 2010; McGrath-Spangler and Denning., 2012; Liu et al., 2015). Nonetheless, the local maximum in the standard deviation of the backscatter profile is not always located at the PBL top, because of signal attenuation and random noises of backscatter. On the other hand, WCT is less subject to random noise and usually achieves better results under low SNR conditions, yet it is a scale-dependent approach and involves the choice of the wavelet function and a few other parameters. However, our study shows that WCT can be effective for retrieving PBLHs from CALIPSO with an appropriate selection of dilation *a*.

Finally, by eliminating cases when the difference between MSD and WCT is above 100 m, the agreement between CALIPSO and MPL/RS results is much improved (Figure 4.4). Thus, jointly using two methods can produce more reliable results. In general, we can apply the MSD method to retrieve the PBLH, and the retrieval from WCT serves as a constraint. We can eliminate cases when the difference between the MSD and WCT retrievals exceeds 100 m, to increase the reliability of the MSD retrievals. Nonetheless, this is still partly related to aerosol loading since when the two methods are more consistent, aerosol loading is usually high and thus the agreement between CALIPSO and MPL/RS also becomes better.

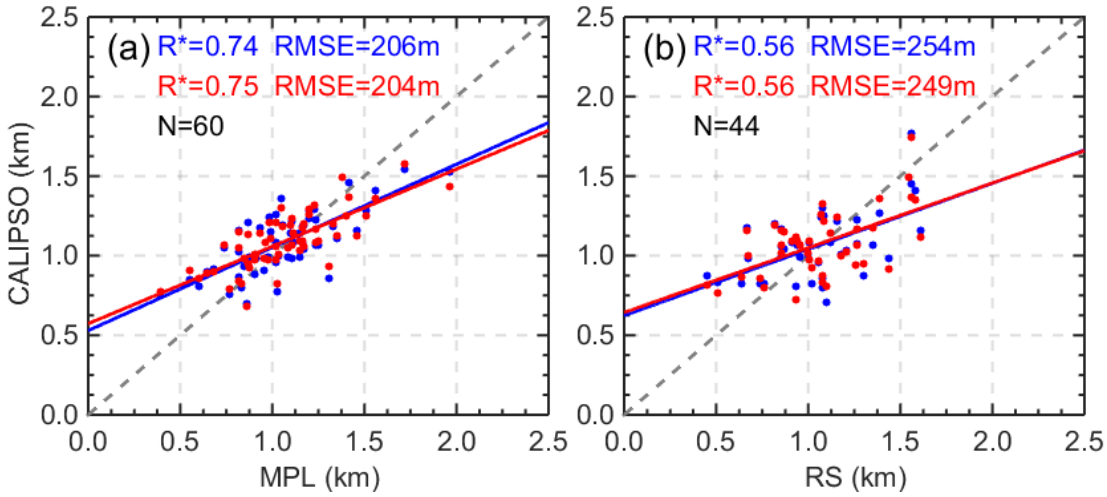


***Figure 4.4**: (a) The updated correlation of PBLH derived from CALIOP and MPL after eliminating cases when the difference between MSD and WCT is above 100 m. (b) The same as (a) but for the comparison between CALIPSO and radiosonde.*

4.2 PBLH derived from ground-based lidar

In addition to the development of PBLH retrievals for space-borne lidar, we attempt to develop advanced PBLH detection methods for ground-based lidar. To address the limitations stated in Section 1.3, the current study presents a new method to retrieve PBL variability from MPL under various thermodynamic stability conditions. This section first describes the method in terms of its physical principles, implementation, and determination of the thresholds required by the DTDS method. A comprehensive evaluation is presented in Section 4.2.3.

4.2.1 A brief overview of existing algorithms

Numerous methods have been proposed to retrieve the PBLH from backscattering lidar. These methods rely on the principle that a temperature inversion often exists at the top of the PBL, trapping moisture and aerosols (Seibert et al. 2000), which causes a sharp decrease in the aerosol backscattering signal from the PBL upper boundary to the free atmosphere. Traditionally, the methods fall into two categories: (i) gradient-based algorithms that track gradients in the vertical distribution of aerosols (i.e., gradients of the backscattering profiles (Flamant et al. 1997; He et al. 2006; Sicard et al. 2006; Summa et al. 2013; Yang et al. 2017), and (ii) wavelet transform algorithms that track the covariance transform of the Haar function (Baars et al. 2008; Brooks 2003; Cohn and Angevine 2000; Davis et al. 2000). The merits and limitations of these approaches were evaluated by (Sawyer and Li 2013), who proposed a joint approach taking advantage of the two that can be applied to any type of measurements, including radiosonde and lidar, and compared their performance.

The gradient method is adapted from Amiridis et al. (2007) and Flamant et al. (1997), and we set the first derivative of lidar backscattering signals as the gradient profile. The wavelet method is another commonly used, more sophisticated method. As a method of detecting step changes, the wavelet method was introduced by Gamage and Hagelberg (1993). The techniques of wavelet method also are detailed in Section 4.1.1.

Despite the dominance of the CBL in the afternoon, SBL and NBL can also occur under some meteorological conditions. Figure 4.5 shows a comparison of afternoon PBLH derived from MPL and radiosonde under SBL, NBL, and CBL conditions. Hereafter, “afternoon” refers to the period from 1200 to 1830 LT. Again, MPL-PBLH shows much better agreement with radiosonde-PBLH for CBL and NBL conditions than for SBL, especially where radiosonde-PBLH < 1.0 km. Therefore, MPL has a major problem retrieving PBLH for weak convective cases. Note also that for all three PBL types, RMSE values are around 0.6 km.

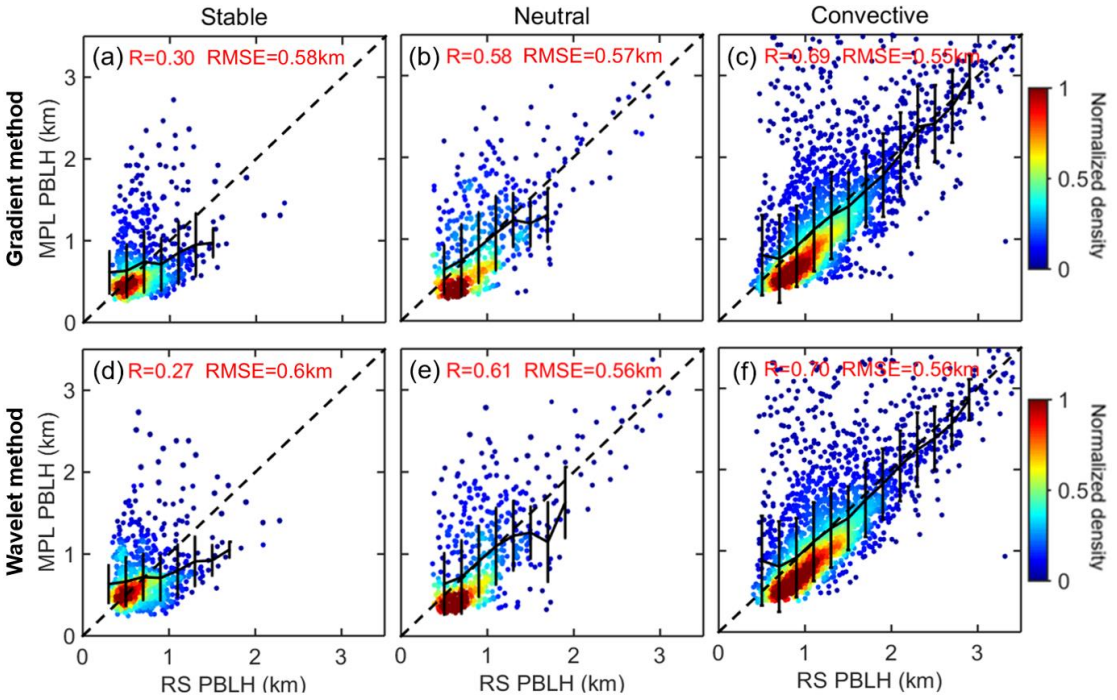


Figure 4.5: *Comparisons between MPL PBLH (derived by the gradient method) and radiosonde PBLH for (a) stable (SBL), (b) neutral (NBL), and (c) convective (CBL) boundary layers during the afternoon (1200-1900 LT). Comparisons between MPL PBLH (derived by the wavelet method) and radiosonde PBLH for (d) SBL, (e) NBL, and (f) CBL.*

4.2.2 Description of our new method (DTDS)

In order to overcome problems with SBL conditions, we develop a new approach for retrieving MPL-PBLH that performs better and more uniformly under all the thermodynamic stability conditions. The physical basis of our DTDS method is to account for both the vertical coherence and temporal continuity of the PBLH. Given that the aerosol within the boundary layer is usually well mixed, a distinct step is sought in the backscattering signal by aerosol as the top of a PBL layer, which follows the gradient and wavelet transform methods. The development of the PBL can be divided into three phases: growing, decaying, and "other" (i.e., maintenance), and apply different selection schemes to these different phases. A schematic diagram of the DTDS is presented in Figure 4.6, and the procedures are detailed below.

First, the MPL backscattering profiles are averaged over 10-minute intervals. Then, we identify the local maximum positions (LMPs) (range: 0.25-4km) in the covariance transform function

collocated with a signal gradient larger than a certain threshold. The aerosol vertical structure can be represented by multiple LMPs. Each LMP is the potential position of PBLH.

$H(i)$ represents the retrieved PBLH for time i. For the starting point (0700 LT), the PBLH ($H(0)$) is given by the morning radiosonde profile (or the lowest LMP if radiosonde is not available), after which we select the closest LMP at 0710 to $H(0)$ as $H(1)$. If we find suitable $H(1)$, we identify the closest LMP at the next timestep as $H(2)$. Generally, for time i+1, we select the closest LMP to $H(i)$ as $H(i+1)$. The maximum variation rate (Max) is set to 0.2km, so the difference between $H(i+1)$ and $H(i)$ needs to be smaller than Max. Otherwise, $H(i+1)$ is set to equal to $H(i)$, and is called a "quasi retrieval". If we get two consecutive quasi retrievals, $H(i)$ and $H(i+1)$, we set $H(i+1)$ to the lowest LMP in time i+1, and then, determine $H(i+2)$ using above-mentioned algorithms.

As shown in Figure 4.6b, the selection scheme is different for the growing (0830-1400) and decaying (1630-1900) PBL phases, defined by these pre-selected local time periods. Due to surface heating by solar radiance, the PBLH tends to rise quickly during the growing period, and generally reaches its daily maximum monotonically. For time i+1, if we find two LMPs adjacent to the previous $H(i)$, we preferentially select the one higher than $H(i)$ as $H(i+1)$ during this period. However, we preferentially select the one lower than $H(i)$ as $H(i+1)$ during the decaying period. Details of the selection scheme for different periods are shown in Figure 4.6b. Generally, during the growing (decaying) phase, the nearest LMP above (below) *H(i)* is selected for *H(i+1)*. However, if the two nearest LMPs during the growing (decaying) phase are above and below *H(i)*, the upper (lower) LMP must be twice as far from *H(i)* as the lower (upper) one for the lower (upper) one to be selected. If we cannot find two LMPs adjacent to *H(i)*, a potential PBLH position is selected as the nearest LMP.

A restart mechanism also is introduced to this algorithm, controlled by the parameter "Re," which remains at 0 as long as the selection scheme in Figure 4.6b is successful. For two contiguous times (i and i+1), if we cannot find the suitable $H(i)$ and $H(i+1)$, Re becomes larger than 1, and the algorithm "restarts." Specifically, the lowest LMP in the time i+2 is selected as $H(i+2)$, and the iterative process is restarted from there.

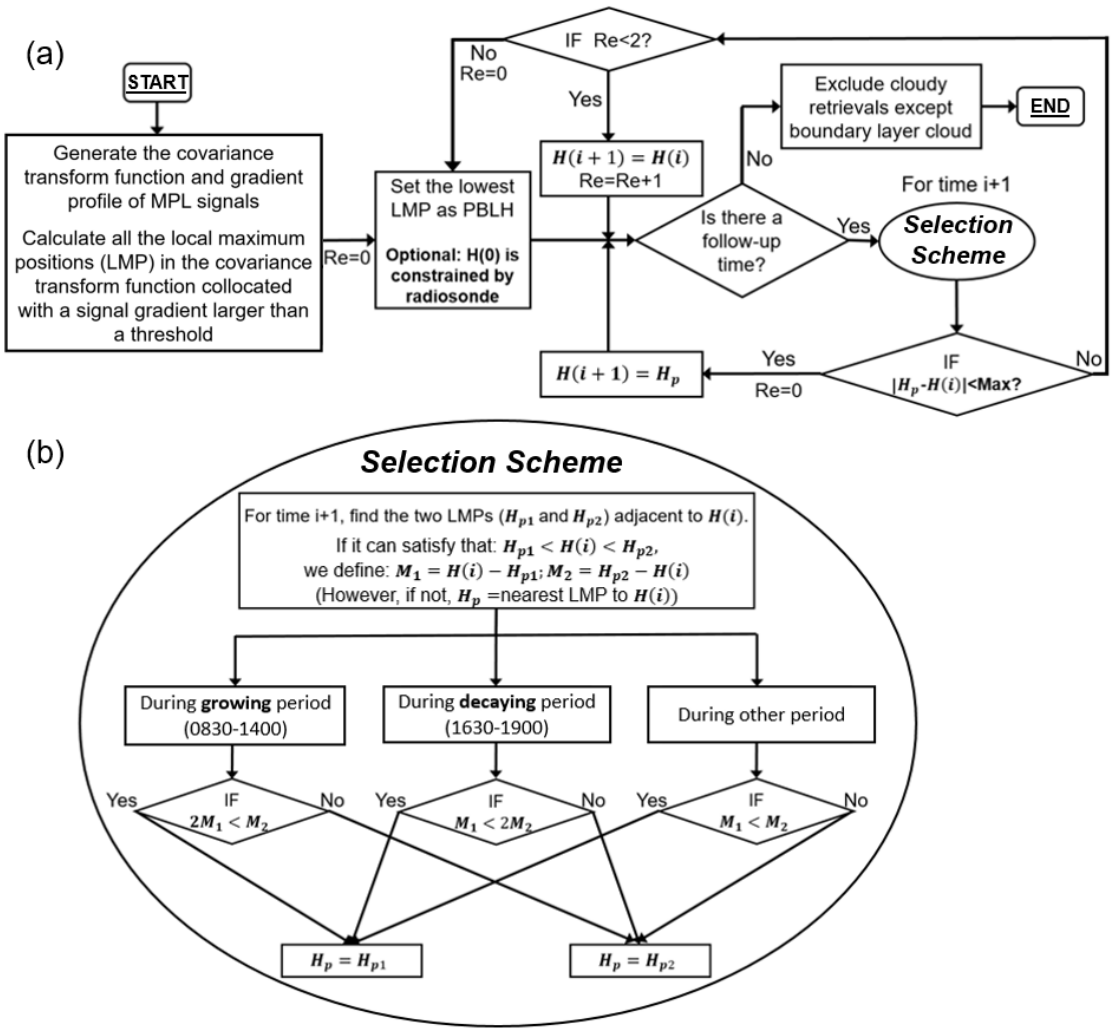


Figure 4.6: *(a) The overview schematic diagram of retrieving the PBLH by the DTDS. In this diagram, $H(i)$ indicates the retrieved PBLH for time i, and H_p indicates the potential PBLH position. The PBLH is retrieved at 10-minute intervals. Max indicate the maximum variation of PBLH allowed over 10-minutes, and is set as 0.2 km in this study. (b) Schematic diagram of the selection scheme. The selection scheme is a process in the retrieval algorithm that describes how an appropriate local maximum position (LMP) as H_p is selected.*

4.2.3 Evaluation of the DTDS

By using the DTDS method, we can continuously retrieve the daily PBLH during 0700-1900. Figure 4.7 illustrates three typical cases of backscattering profiles and PBLH retrievals under different stability conditions on 25 Dec. 2014, 13 Sep. 2010, and 1 Jul. 2010 respectively

Relatively complex aerosol stratification is indicated by multiple LMPs for each profile, especially during the morning periods.

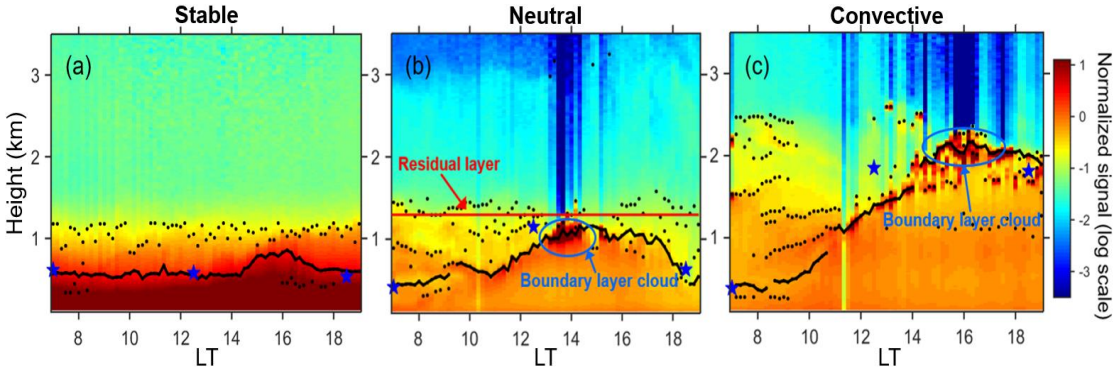


Figure 4.7: *Daily backscattering profiles for (a) SBL, (b) NBL, and (c) CBL cases. Backscattering is presented as a normalized signal on a log scale, in arbitrary units. Black lines mark the PBLH retrieved from the DTDS. Black dots indicate the LMPs identified in the backscattering profiles. The blue stars indicate the PBLH as derived from the radiosonde.*

For the SBL case, a clear stratification structure of aerosol is shown, with a little diurnal variation. Due to relatively weak convection, the PBL does not develop well during the entire day. However, there are multiple LMPs at different elevations; aerosol step signals are found around 0.9 km, 1.9 km, and 2.5 km. If we retrieve PBLH based on single profiles, the retrievals would jump between different atmospheric layers, and would not provide accurate results. With initial input of radiosonde in the morning, we coherently track the PBLH by the DTDS during the entire day, and the result agrees well with subsequent radiosonde measurements.

For the NBL and CBL cases, the PBLH grows continuously from the initial points constrained by radiosonde measurements, and reaches a maximum around 1500LT. The retrievals derived from the DTDS exclude interference from the complex aerosol structures that occur in the morning. The PBLH generally maintains its highest value during 1500-1900LT for this case. Nonetheless, the PBLH starts to decay after 1600LT for NBL. The selection scheme for the decaying period assures that our method can distinguish the actual PBL from the aerosol residual layer. Moreover, boundary layer clouds are identified for the NBL and CBL cases, and the retrieved PBLH location remains consistent with the values under a cloud-free condition.

In general, DTDS reliably tracks the true PBLH development during daytime for these three typical cases, and matches well the radiosonde measurements. In some periods, multiple gradients are found in the lidar backscatter observations. Despite the complex aerosol structure, the DTDS successfully identifies the diurnal variation of the PBLH in each case, given the initial point constraint provided by the radiosonde data.

In this section, we further compare the performance of the DTDS with the traditional algorithms over the diurnal range of stability conditions. We apply DTDS to the 8-year MPL dataset at the ARM SGP site, and compare the PBLH results with those from radiosonde measurements. We prefer to use the morning radiosonde to constrain the initial position of PBLH. However, the DTDS method still performs without morning radiosonde input. Figure 4.8a-c show comparisons between PBLH derived from the DTDS without radiosonde constraints and radiosonde PBLH, under different thermodynamic stability conditions during the afternoon. Figure 4.8d-f show comparisons between PBLH derived from DTDS with radiosonde constraints and radiosonde PBLH.

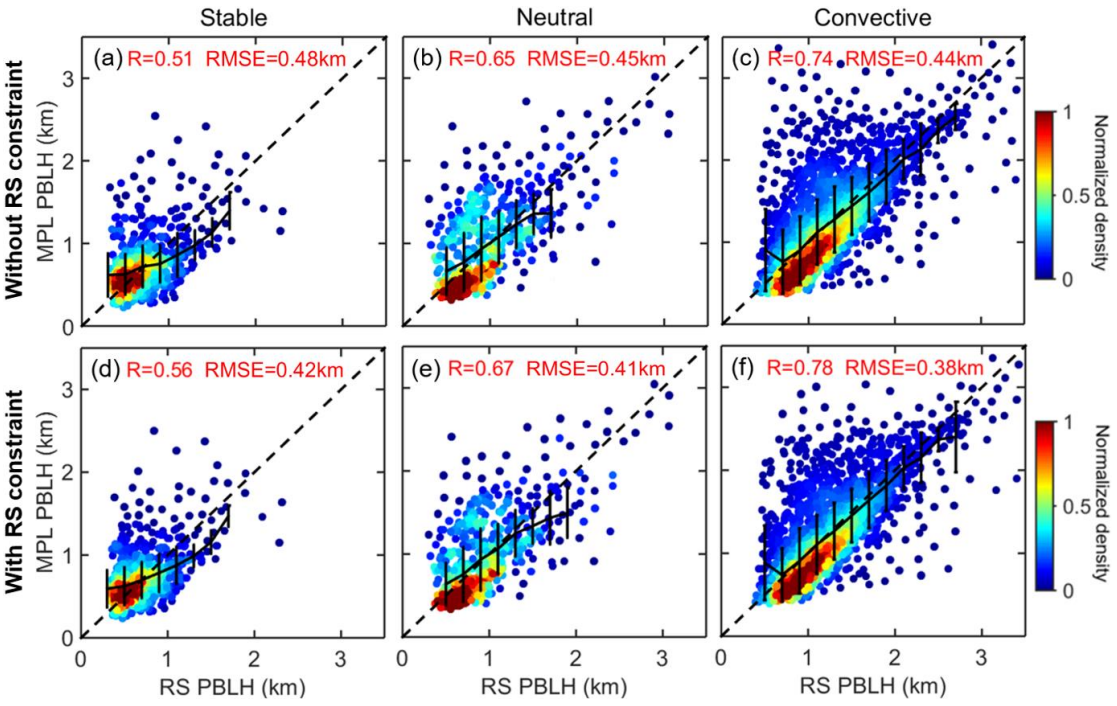


Figure 4.8: *Comparisons between afternoon PBLH derived from MPL and from the radiosonde under different thermodynamic stability conditions. In (a, b, c), PBLH is calculated by the DTDS without the constraint of morning radiosonde data. In (d, e, f), PBLH is calculated by the DTDS with a morning radiosonde constraint.*

Compared to the gradient and wavelet methods (Figure 4.5), the correlation coefficients between PBLH derived from the DTDS and radiosonde are higher for all PBL types, and are improved especially for the SBL cases. Although the correlation is still relatively weak for weakly convective cases, it is substantially improved relative to the gradient and wavelet methods. Moreover, compared to the RMSE for the gradient and wavelet methods, the RMSE values for the DTDS vs. radiosonde are considerably reduced under SBL and CBL conditions. Although the DTDS method is valid without the morning radiosonde input, better performance is achieved after assimilating morning radiosonde input, especially under SBL and CBL conditions. Due to the availability of radiosonde, we analyze the PBLH retrievals derived from the DTDS with the morning radiosonde constraint in the following sections.

In general, the PBLH retrievals derived from the DTDS are improved for all PBL types. Although the correlation coefficients are generally higher only under SBL conditions, the RMSE values are considerably reduced for both SBL and CBL in the study dataset. The diurnal cycles derived from the DTDS also match well with RS results under different thermodynamic stability conditions (Figure B.3).

The PBLH growth rate during transient periods is directly linked to the entrainment process, and is a key factor associated with boundary layer turbulence and eddies, as well as aerosol dilution (Carson 1973; Driedonks 1982; Martin et al. 1988). We show the distribution of PBLH growth rates during 0700-1230 LT under different thermodynamic stability conditions derived from different methods (Figure 4.9). Clearly, strong and positive buoyancy forcing under CBL conditions contributes to the large PBLH growth rate, which is near zero under SBL conditions. However, conventional lidar algorithms (gradient and wavelet methods) cannot capture well such phenomena, and the derived PBLH growth rates are similar with significantly large variations under different thermodynamic stability conditions. The PBLH growth rate can be estimated by DTDS with much better accuracy than the traditional methods. The absolute basis in PBLH growth rate between radiosonde and the wavelet/gradient methods is around 0.06 km/h, and is largely reduced to 0.033 km/h between radiosonde and the DTDS.

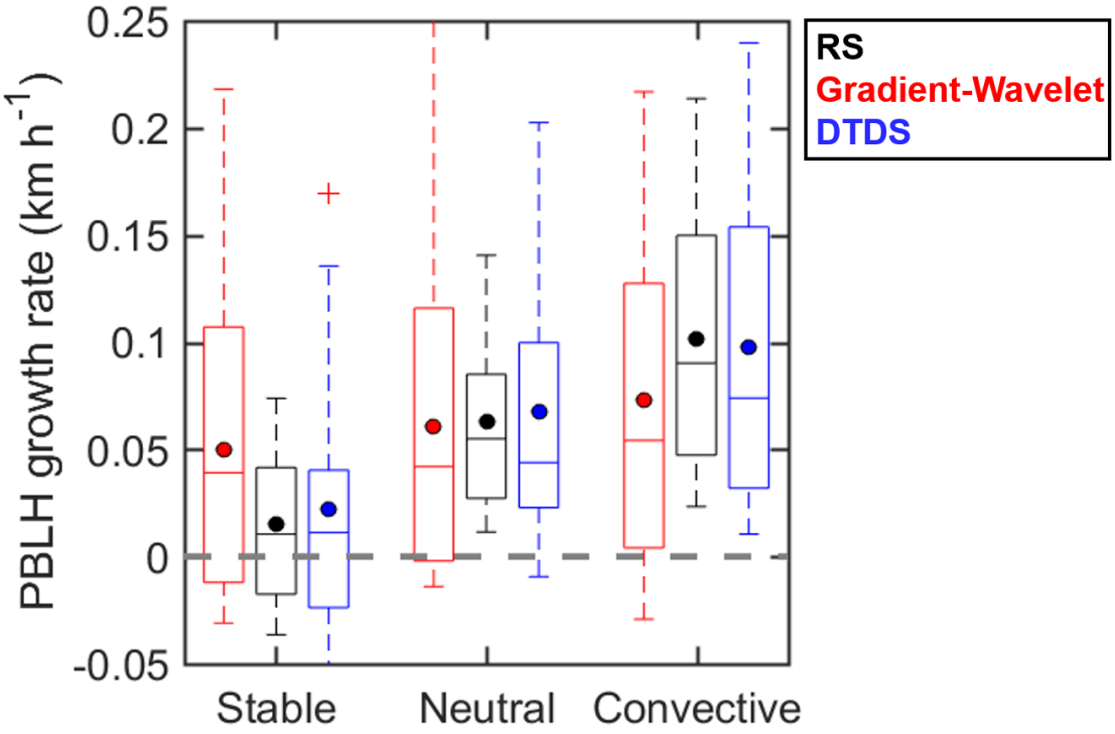


***Figure 4.9**: Seasonal box-and-whisker plots showing 10th, 25th, 50th, 75th, and 90th percentile values of PBLH growth rate during 0700-1230 LT derived from the radiosonde (black), gradient and wavelet methods (red), and the DTDS from MPL (blue) under SBL, NBL, and CBL conditions.*

Chapter 5 Remote sensing of continental clouds coupling with surface

Following the development of lidar-based PBLH detection methods, this chapter further investigates the identification of states of coupling between clouds and surface or PBL. The cloud coupling has been investigated much more extensively for marine stratocumulus clouds than for continental low clouds, partly due to more complex thermodynamic structures over land. A manifestation is a lack of robust remote sensing methods to identify coupled and decoupled clouds

over land. In Section 5.1, we try to generalize the concept of coupling and decoupling to low clouds over land, based on potential temperature profiles. In Section 5.2, we use long-term lidar measurements to develop a method to simultaneously retrieve the PBLH and coupling states under cloudy conditions during the daytime. The retrievals of coupling states are presented in Section 5.3. We found that the diagnosis of cloud coupling regimes can be beneficial for determining PBLH under cloudy conditions.

5.1 Definition of coupled and decoupled clouds based on thermodynamics

The definition of the state of cloud-surface coupling over land is a critical question. For marine stratocumulus, coupled clouds are identified when the liquid water potential temperature varies less than a certain threshold (i.e., 0.5 K) below the cloud base (Jones et al., 2011). We try to extend the concept of coupling and decoupling to clouds over land. The PBL over land is typically buoyancy-driven and controlled by surface fluxes during the daytime. We consider a cloud is in the coupled state when it strongly interacts with the buoyancy fluxes within the PBL.

Figure 5.1 presents the idealized vertical profiles of virtual potential temperature (θ_v) under the clear-sky, coupled cloud, and decoupled cloud. A superadiabatic surface layer exchanges the heat fluxes between the surface and PBL. The outer layer and entrainment zone are turbulently coupled with the surface, and thus, are considered as the coupled regime. Meanwhile, the free atmosphere is considered as the decoupled regime. Theoretically, θ_v is constant in the outer layer, and follows the wet adiabatic lapse rate in the cloud layer. Although the profiles of θ_v in the real atmosphere can largely differ from the idealized profiles, the relative position between the cloud layer and capping inversion of the entrainment zone is clear. For the coupled cases, the cloud base is below the capping inversion of the entrainment zone. For the decoupled cases, the cloud base is above the capping inversion. Based on this feature, we can use the virtual potential temperature profiles to diagnose the coupling state of low clouds.

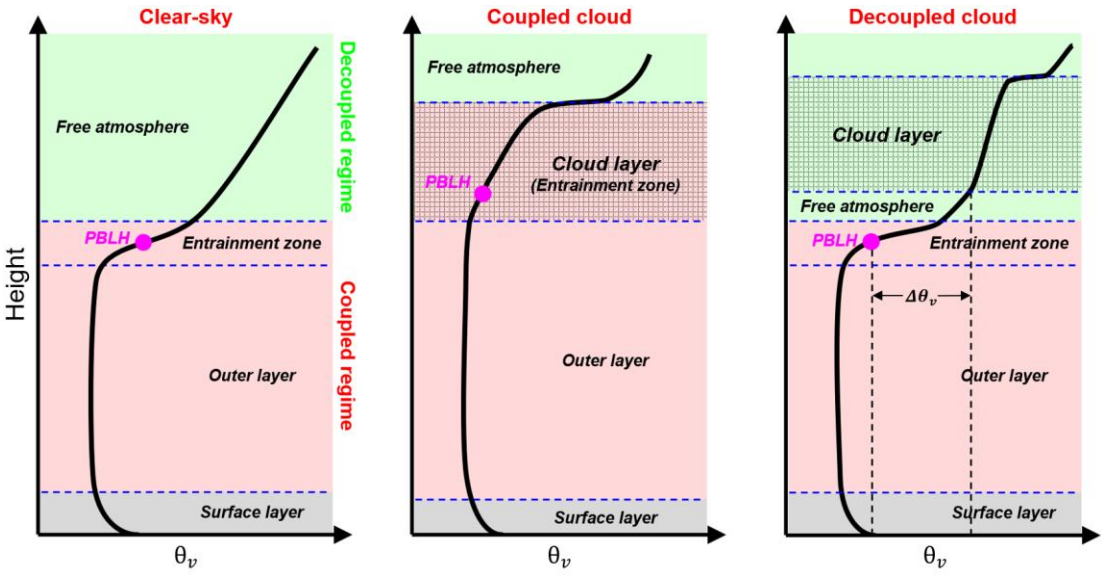


Figure 5.1: *Idealized vertical profiles of virtual potential temperature (θ_v) under the clear-sky, coupled cloud, and decoupled cloud over the land. The surface layer, outer layer entrainment zone, and the free atmosphere are divided by the blue dash lines. The cloudy layer is marked as the shaded area, and PBLH is marked as the pink point. Red and green zones indicate the coupled and decoupled regime, respectively. Elements (e.g., turbulence, heat fluxes, cloud) in the coupled regime are directly affected by the PBL processes, while these elements are not directly affected by the PBL processes in the decoupled regime. For the coupled cases, the cloud base is below the capping inversion of the entrainment zone. For the decoupled cases, the cloud base is above the capping inversion.*

We first look at several examples of profiles of θ_v and WVMR from the radiosonde (Figure 5.2). If the CBH is lower than the PBLH, the cloud is affected by turbulence and buoyancy fluxes in the PBL, such as the cases shown in Figure 5.2a. Note that the PBLH is not an absolute boundary limiting turbulence and buoyancy fluxes. Due to the overshooting of rising air parcels, we use a range to screen the condition of coupled clouds. As shown in Figure 5.2b, even when the CBH is slightly above the PBLH, WVMR and θ_v are still relatively consistent between the cloud layer and the PBL and show large step signals at the cloud top.

Figure 5.2c-d shows a clear inversion layer between the cloud base and the PBL top, and the difference in θ_v between the CBH and the PBLH ($\Delta\theta_v$) is relatively large. Such a notable inversion

layer prevents the buoyancy fluxes within the PBL from reaching the cloud base, leading to the decoupling between the cloud and the PBL. Overall, whether there is a clear inversion between the cloud base and the PBL top is the key factor in determining coupling and decoupling. In this aspect, $\Delta\theta_v$ is the key factor. In Figure 5.2, $\Delta\theta_v$ for coupled cases (a-c) is -0.32 K and 0.31 K, respectively, and $\Delta\theta_v$ for decoupled cases (d-e) is 1.47 K and 5.0 K, respectively.

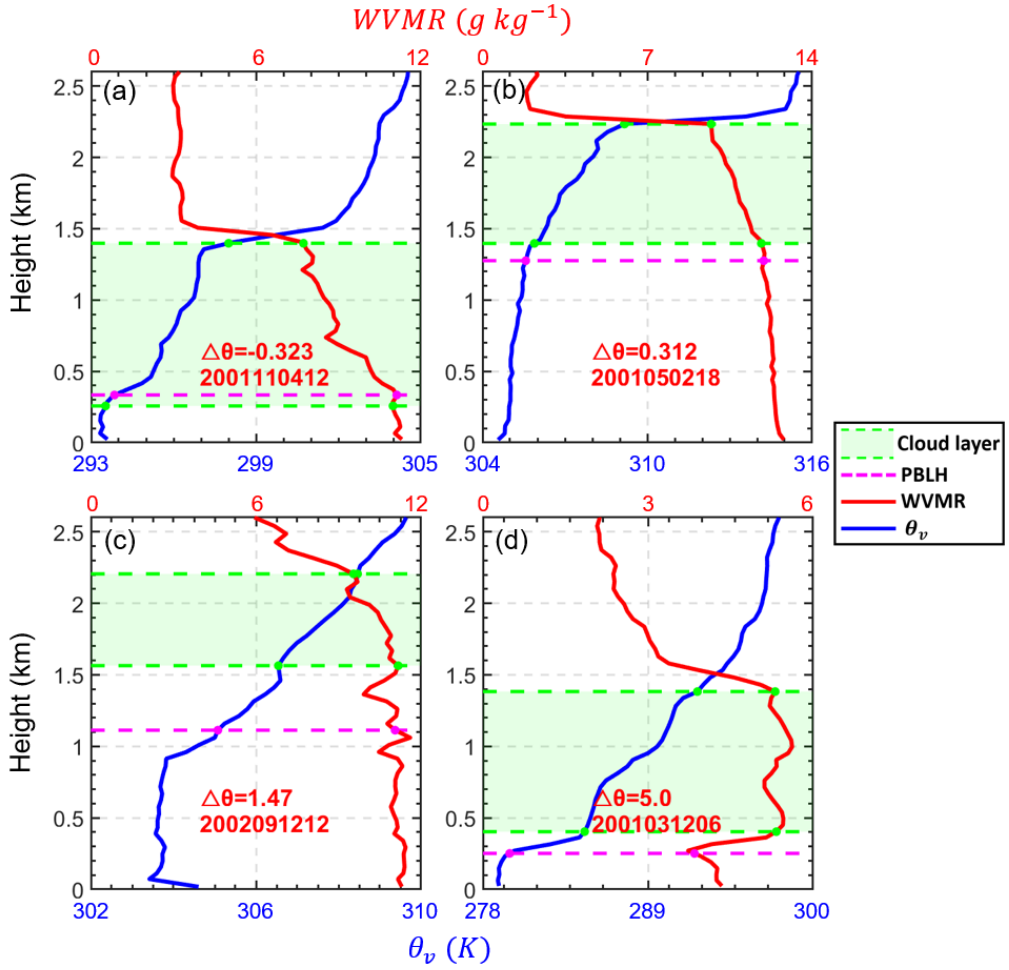


Figure 5.2: *Virtual potential temperature (θ_v, red lines) and water vapor mixing ratio (WVMR, blue lines) profiles from the radiosonde over the Southern Great Plains site for different cases. The differences in virtual potential temperature between the cloud base and the planetary boundary layer (PBL) top are expressed as $\Delta\theta_v$ ($\theta_v^{CBH} - \theta_v^{PBLH}$). The time of each radiosonde launch is marked in each panel. Green regions are cloud layers, and green dashed lines indicate their boundaries. PBL heights are marked as dashed pink lines.*

Therefore, instead of giving a height range to limit the differences between CBH and PBLH, we consider using the differences in θ_v between CBH and PBLH ($\Delta\theta_v = \theta_v^{CBH} - \theta_v^{PBLH}$) to determine the threshold for distinguishing coupled and decoupled clouds. For decoupled cases, the cloud base is above the capping inversion of the entrainment zone. There is a notable inversion in θ_v between PBL top and decoupled cloud base. Thus, we identify the cases satisfying $\Delta\theta_v > \delta_s$ as being in a decoupled state. Correspondingly, we identify the cases satisfying $\Delta\theta_v < \delta_s$ as being in a coupled state. We set the range of CBH to between 0 and 4 km and excluded cases of deep convection (i.e., CBH < 4 km and CTH > 6.5 km).

For identical cases, θ_v is constant within the PBL. Under this assumption, δ_s can be set as 0. However, there are certain variations in θ_v within the PBL, which can cause inversions with relatively small magnitudes between the cloud base and PBL top. Besides the capping inversion and surface inversion, the inversion strength within PBL is typically below 1K (Figure C.1a). Therefore, we set δ_s as 1 K, which is the same as the criterion for determining stable or convective conditions (Liu & Liang, 2010). Furthermore, we demonstrate the probability density function (PDF) of $\Delta\theta_v$ for the low cloud cases. Coupled and decoupled clouds are classified by the threshold of δ_s (1 K). Through the development of PBL, boundary layer clouds frequently occur in the entrainment zone, and form a coupled cloud-PBL system. For such coupled systems, θ_v at cloud top and PBL top is highly consistent for the majority of cases. Thus, the PDF of $\Delta\theta_v$ shows significantly high values for the range of -2 K to 0.5 K in the coupled regime. Meanwhile, the PDF of $\Delta\theta_v$ is evenly distributed in the decoupled regime. Since we only analyze low clouds, the PDF of $\Delta\theta_v$ slowly decrease when $\Delta\theta_v$ is above 10 K.

Based on the variations in θ_v within PBL, we set δ_s as 1 K. However, it should note that it is not an absolute value. A similar threshold of 0.5 K has been used for marine stratocumulus (Jones et al., 2011; Dong et al., 2015). Comparing to the marine condition, θ_v show greater variabilities over land. Hence, the threshold is correspondingly larger. On the other hand, since the threshold of 1 K is in the low PDF regime (Figure C.1b), the small changes in this value would not notably affect the identifications. Specifically, a 0.1 K difference in δ_s will lead to a 0.5% difference in the identification of coupled cloud.

5.2 Lidar-based method to identify coupled and decoupled clouds

5.2.1 Method description

Given the rapid change in clouds over land, radiosonde observations have limitations when it comes to tracking cloud development due to the coarse temporal resolution and drifting of the balloon. We thus further developed a lidar-based method to identify the coupled states of clouds based on our new algorithm for retrieving the PBLH that can better track the diurnal variations in PBLH than conventional lidar-based approaches (Su et al., 2020a). We adapted this algorithm for retrieving the PBLH and developed a new scheme to deal with cloudy conditions. The principles behind the PBLH algorithm are stated next for completeness.

Our new PBLH algorithm can retrieve the PBL variability from the MPL under Different Thermo-Dynamic Stability (thus named the DTDS algorithm) conditions, taking into account the vertical coherence and temporal continuity of the PBLH. First, we identify the LMPs (range: 0.25–4 km) in profiles of the wavelet covariance transform function derived from lidar backscatter (Brooks, 2003). These LMPs are the potential positions of the PBLH. We can use the PBLH derived from morning radiosonde as the starting point. Without morning radiosonde, the algorithm can still work well, with the lowest LMPs selected as the starting point, which reduces by 0.02–0.05 the correlation coefficient between MPL-derived and radiosonde-derived PBLHs (Su et al., 2020a).

To ensure good continuity, we select the closest LMP to the earlier position of the PBLH. Different stages of PBL development are considered. DTDS-derived PBLHs likely increase during the growth stage and decrease during the decaying stage, but the algorithm is also able to identify decreases during the growth stage or increases during the decaying stage based on the selection scheme described by Su et al. (2020a). There are multiple step signals in the backscatter profiles when complex aerosol structures (e.g., the residual layer) are present, leading to multiple LMPs. Based on temporal continuity, we select the appropriate LMP as the position of the PBL top. However, PBLH retrievals still suffer from relatively low accuracies under stable conditions because of the weak vertical mixing and residual layer.

Clouds induce strong step signals in the lidar backscatter, further considerably affecting PBLH retrievals. Su et al. (2020a) only considered cases where the low cloud top coincided with the previous PBL top, excluding other low-cloud cases (> 60% of all low-cloud cases). Here, we

specifically consider coupled and decoupled states of low clouds. Due to the MPL's ~0.2-km blind zone, we only analyze the PBLH and CBH above 0.2 km. Figure 5.3 presents the flow chart describing the updated DTDS algorithm. In particular, we jointly use PBL development and the lifted condensation level (LCL) to diagnose the states of coupling or decoupling. In ideal situations, LCL, PBLH, and CBH are highly consistent with each other for coupled clouds. But for real conditions, we only require that either the LCL or the PBLH coincides with the CBH for identifying coupled cases, with another parameter serving as an additional constraint. Specifically, a coupled cloud needs to occur within a certain range of LCL and the previous position of the PBL top. For the DTDS algorithm, five empirical parameters are used, including A_1, A_2, A_3, A_4, A_5. As listed in the Table 5.1, $A_1 - A_5$ are set as 0.7, 0.2, 0.15, 1.35, and 1.1, respectively. A cloud at time *i* is identified as being in the coupled state if the CBH is less than $[H(i-1) + 0.2 \text{ km } (A_2)]$ and $[\text{LCL} + 0.7 \text{ km } (A_1)]$. A cloud is also considered to be in a coupled state if the CBH is coincident with the LCL within 0.15 km (A_3), and the CBH is less than $[H(i-1) + 0.7 \text{ km } (A_1)]$, where $H(i-1)$ represents the PBLH at time $(i-1)$. The LCL is calculated from surface meteorological data (relative humidity, temperature, pressure) at the SGP site based on an exact expression (Romps, 2017).

Table 5.1. *List of parameters in the flow chart of DTDS (Figure 5). These parameters are related to three factors, including LCL, PBLH, CBH. The sensitivity of selection of these parameters is presented. The detailed impacts of variations in these parameters on the retrievals of cloud coupling and PBLH are presented in Figure 7 and Figure 8, respectively.*

	Unit	Related factors	Value	Sensitivity (coupled states)	Sensitivity (PBLH)
$\boldsymbol{A_1}$	km	LCL / PBLH	0.7	Low	Low
$\boldsymbol{A_2}$	km	PBLH	0.2	**High**	Low
$\boldsymbol{A_3}$	km	LCL	0.15	Low	Low
$\boldsymbol{A_4}$	dimensionless	CBH	1.35	Low	Low
$\boldsymbol{A_5}$	dimensionless	CBH	1.1	Low	**High**

After determining the coupling or decoupling state of a cloud, we retrieve $H(i)$ (i.e., PBLH at time *i*) based on the cloud state. For decoupled cases, we use the same strategy for a clear sky to

retrieve the PBLH. Based on the selection scheme in the DTDS algorithm, the LMP below the CBH is selected as $H(i)$. For coupled cases, we jointly use CBH and CTH to determine PBLH. During the warm season, shallow cumulus often occurs in the upper part of the PBL with strong surface heating, so the CTH can be regarded as the PBLH (Stull, 1988; Wallace & Hobbs, 2006). Under this condition, the CTH coincides with the previous PBL top. Therefore, if [CTH $<$ $PBLH_{30min} + 0.2\ km\ (A_2)$], we set $H(i)$ equal to the minimum between CTH and the product A_4*CBH, where $PBLH_{30min}$ is the average value of the PBLH within 30 min of the prior time i. Since this does not work for the active cloud (i.e., a cloud that reaches its level of free convection), we set $H(i) = A_5$CBH , if [CTH $\geq PBLH_{30min} + 0.2\ km\ (A_2)$]. After retrieving $H(i)$, we consider that the cloud above the PBLH is still coupled if [CBH $< H(i) + 0.2\ km\ (A_2)$] . Moreover, we added an upper limit for all PBLH retrievals. If [$H(i) >$ LCL $+ 0.7\ km\ (A_1)$], we adjust $H(i)$ as the maximum LMP below the LCL. The new DTDS method combines lidar measurements and surface meteorological observations and can simultaneously retrieve the PBLH and cloud states.

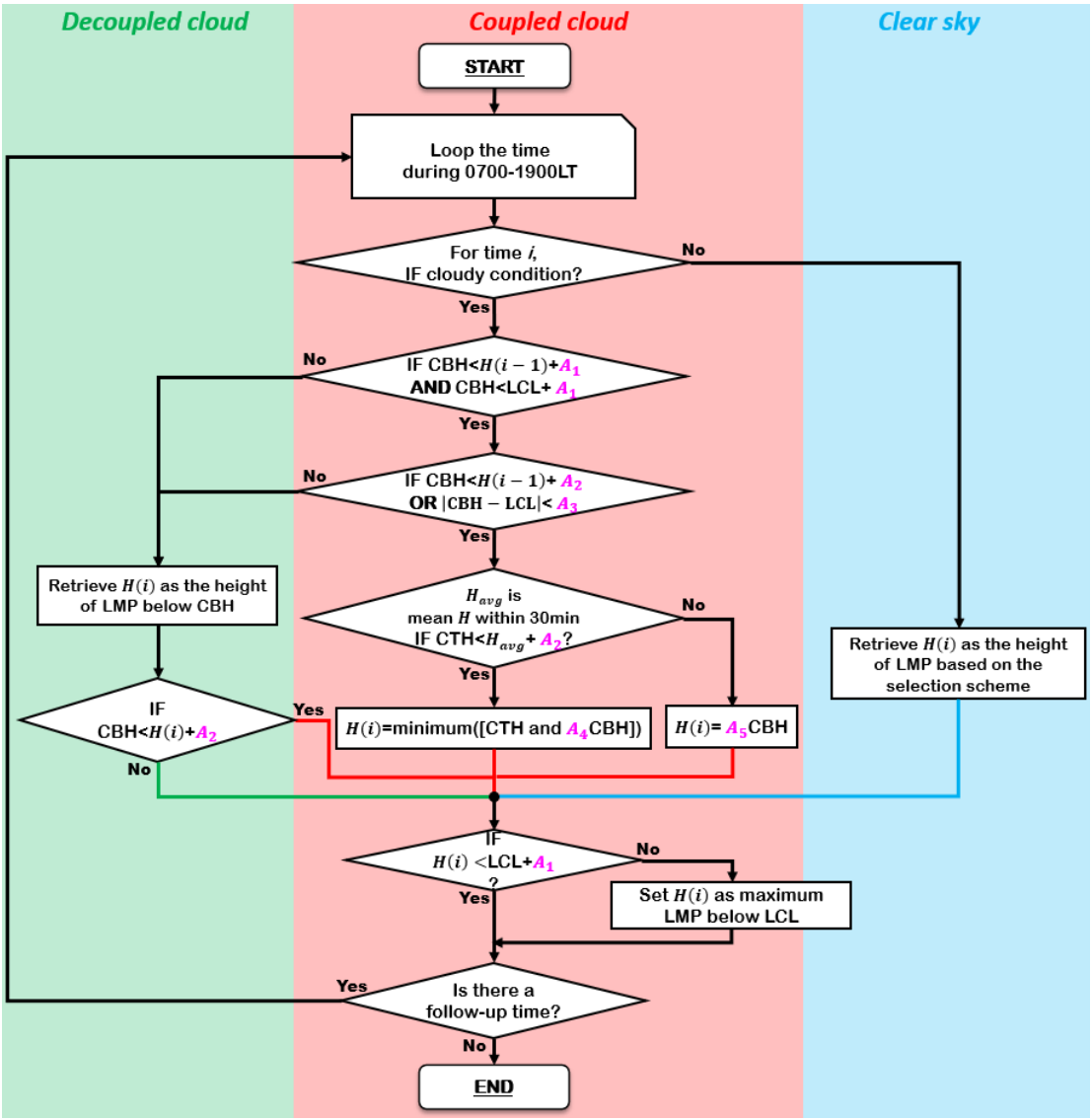


Figure 5.3: *The flow chart of the updated DTDS algorithm. In this diagram,* $H(i)$ *is the retrieved planetary boundary layer height (PBLH) at time* i. *CBH and CTH represent the base and top heights, respectively, of the lowest cloud at time* i. *The PBLH part for selecting the suitable local maximum position (LMP) follows Su et al. (2020a), and a detailed scheme for identifying a coupled cloud is added to the DTDS algorithm. LCL stands for lifted condensation level. Five empirical parameters* $(A_1, A_2, A_3, A_4, A_5)$ *are set as 0.7, 0.2, 0.15, 1.35, 1.1, respectively.*

5.2.2 Selection of empirical parameters in the algorithm

The states of coupling and decoupling are diagnostic parameters rather than explicit expressions. Similar to the other methods for retrieving PBLH (e.g., Brooks, 2003; Liu & Liang, 2010), multiple empirical parameters are used to determine PBLH. Here we discuss the selection of empirical parameters in the algorithm.

Note that we used the CTH and A_4*CBH as the upper limits for PBLH retrievals in the DTDS algorithm. For coupled cases, these two limits are generally close to or above the position of the PBL top. Only 2% (3%) of total cases meet the condition that the radiosonde-derived PBLH is 0.25 km higher than the CTH (A_4*CBH). Section 4 presents the detailed relationships between CBH, CTH, and PBLH. In the DTDS method, CTH serves as the upper limit for PBLH under the condition of coupled shallow cumulus.

Similar to previous studies, we can also use the LCL as the standard to identify coupled clouds (Dong et al., 2015; Zheng & Rosenfeld, 2015). We assume a cloud is coupled if $|\mathrm{CBH} - \mathrm{LCL}| <$ some criteria. Figure 5.4a shows the commission errors and omission errors for different criteria. Here, the commission error is calculated as the percentage of decoupled clouds misidentified as coupled clouds. The commission error can also be called a "false positive", as the former is a common term for describing the nature of an error in identification. The omission error is calculated as the percentage of coupled clouds that have not been identified under this criterion. By using the LCL, we can obtain a relatively low commission error if the criterion is less than 0.15 km and a relatively low omission error if the criterion is greater than 0.7 km. Thus, we set A_1 and A_3 as 0.7 and 0.15 in the DTDS method to exclude and to select cases of coupled clouds. We can also use the radiosonde-derived PBLH as the criterion (Figure 5.4b). Despite the coarse temporal resolution, the radiosonde-derived PBLH can be a good criterion to use to distinguish between coupling and decoupling. If we consider a coupled cloud as a cloud where (CBH < radiosonde-derived PBLH + 0.2 km), both commission and omission errors are ~5%. Therefore, we primarily use [PBLH+0.2 km (A_2)] in the DTDS method to identify coupled and decoupled regimes.

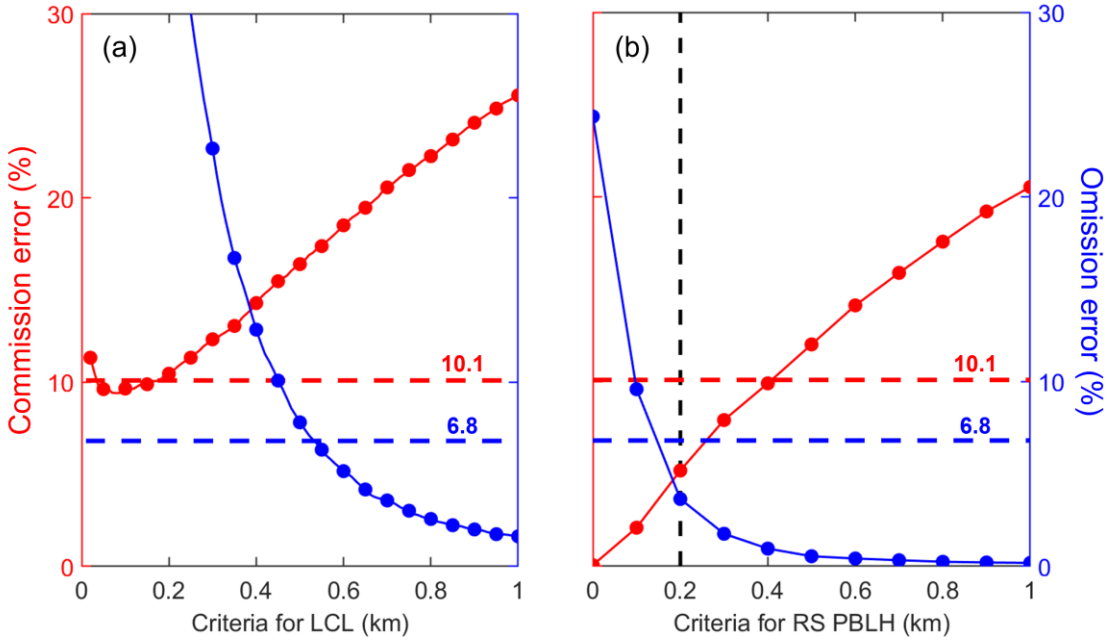


Figure 5.4: *Commission errors and omission errors of coupled cloud identifications (a) for different criteria for the lifted condensation level (LCL) and (b) for different criteria for the planetary boundary layer height (PBLH). "Criteria for LCL" means coupled clouds are identified if* $|CBH - LCL| <$ *Criteria for LCL. Similarly, "Criteria for radiosonde PBLH" means coupled clouds are identified if* $CBH - radiosonde\ PBLH <$ *Criteria for radiosonde PBLH. The red and blue dashed lines indicate the commission and omission errors, respectively, for the DTDS algorithm. CBH stands for cloud-base height, and radiosonde stands for radiosonde.*

As a by-product of this method, we also pay attention to the PBLH retrievals under cloudy conditions. Figure 5.5 presents the mean absolute biases and correlation coefficients between PBLH derived from lidar and radiosonde for selecting different values of empirical parameters. To match the scope of this study, we only analyze the low cloud conditions. For retrieving PBLH under cloudy conditions, A_2 is the critical parameter. The variations in correlation coefficients under different values of empirical parameters are small with a range of 0.81-0.82. However, the absolute biases can considerably differ under different values of A_5. In general, A_5 represents the ratio between CBH and PBLH under coupled conditions. If A_5 is above 1.1, PBLH retrievals under cloudy conditions are overestimated. We set A_5 as 1.1 to achieve a relatively low biases and a

relatively high correlation coefficient at the same time. For other parameters, the selections from reasonable ranges would not notably affect the PBLH retrievals.

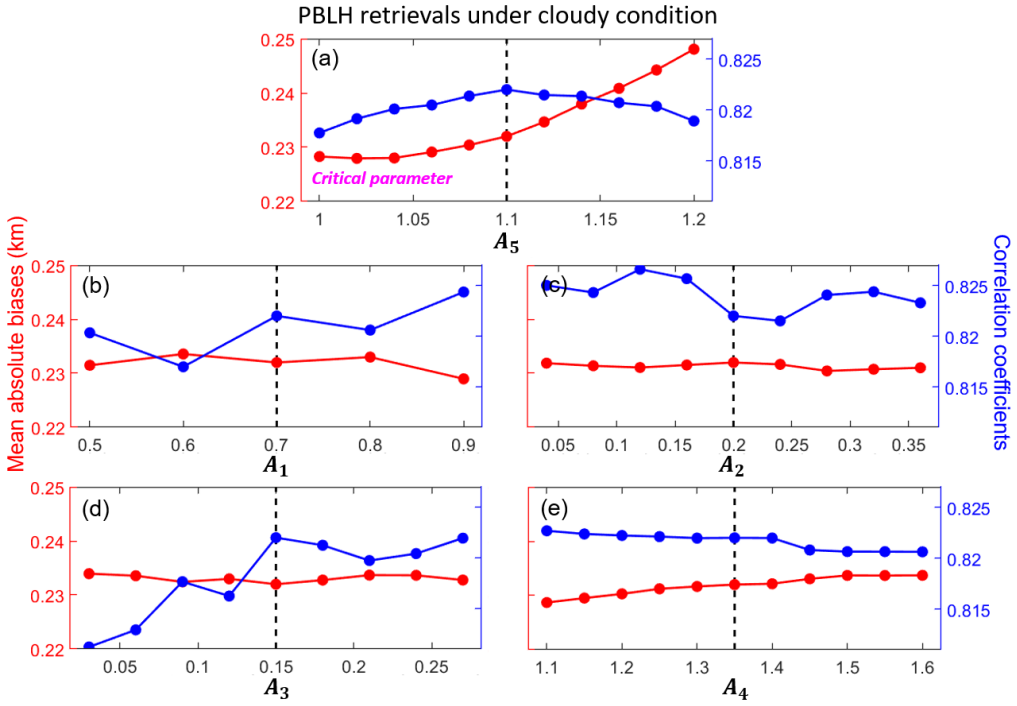


Figure 5.5: *The sensitivities of empirical parameters for PBLH retrievals. Red lines indicate the mean absolute biases between PBLH derived from lidar and radiosonde for selecting different values of empirical parameters (A_1, A_2, A_3, A_4, A_5) in the DTDS algorithm. Here, we only analyze the low cloud cases. Blue lines indicate the corresponding correlation coefficients between PBLH derived from lidar and radiosonde. Black dash lines indicate the default values. For each test, one parameter is variable, while other parameters are set as default values. For PBLH retrievals under cloudy conditions, A_5 is the critical parameter.*

In short, selections of these empirical parameters are based on the overall relationship between cloud and PBL under the coupled and decoupled states. In our method, the selection of A_2 is critical for the identifications of coupled clouds, while the selection of A_5 is critical for the PBLH retrievals under cloudy conditions. The selections of other parameters are not sensitive.

5.3 Retrievals of PBLH and coupling states under cloudy conditions

Figure 5.6 illustrates four examples of PBLH retrievals and cloud states derived from the DTDS algorithm for 27 October 2011, 31 July 2002, 19 March 2000, and 1 May 2012. Figure 5.6a depicts coupled shallow cumulus occurring at noontime at the PBL top. With a weak surface flux of ~200 W m^{-2}, this shallow cumulus cloud appeared for less than an hour. Figure 5.6b shows a developed coupled cumulus cloud. With a strong surface flux of ~500 W m^{-2}, this coupled cloud continuously developed during the daytime. Figure 5.6c presents the case of a daylong coupled cloud. After the passage of a frontal system that day, stratocumulus occurred during the morning with a cloud thickness of 0.5 km. Through the development of the PBL, the thick stratocumulus cloud was broken up by the strong turbulences, transforming into shallow cumulus clouds. Figure 5.6d shows the case of an active coupled cloud, which is generally associated with a large amount of convective available potential energy. Even though coupled clouds can differ in appearance and variability throughout the day, the common feature is the coherent variation between the cloud base and the PBL top. The LCL is a relevant parameter and can differ from the PBLH and the CBH for some coupled cases (e.g., Figure 5.6b-c).

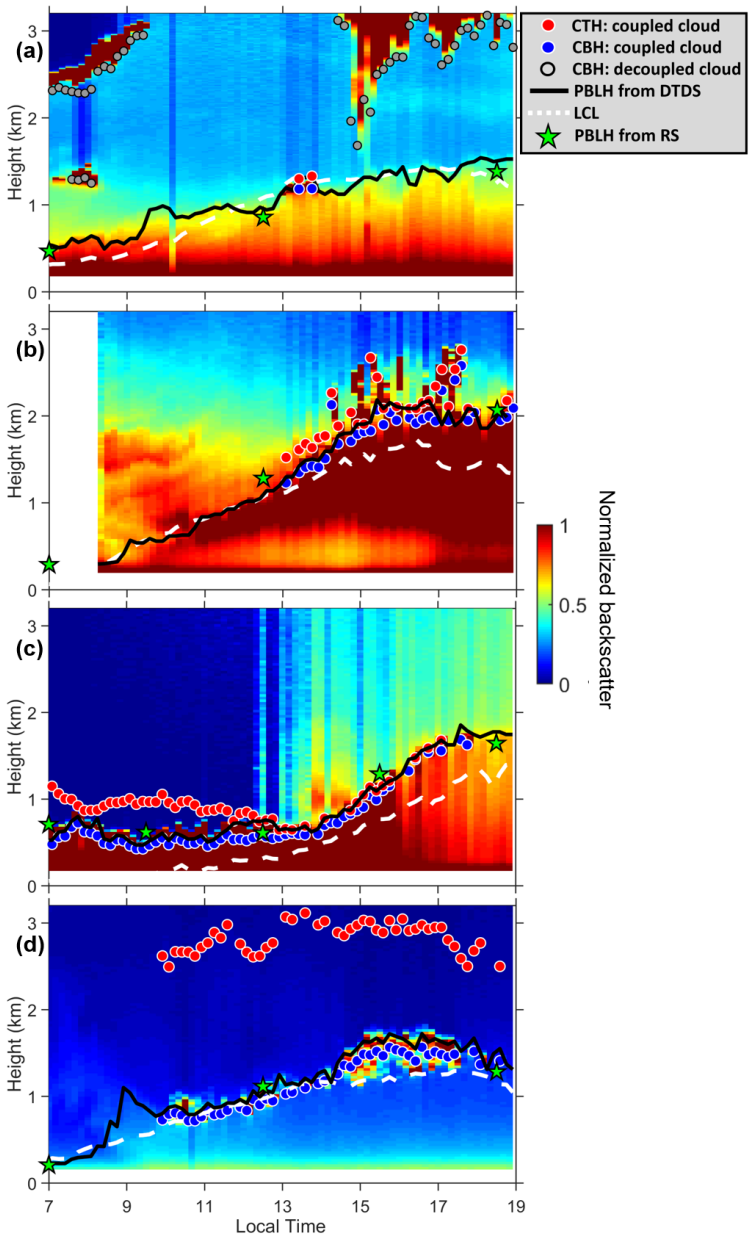


Figure 5.6: *Daily backscatter profiles: (a) short-lived coupled cloud, (b) developed coupled cloud, (c) daylong coupled cloud, and (d) active coupled cloud. Backscatter is normalized to a range of 0–1 in arbitrary units. Red dots and blue dots indicate cloud-top heights (CTHs) and cloud-base heights (CBHs) of coupled clouds. Grey dots mark CBHs for decoupled clouds. Black lines and green stars mark the planetary boundary layer height (PBLH) retrieved from the DTDS algorithm and from radiosonde soundings, respectively. White dashed lines represent lifted condensation levels (LCLs).*

The identification accuracy, or disparity between different methods, are evaluated in terms of the selected criteria, for which the identification method based on $\Delta\theta_v$ is regarded as the "truth". The commission error is 10.1%, and the omission error is 6.8% for the DTDS method. Note that lidar-based PBLH methods generally suffer from relatively low accuracy under stable atmospheric conditions. Following Liu and Liang (2010), we identified stable PBLs from radiosonde measurements. Since coupled clouds are driven by relatively strong buoyancy fluxes, only 1% of total cases of coupled clouds occurred under stable PBL conditions during the study period (0700–1900 LT). Therefore, the relatively low accuracy for stable PBLs is not a major problem in this study.

Figure 5.7 shows scatter plots between CBH, CTH, PBLH, and LCL for coupled and decoupled clouds. For coupled clouds, there is a generally strong correlation between CBH, LCL, and PBLH, contrary to the weak relationships of decoupled cases. The relationship between CTH and radiosonde-derived PBLH is complicated. For shallow cumulus clouds, their tops can be considered as PBL tops for the coupled state, while the cloud top is considerably above the position of the PBL top for active cumulus clouds. We also note that the accuracy of CTH retrievals is generally lower than the accuracy of CBH retrievals (Clothiaux et al., 2000). As CTH is not a criterion for cloud coupling, the accuracy of CTH would not affect the identification of coupled cloud, but may affect the PBLH retrievals for the coupled cloud cases.

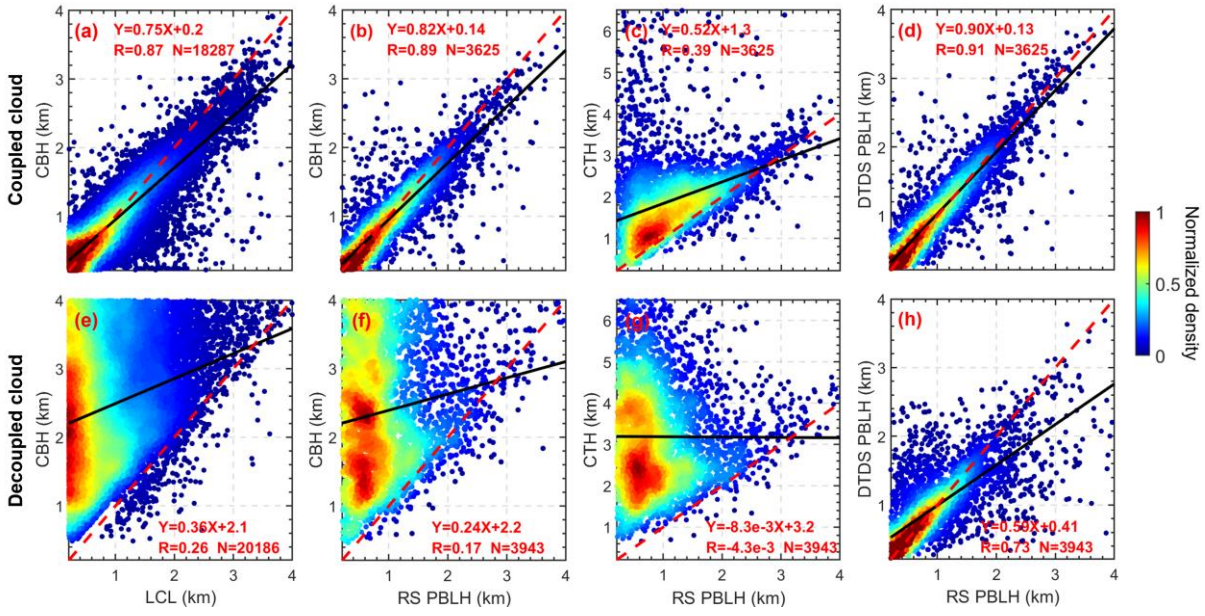


Figure 5.7: *The relationships between (a) LCL and CBH, (b) CBH and radiosonde-derived PBLH, (c) CTH and radiosonde-derived PBLH for coupled clouds, and (d) DTDS-derived PBLH and radiosonde-derived PBLH. Panels (e-h) are similar to panels (a-d) but for decoupled clouds. Black lines represent the linear regressions. The linear fitting functions, correlation coefficients (R), and sampling numbers (N) are given in each panel.*

For coupled clouds, the DTDS-derived PBLH relies on the cloud position, showing a strong correlation with the radiosonde-derived PBLH with a correlation coefficient (R) of ~0.9 (Figure 5.7d), since the position of the coupled cloud serves as a good reference for determining the PBLH. For decoupled cases, the correlation between the DTDS-derived PBLH and the radiosonde-derived PBLH is generally good (R = 0.73) but worse than the correlation for coupled cases. As pointed out in previous studies (Chu et al., 2019; Hageli et al., 2000; Lewis et al., 2013; Su et al., 2017a), it has been a persistent problem to retrieve the PBLH under cloudy conditions since the large backscatter and step signals from cloud interference would be excluded to avoid interfering with the retrievals. After identifying the coupling or decoupling of clouds, the PBLH can be successfully retrieved under cloudy conditions. Specifically, in coupled conditions, the position of the coupled cloud serves as a good reference for identifying the PBLH. In decoupled conditions, the large backscatter and step signals from clouds would be excluded to avoid interfering with the retrievals. With our method, PBLH retrievals under cloudy conditions are well handled during the study period (1998-2019).

Compared to the clear-sky cases discussed in previous studies (e.g., Chu et al., 2019; Yang et al., 2017), the DTDS-derived PBLH shows a much higher correlation with radiosonde-derived PBLH for coupled cloud cases and has a similar R as the radiosonde-derived PBLH for decoupled cloud cases. Moreover, due to the different definitions of the PBLH and aerosol stratification within the PBL, there are always considerable differences between lidar- and radiosonde-derived PBLHs, which cannot be eliminated by a specific algorithm (Chu et al., 2019; Su et al., 2020a).

Chapter 6 Investigation of aerosol-PBL interactions

The aerosols and PBL remote sensing techniques presented in previous chapters lay a solid foundation for investigating aerosol-PBL interactions. By fusing the PBLH product developed in Chapter 4 with synergistic field measurements, this chapter addresses some important scientific questions concerning the API, which is critical for understanding the accumulation and evolution of surface pollutants. The two key parameters of our study, AOD and PBLH, are both closely related to surface PM2.5 (Su et al., 2017c). Therefore, we analyze the relationships between PBLH, surface pollutants, and AOD over China in Section 6.1. Regional patterns and influencing factors for these relationships are comprehensively discussed. Furthermore, Section 6.2 investigates the role of aerosol vertical distribution in thermodynamic stability and PBL development by jointly using observations and a radiative transfer model. In Section 6.3, we present the association between PBL and air pollution during the Coronavirus Disease 2019 (COVID-19) lockdown over China. It should note that our AOD and surface-cloud coupling methods were not applied in Chapter 6, but are expected to be utilized in future studies relating to API.

6.1 Relationships between PLBH and surface pollutants over China

Although several studies have explored the relationship between PBLH and surface pollutants in China, the interaction between PBLH and surface pollutants under different topographic and meteorological conditions is not well characterized. To address this issue, this section presents a comprehensive exploration of the relationship between the PBLH and surface pollutants over China, for a wide range of atmospheric, aerosol, and topographic conditions.

6.1.1 Climatological patterns of PBLH and surface pollutants

The climatology of the PBLH, especially its seasonal variability, is very important for air-pollution-related studies. We utilized the CALIPSO measurements from 2006 through 2017 to represent the spatial distribution of seasonal mean PBLH with interpolation, as shown in Figure 6.1. A horizontal smoothing window of 20 km was applied to the original PBLH data at 1/3 km horizontal resolution. The seasonal climatological patterns of MERRA-derived PBLH are presented in Figure 6.1e-h for the same period. In general, the climatological pattern of MERRA PBLH is similar to that of CALIPSO, though the MERRA values are higher in spring and summer,

and the peak values are lower in autumn and winter. Both CALIPSO and MERRA PBLHs are generally shallower in winter, when the development of the PBL is typically reduced due to the weaker solar radiation reaching the surface, and are generally higher in summer, especially for inland regions. Correspondingly, the spatial distributions of seasonal mean $PM_{2.5}$ also exhibit large spatial and seasonal variations (Figure D.1). The $PM_{2.5}$ seasonal pattern is generally coupled to that of PBLH; the lowest values occur in summer and the highest in winter. As a high PBLH facilitates the vertical dilution and dissipation of air pollution, the contrasting patterns of PBLH and $PM_{2.5}$ are qualitatively consistent with expectations. Meanwhile, the seasonality of PM2.5 also is contributed by other factors, such as aerosol vertical mixing, humidity, and horizontal transport.

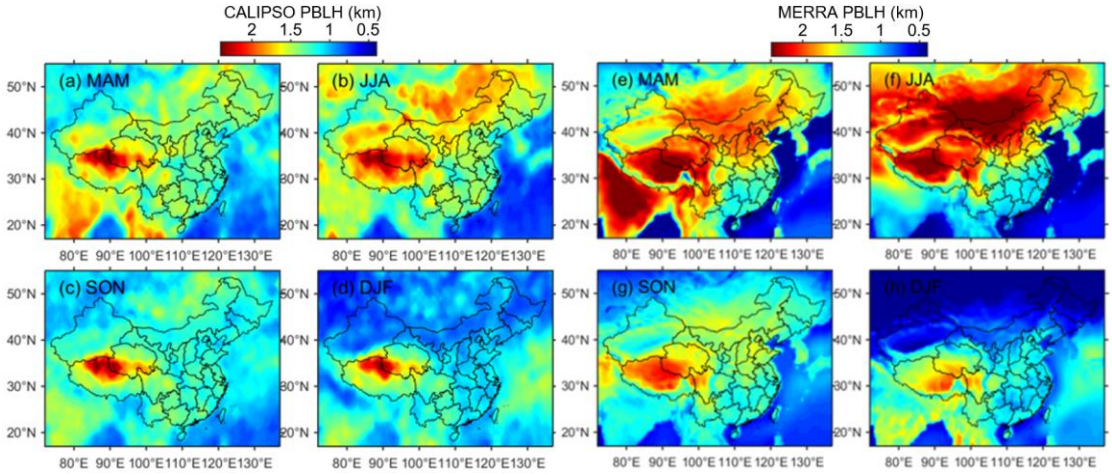


Figure 6.1. *Spatial distributions of climatological mean PBLH derived from CALIPSO for (a) March-April-May (MAM), (b) June-July-August (JJA), (c) September-October-November (SON), and (d) December-January-February (DJF) during the period 2006–2017. Spatial distributions of climatological mean of early-afternoon PBLH obtained from MERRA for (e) MAM, (f) JJA, (g) SON, and (h) DJF during the same period.*

From the seasonal climatologies, we find a coupling pattern between PBLH and $PM_{2.5}$, although one cannot assume a causal relationship from these plots alone. In subsequent sections, we use the lidar PBLH retrievals to investigate the PBLH-PM relationships in more detail.

6.1.2 Regional relationships between PM and PBLH

If the common factor driving large-scale variations in both PM and PBLH is meteorology, a regional analysis of their relationship could elucidate the meteorological impacts. We investigate

the CALIPSO-PBLH and surface $PM_{2.5}$ data case by case. By matching the available CALIPSO retrievals within 35 km of the surface $PM_{2.5}$ observations, we show the scatterplots for PBLH versus surface $PM_{2.5}$ for the four ROIs in Figure 6.2. Despite the overall negative correlations, the correlations between PBLH and $PM_{2.5}$ have large spreads and differences. Several factors may contribute to the regional differences, as noted in the following discussions. Both regular linear regression and inverse fit are applied to characterize the PBLH-PM relationships. Significant negative correlations between $PM_{2.5}$ and PBLH are found over NCP with a Pearson correlation coefficient of -0.36. In addition, the nonlinear inverse function shows high consistency with the average values for each bin, and characterizes the PBLH-PM relationship with a somewhat higher correlation coefficient (-0.49). PBLH also shows a significant negative correlation with $PM_{2.5}$ over YRD and NEC, whereas the weak PBLH correlation with $PM_{2.5}$ over the PRD is not statistically significant (significance level>99%) . The correlation coefficients for the inverse fit are generally larger than the Pearson correlation coefficients, indicating that the nonlinear fit may be more suitable for characterizing the PBLH-PM relationships. Such improvements are obvious for NCP and YRD, but are not significant over PRD and NEC. The non-linearity of the relationships is notably contributed by the high PM2.5 concentration when the PBLH is below 1 km. The aerosol vertical structure also should play an important role in the PM-PBLH relationships, which will be discussed in the next section.

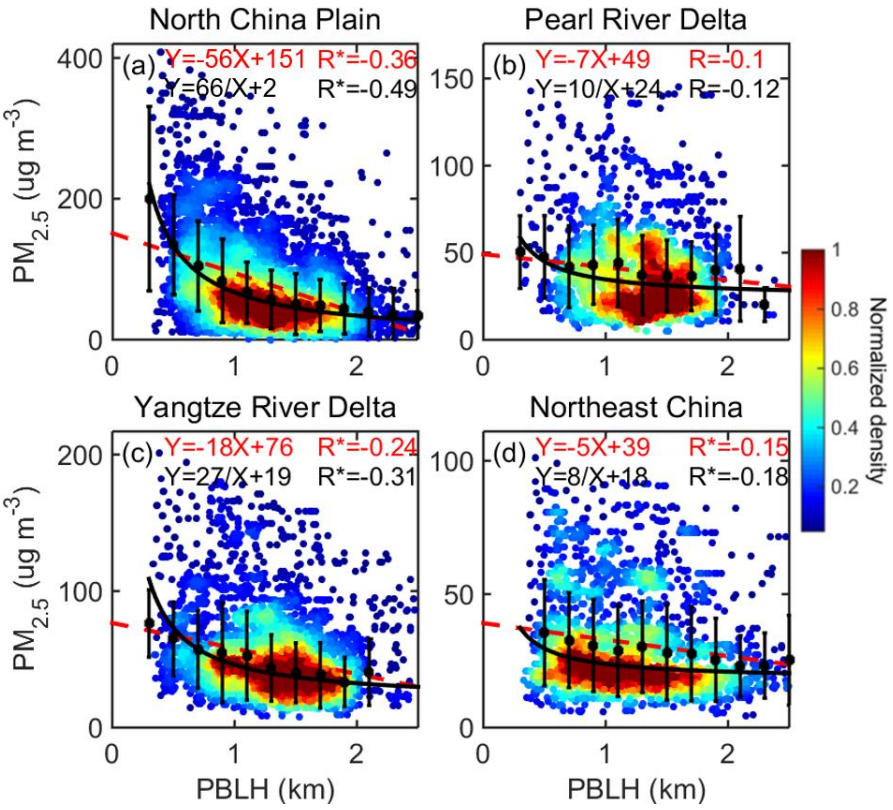


Figure 6.2: *The relationship between CALIPSO-derived PBLH and early-afternoon $PM_{2.5}$ over (a) NCP, (b) PRD, (c) YRD, and (d) NEC. The black dots and whiskers represent the average values and standard deviation for each bin. The red dash lines indicate the regular linear regressions, and the black lines represent the inverse fit ($f(x) = A/x + B$). The detailed fitting functions are given at the top of each panel, along with the Pearson correlation coefficient (red) and the correlation coefficient for the inverse fit (black). Here and in the following analysis, R with asterisks indicates the correlation is statistically significant at the 99% confidence level. The color-shaded dots indicate the normalized sample density.*

We note that the ranges of $PM_{2.5}$ for these ROIs are significantly different; therefore, the background pollution level is likely to be an important factor for the PBLH-PM relationship. We thus normalize the $PM_{2.5}$ by MODIS AOD, a widely used parameter to represent the total-column aerosol amount, to qualitatively account for background or transported aerosol that is not concentrated within the PBL. The relationships between PBLH and $PM_{2.5}$/AOD over four ROIs are presented in Figure 6.3. Clearly, after normalizing $PM_{2.5}$ by AOD, the spread of these scatter

plots and the regional differences are significantly reduced, and the correlations become more significant for all ROIs, especially for PRD. This is because transported aerosol aloft can contribute to variability in total column AOD that is unrelated to the PBLH.

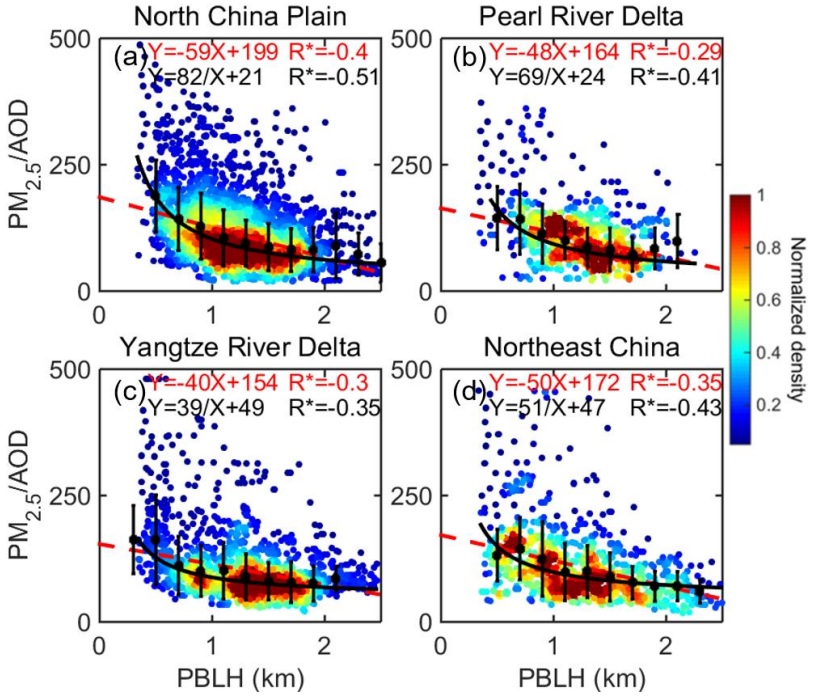


***Figure 6.3**: Similar to Figure 6.2, but for the relationship between CALIPSO PBLH and early-afternoon $PM_{2.5}$/AOD (unit: $\mu g\ m^{-3}$ per AOD) over four ROIs. Here, the AOD data are obtained from MODIS.*

As the NCP experiences the most pronounced seasonality in both PBLH and $PM_{2.5}$, the relationship over this region also shows the most prominent seasonal differences (Figure D.2). Figure 6.4 focuses on the seasonal dependence of the PBLH and $PM_{2.5}$ relationship over the NCP. The magnitude of the slope between $\frac{1}{\text{PBLH}}$ and $PM_{2.5}$ for this region is ~90 (unit: km*ug m^{-3}) with a correlation coefficient of -0.55 during winter, and is only ~40 in summer. For comparison, the seasonally aggregated relationship between PBLH and $PM_{2.5}$ is presented in Figure 6.4e. $PM_{2.5}$ concentrations do not increase linearly with decreasing PBLH. Specifically, $PM_{2.5}$ increases rapidly with decreasing PBLH when PBLH is lower than 1 km, but changes much more slowly for PBLH > 1.5 km. As we noted previously, such a relationship could be contributed by the aerosol vertical mixing. The seasonal mean values for $PM_{2.5}$ and PBLH are presented as colored

dots in Figure 6.4e, and the whiskers represent the standard deviations. For winter, the PBLH is generally shallow, $PM_{2.5}$ concentrations are high, and thus PBLH shows the most significant negative correlation with $PM_{2.5}$. Conversely, in summer, the PBLH is generally higher, $PM_{2.5}$ concentrations are lower, and the PBLH-$PM_{2.5}$ relationship is virtually flat. Such seasonally distinct PBLH-$PM_{2.5}$ relationships have not previously been studied quantitatively, and have the potential for improving $PM_{2.5}$ monitoring and predictions.

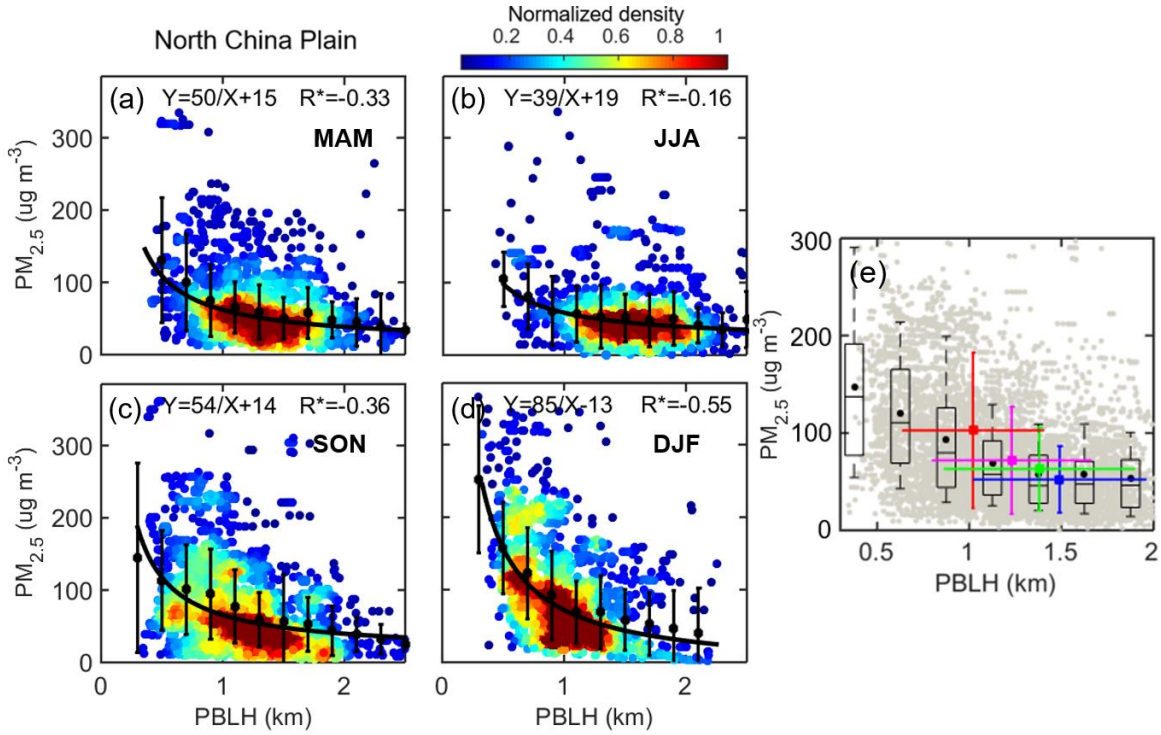


***Figure 6.4**: The relationship between CALIPSO PBLH and $PM_{2.5}$ over the NCP for (a) MAM, (b) JJA, (c) SON, and (d) DJF. (e) The general relationship between $PM_{2.5}$ and PBLH aggregated over all seasons, with individual observations for each day plotted as gray dots. The box-and-whisker plots show 10th, 25th, 50th, 75th, and 90th percentile values of $PM_{2.5}$ for each bin. The green, blue, pink, and red dots present the mean values for MAM, JJA, SON, and DJF, respectively.*

6.1.3 Influencing factors for PBLH–$PM_{2.5}$ relationships

The PBLH affects mainly the vertical mixing and dispersion of air pollution, but horizontal transport also plays a critical role in surface air quality. Figure 6.5a-b present the PBLH-$PM_{2.5}$ relationships over China under strong wind (WS>4m s^{-1}) and weak wind (WS<4m s^{-1}) conditions. Under strong wind conditions, $PM_{2.5}$ is found to be much less sensitive to PBLH than for weak

wind. In addition, Figure 6.5c-d show the aerosol extinction profiles as a function of PBLH under strong and weak wind conditions, as retrieved by the MPLs at Beijing, with the Klett method applied (Klett, 1985). In both strong and weak wind conditions, we found clear aerosol extinction gradients appear at the top of the PBL. Nonetheless, under strong wind, the aerosol extinction is typically low in the PBL, and the surface extinction does not change significantly with different PBLH. In this situation, the strong wind likely plays a dominant role in affecting $PM_{2.5}$ concentrations by ventilating the PBL. Under weak wind, the response of near-surface pollutants to PBLH is more nonlinear, and both aerosol extinction and $PM_{2.5}$ fall rapidly as the PBLH increases from 600m to 1200m.

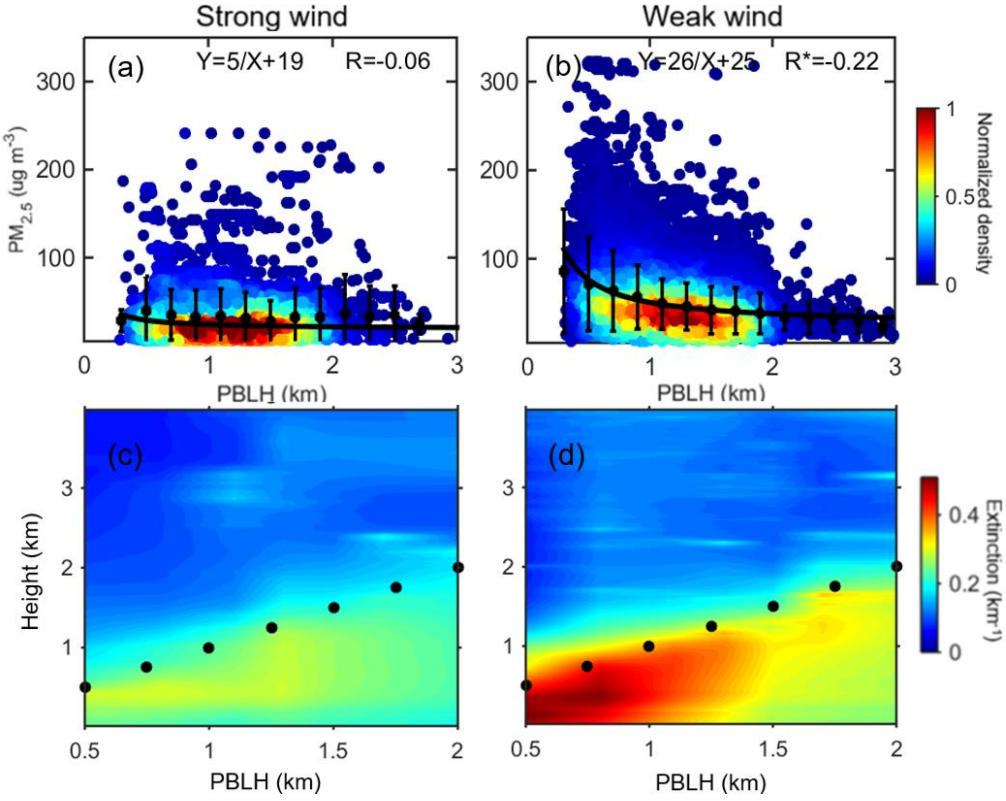


Figure 6.5: *The relationship between CALIPSO PBLH and $PM_{2.5}$ over China for (a) strong wind (WS>4m s^{-1}) and (b) weak wind (WS<4m s^{-1}). The aerosol extinction profiles at ~550 nm derived from MPL at Beijing change with different MPL-derived PBLH under (c) strong wind and (d) weak wind conditions. In (c, d), the black dots indicate the location of PBL top.*

We further consider the relationship between PBLH-$PM_{2.5}$ under different wind-direction regimes for Beijing. Two different regimes are easy to identify: a northerly wind and a southerly

wind; these are divided by the red line in Figure 6.6a. The northerly air comes from arid and semiarid regions in northwest China and Mongolia, and is usually strong and clean. The southerly wind comes from the southern part of the NCP, with high humidity and aerosol content. To relate the connections between WS, PBLH, and surface air quality, at least qualitatively, the ventilation rate (VR) can be represented as VR = WS × PBLH (Tie et al., 2015). By jointly considering the boundary layer extent and horizontal transport, the VR represents how fast the air pollution can dissipate within the boundary layer. Figure 6.6b-c and d-e present the PBLH-$PM_{2.5}$ and VR-$PM_{2.5}$ relationships under southerly wind and northerly wind conditions, respectively. For all wind conditions, VR shows a reciprocal relationship with surface $PM_{2.5}$. Under northerly wind conditions, both PBLH-$PM_{2.5}$ and VR-$PM_{2.5}$ relationships are flatter and have lower correlation coefficients. The northerly wind is apparently effective in removing pollutants and may play a dominant role in affecting air quality. For the southerly wind, the $PM_{2.5}$ concentrations are highly sensitive to PBLH and VR values.

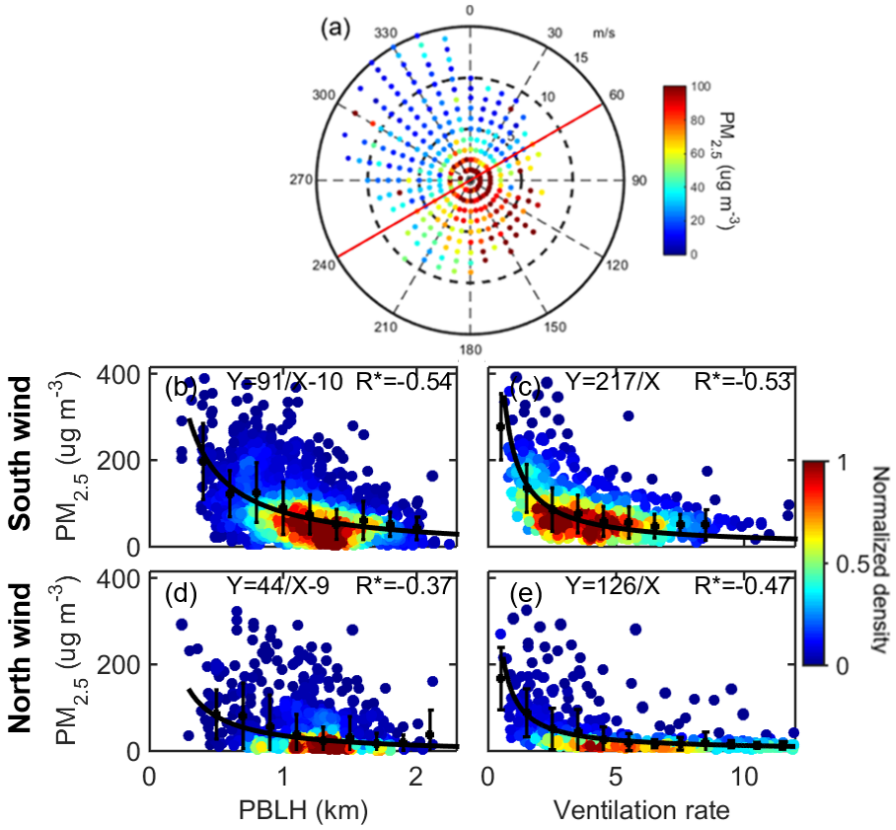


***Figure 6.6**: (a) Relationship between wind direction/wind speed and $PM_{2.5}$ over Beijing. The red line divides the northerly wind and southerly wind. (b-c) The relationship between $PM_{2.5}$ and MPL-*

*PBLH/ventilation rate (VR = WS × PBLH, unit: km*m s^{-1}), for southerly winds over Beijing. (d-e) The relationship between $PM_{2.5}$ and MPL-PBLH/VR, for northerly winds over Beijing.*

The PBL structure and $PM_{2.5}$ concentrations can both be affected by topography. We divided the sites into two categories based on elevation: plains (elevation < 0.5 km) and highland (elevation > 1 km). Figure 6.7a-d presents the correlation coefficients and slopes in the inverse fit between $PM_{2.5}$ and PBLH for the plains and highland areas. For calculating the correlation coefficient and slope, we require that the number of matched CALIPSO PBLH and $PM_{2.5}$ samples is larger than 15 for each site. Much higher correlation coefficients are found in the plains than those in the highlands, and the slope (i.e., linear slopes between $-\frac{1}{\text{PBLH}}$ and $PM_{2.5}$) in the plains is ~3 times that in highlands. A reciprocal relationship is shown between station elevation and the slope between $-\frac{1}{\text{PBLH}}$ and $PM_{2.5}$ (Figure 6.7e). The magnitudes of slopes decrease dramatically with elevation increase, for elevations between 0 and 500 m.

Local emissions also affect aerosol loading, and differences between plains and highland areas regarding local source activity could be important here as well. Figure 6.7e shows that the low-elevation regions are typically more polluted than highland areas, and the magnitudes of the slopes tend to be higher. As stronger correlations for PBLH-$PM_{2.5}$ relationships are found over polluted regions, the strong local emission over plains areas is likely to be another factor contributing to the strong correlation between PBLH and $PM_{2.5}$, whereas the low $PM_{2.5}$ concentrations may contribute to the weak PBLH-$PM_{2.5}$ correlation over the highlands.

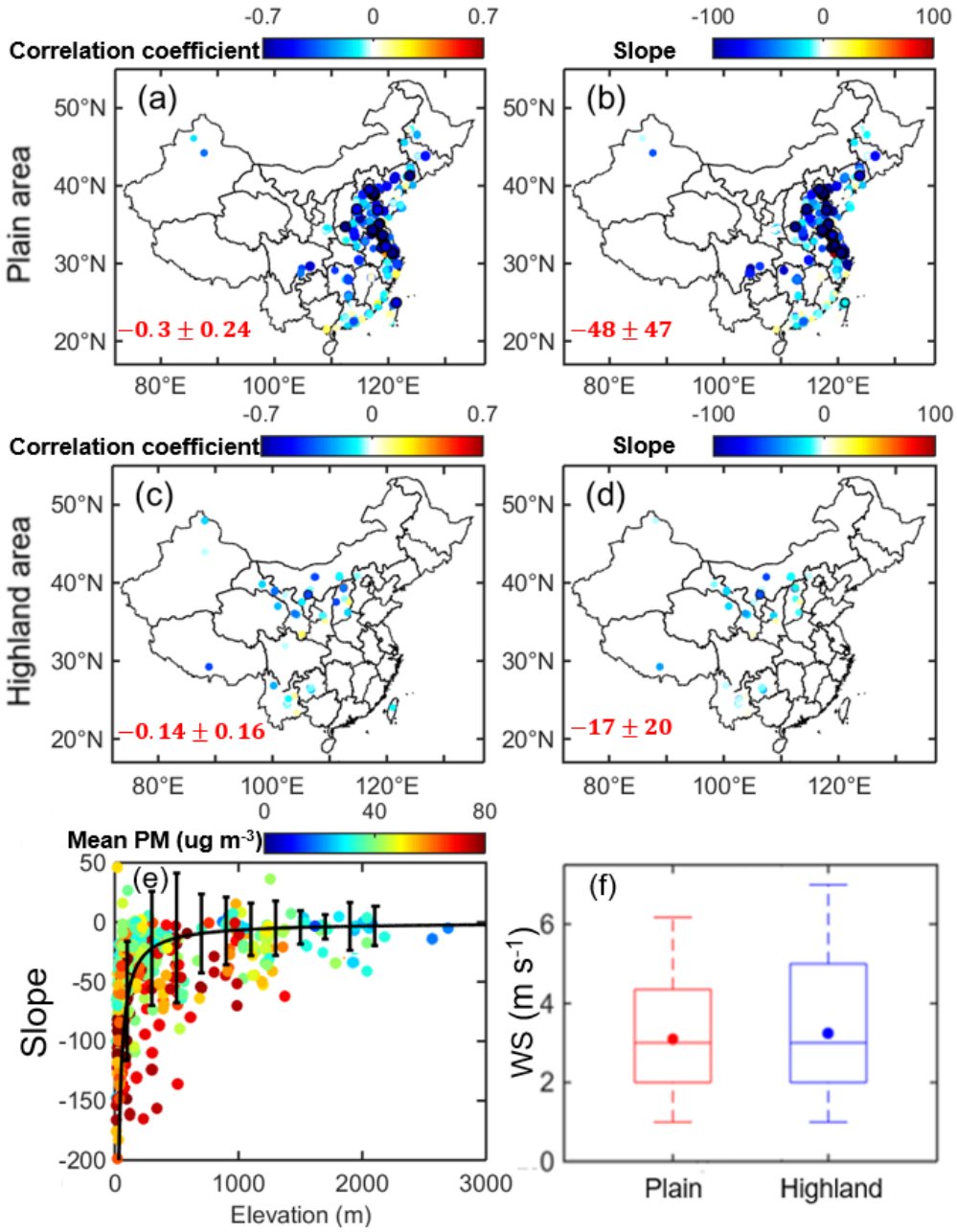


Figure 6.7: *Stratification by terrain elevation. The correlation coefficients (R) and slopes (unit: km*ug m^{-3}) between CALIPSO PBLH and $PM_{2.5}$ for the inverse fit ($f(x) = {}^{A}/_{x} + B$) are shown for the (a-b) plains and (c-d) highland areas. Noted the slope in the inverse fit is defined as – A. (e) The slopes in the inverse fit (i.e., linear slopes between $-\frac{1}{PBLH}$ and $PM_{2.5}$) under different station elevations, with color-shading indicating station mean $PM_{2.5}$ concentration. (f) Box-and-whisker plots showing the 10th, 25th, 50th, 75th, and 90th percentile values of the early-afternoon WS for plain and highland regions. The dots indicate the mean values.*

In addition, horizontal transport is associated with topography. Thus, we illustrate the distribution of WS for plains and highland areas in Figure 6.7f. WS is generally larger for highland areas, especially for the strongest wind cases. In fact, the 10% and 25% quantiles of WS are nearly the same between plains and highland areas, whereas there are clear differences in the 75% and 90% quantiles. Strong wind cases account for 37% of the total over highland areas, but only 27% of the total over the plains. As discussed previously, a strong wind can effectively remove surface pollutants, and can play a dominant role in determining local pollution levels. In this situation, PBLH might not play as critical a role in $PM_{2.5}$ concentrations. Thus, mountain winds, along with less local emission, are likely to be leading factors accounting for the differences in PBLH-$PM_{2.5}$ correlations between plains and highland areas.

6.2 The role of aerosol vertical structure in API

As mentioned in the previous section, despite the general negative correlations, the relationship between PBL and surface pollutants is subject to confounding factors such as aerosol, the underlying surface, and meteorological conditions. Among these factors, the vertical variabilities in aerosol loading are especially complex and challenging to estimate, and hence, serve as an important uncertainty source in estimating aerosol radiative effects. By using comprehensive observations, this section particularly addresses this issue (Su et al. 2020c). Based on the observational constraint, a radiative transfer model is used to simulate the vertical profiles of aerosol radiative forcing that are employed to investigate the impact of aerosols on buoyancy in the lower atmosphere.

6.2.1 Classification of different aerosol structure scenarios

By altering the adiabatic heating rate of the atmosphere, the aerosol vertical distribution is of great importance to the PBL. Based on cloud-free AEC profiles in the PBL, aerosol vertical structures can be classified into three types: well-mixed, decreasing with height, and its inverse, increasing with height. If AEC varies by less than 20% within the lowest 80% of the PBL, it is considered a well-mixed structure. For the other cases, a decreasing structure indicates a peak in AEC near the surface, and the inverse structure indicates a peak in AEC in the middle or upper PBL.

To investigate the vertical variation in AEC within the PBL, the evolution of the PBLH has to be taken into account. Following previous studies (Ferrero et al., 2014; Kuang et al., 2016), vertical profiles were normalized by introducing a standardized height (H_s), calculated as follows:

$$H_s = \frac{z - PBLH}{PBLH}, \qquad (6.1)$$

where z is the height above the ground, and H_s is 0 at the PBL top and -1 at ground level. Figure 6.8 shows the normalized vertical profiles of AEC derived from MPL data for different aerosol structures around noontime. The number of samples and percentages of decreasing, well-mixed, and increasing aerosol structures are 998 (51%), 611 (32%), and 330 (17%), respectively. Since a temperature inversion located at the PBL top traps moisture and aerosols, there is a sharp decrease in the AEC profile from the PBL upper boundary to the free atmosphere. Variations in the aerosol vertical distribution largely depend on different conditions, but share similar features among the different aerosol structure patterns. Despite complex aerosol vertical distributions, these three types of profiles can account for most of the cloud-free cases.

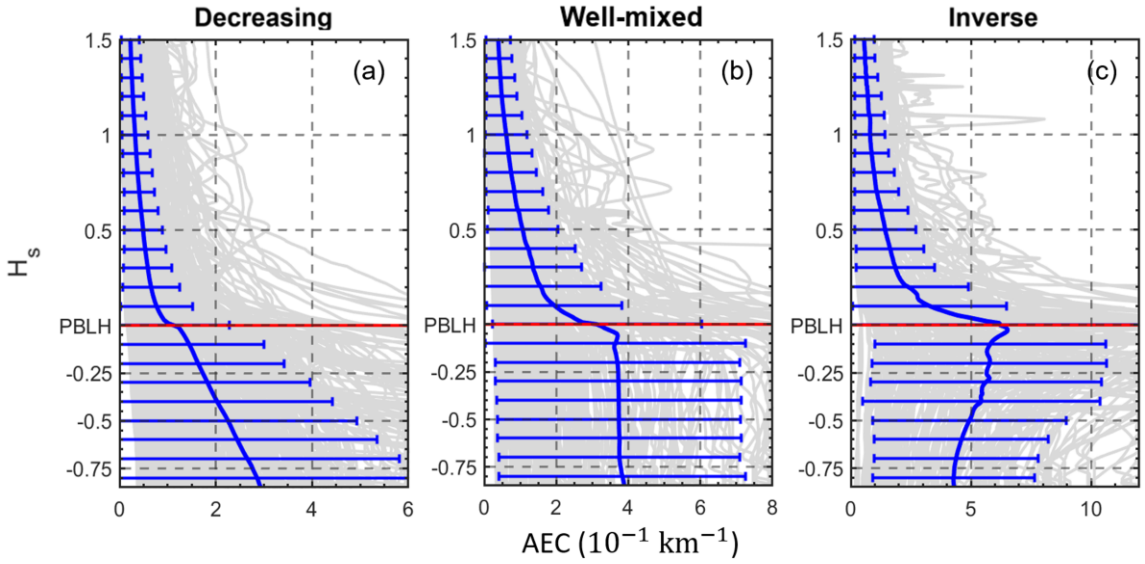


Figure 6.8: *Normalized vertical profiles of aerosol extinction coefficient (AEC) for (a) decreasing, (b) well-mixed, and (c) increasing (i.e., inverse) aerosol structures. Red lines mark the position of the PBLH, solid blue lines represent the average profiles of corresponding profiles, and error bars represent the standard deviations.*

6.2.2 PBLH and $PM_{2.5}$ under different aerosol structure scenarios

Absorbing aerosols tend to have a positive feedback with the PBLH, and the aerosol vertical distribution plays a critical role in this process. We investigate the relationship between MPL-derived PBLH and $PM_{2.5}$ for absorbing (daily average SSA $\leq$ 0.85) or weakly absorbing (daily average SSA > 0.9) aerosols for increasing/decreasing aerosol structures during 0900–1900 LT (Figure 6.9). The PBLH-$PM_{2.5}$ relationships can represent the intensity of the API. In general, there are stronger correlations between PBLH and $PM_{2.5}$ for the inverse aerosol structure. This is likely caused by substantial heating in the upper PBL, facilitating the formation of a temperature inversion and further increasing the stability of the PBL. For the decreasing aerosol structure, aerosols may not significantly redistribute adiabatic energy. Hence, the PBLH-$PM_{2.5}$ correlation is relatively weak. Significant PBLH-$PM_{2.5}$ correlations are found for both absorbing and weakly absorbing cases, indicating that scattering aerosols may also play an important role in the API, especially for the inverse aerosol structure.

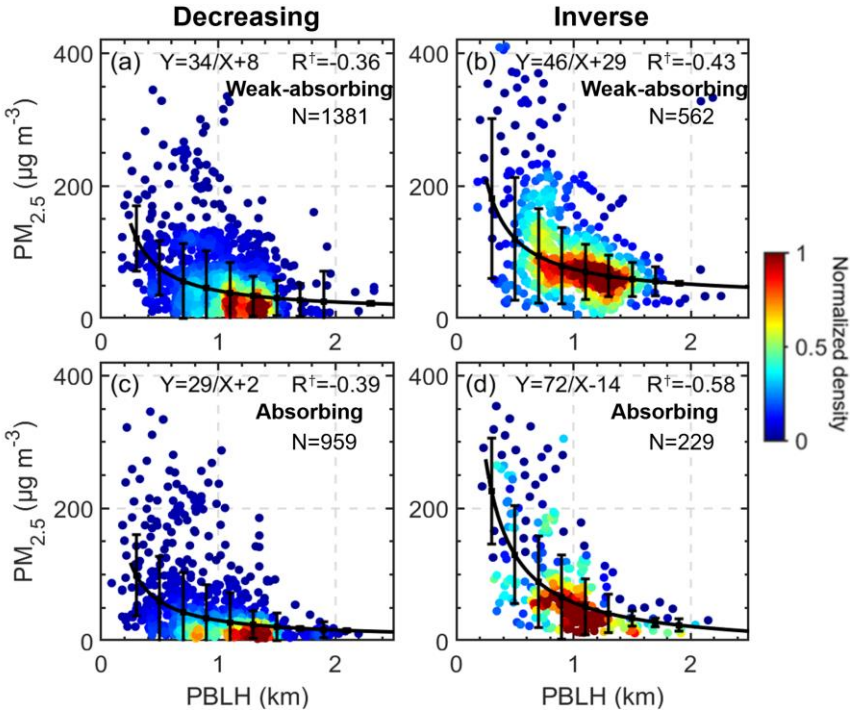


Figure 6.9: *The relationship between MPL-derived PBLH and $PM_{2.5}$ for (a) weakly absorbing and (c) absorbing aerosols for the decreasing aerosol structure. The relationship between MPL-derived PBLH and $PM_{2.5}$ for (b) weakly absorbing and (d) absorbing aerosols for the increasing (i.e., inverse) aerosol structure. Black lines represent the inverse fits, and the whiskers indicate the standard deviations. The fitting functions and number of samples are given in each panel, along with the correlation coefficient ($R^{\dagger}$) for the inverse fit.*

Figure 6.10 presents the averaged diurnal cycles of AEC, PBLH, and $PM_{2.5}$ for different aerosol vertical structures, classified based on the average AEC profiles during noontime. High humidity cases (surface relative humidity > 90%) and strong wind cases (WS > 5m s^{-1}) are excluded. Here, both AEC and PBLH are derived from MPL data. Data are collected on 371 available days, of which 191 days have decreasing aerosol structures, 122 days have well-mixed aerosol structures, and 58 days have inverse aerosol structures. Multiple entangled factors can contribute to the formation of different aerosol structures within the PBL, including synoptic patterns, new particle formation, vertical turbulence, horizontal transport, entrainment rates, to name a few. In general, the inverse structure is characterized by higher aerosol loadings and lower PBLHs, whereas the decreasing structure is characterized by light pollution and a well-developed PBL. In theory, $PM_{2.5}$ should generally decrease with increasing PBLH in the morning due to the dilution effect. This situation is demonstrated clearly for decreasing aerosol structures. However, $PM_{2.5}$ continuously grows during the daytime when an inverse aerosol structure is present, regardless of the PBLH diurnal cycle. Even though many factors control the diurnal variations in aerosols and the PBL, the strong aerosol-stability interaction generates an unfavorable condition for the vertical dissipation of aerosols, so the surface aerosol loading can continuously accumulate due to emissions.

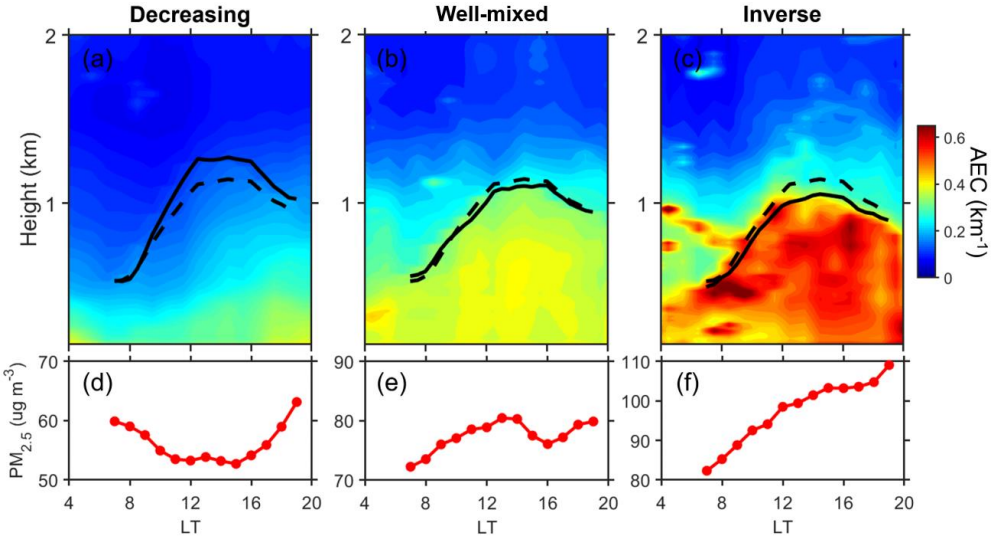


Figure 6.10: *The averaged diurnal variations in AEC for (a) decreasing, (b) well-mixed, and (c) increasing (i.e., inverse) aerosol structures. Solid black lines indicate the averaged diurnal cycles of MPL-derived PBLH under the different aerosol structures. Dashed black lines represent the*

mean MPL-derived PBLH diurnal cycles. (d, e, f) The averaged diurnal variations in surface $PM_{2.5}$ under the different aerosol structures.

The correlations and statistical results concerning the PBLH and $PM_{2.5}$ provide hints about the differences in API for different aerosol structures. However, these results cannot explain the feedback loop and causality. Therefore, we further use the SBDART model with the constraint of ample observations to investigate the vertical profiles of radiative forcing induced by aerosols and their impacts on atmospheric stability.

6.2.3 Aerosol radiative forcing for different aerosol structures

To show vertical profiles of aerosol radiative forcing, the Santa Barbara DISORT Atmospheric Radiative Transfer (SBDART) model (Ricchiazzi et al., 1998) was used to simulate the atmospheric heating rate (dT/dt) induced by aerosols (Liu et al., 2012; Dong et al., 2017). Integrated aerosol inputs include AODs, SSAs (i.e., at 0.44, 0.67, 0.87, and 1.02 μm) retrieved from AERONET measurements, and AEC profiles at 0.5 μm obtained from the MPL. We also use Moderate Resolution Imaging Spectroradiometer surface reflectances as an additional input (https://modis.gsfc.nasa.gov/data/dataprod/mod09.php). We further use heating rates induced by aerosols to estimate the impact of aerosols on buoyancy.

Hence, we can calculate the statistical means of aerosol radiative forcing in the vertical for decreasing, well-mixed, and inverse aerosol structures, derived from the cases presented in Figure 6.10. Figure 6.11 shows that the vertical distributions of the heating rate differ drastically among the different aerosol structures. For the inverse aerosol structure scenario, aerosols cause substantial heating in the upper PBL, facilitating the formation of a temperature inversion and further increasing the stability in the PBL. For the decreasing aerosol structure scenario, the abundance of aerosols at the bottom of PBL heats the lower PBL so can potentially enhance convection in the PBL.

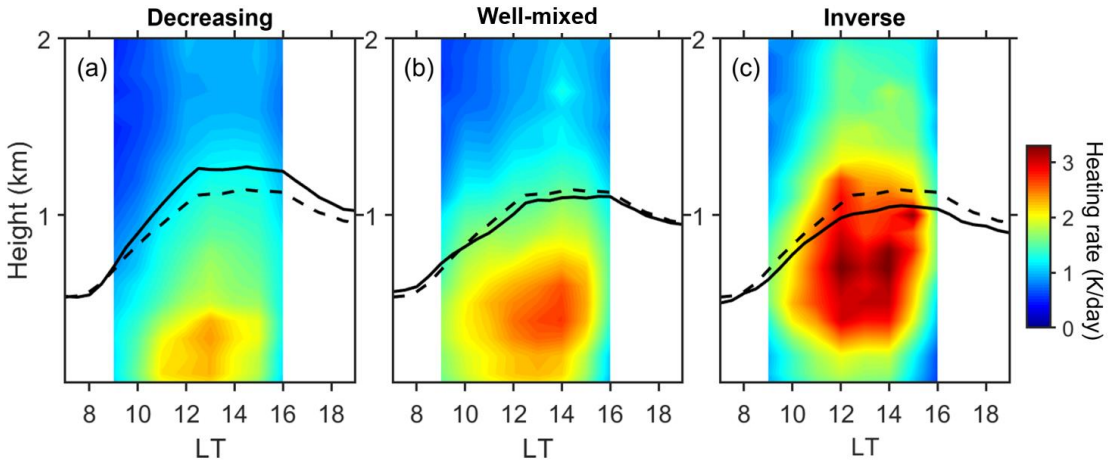


***Figure 6.11**: The averaged diurnal variations in aerosol radiative forcing in the vertical for (a) decreasing, (b) well-mixed, and (c) increasing (i.e., inverse) structures of aerosol loading. Solid black lines indicate the mean diurnal cycles of MPL-derived PBLH under different aerosol structures. Dashed black lines represent the mean MPL-derived PBLH diurnal cycles.*

Moreover, we can estimate the impact of aerosols on buoyancy. Theoretically, the rate of change in buoyancy for a certain layer is expressed as

$$\frac{dB}{dt} = \frac{d}{dt}\left(\frac{T_0 - \Gamma_d \Delta z - T}{T} g\right) = \frac{\left(\frac{dT_0}{dt} - \frac{dT}{dt}\right) T + \frac{dT}{dt}(\Gamma_d - \Gamma)\Delta z}{T^2} g \ , \quad (6.2)$$

where z is the height of the air parcel, and t indicates the time. T represents the temperature of the environment, and Γ_d (Γ) represents the dry adiabatic lapse rate (environmental lapse rate). We primarily focus on the rate of change in buoyancy during the noontime period, when the PBL is well developed, and aerosol radiative forcing is strong. The rate of change in buoyancy (dB/dt) induced by aerosols is largely determined by the aerosol heating rate, which can be produced by the radiative transfer model. Additional inputs include the environmental lapse rate and temperature, obtained from noontime radiosonde soundings.

There are considerable differences in heating rate among the three distinct aerosol structures, which affects the atmospheric buoyancy and stability differently (Figure 6.12). Based on the statistical results during 2016-2018, aerosols generally suppress buoyancy in the lower atmosphere. Such an effect is quite notable for the inverse structure and is insignificant for the decreasing structure with large standard deviations. Absorbing aerosols are not very helpful for

stabilizing the lower atmosphere when a decreasing aerosol structure is present, but they play an important role when an inverse aerosol structure is present. As such, we expect the strongest API to occur for absorbing aerosol cases when an inverse aerosol structure is present, consistent with the results shown in Figure 6.9.

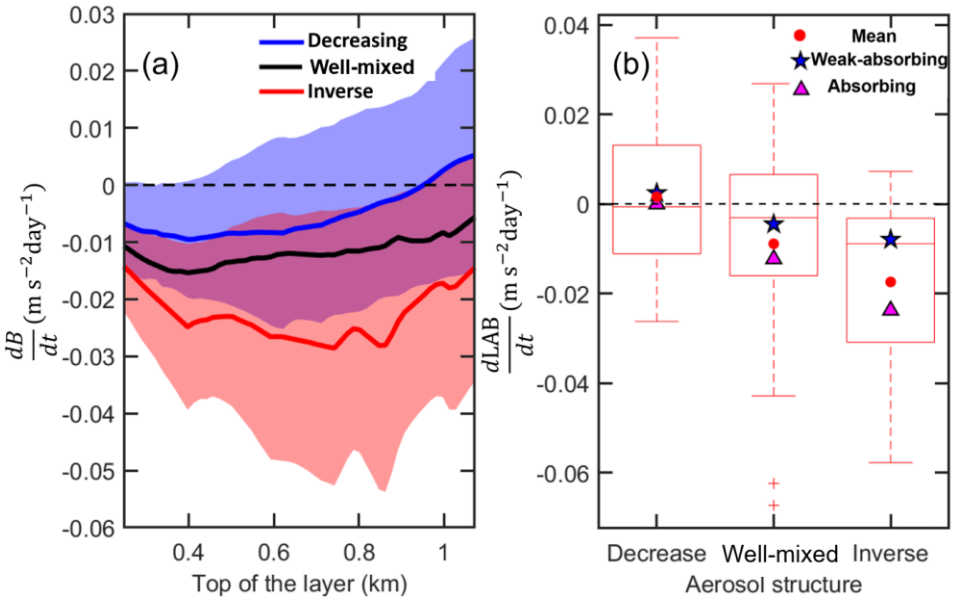


***Figure 6.12**: (a) The rate of change in buoyancy (dB/dt) in a layer of the lowest atmosphere for decreasing (blue), well-mixed (black), and inverse (red) aerosol structures during noontime. The bottom of the layer is the surface, and the rate of change in buoyancy is subjected to the top of the layer. The shaded areas show the standard deviations of the rate of change in buoyancy. (b) Box-and-whisker plots showing 10th, 25th, 50th, 75th, and 90th percentile values of the rate of change in LAB (buoyancy within lowest 1 km) during noontime. Red dots indicate the mean values, and blue stars and pink triangles show the means for weakly absorbing (SSA > 0.9) and absorbing (SSA < 0.85) cases.*

Figure 6.13 shows schematic diagrams of the interactions between aerosols, stability, and the PBL when decreasing/inverse aerosol structures are present. Overall, both decreasing and inverse aerosol structures can cool the surface and suppress sensible heating, thus stabilizing the PBL. In both cases, aerosols have notable stabilizing effects near the surface. When a decreasing aerosol structure is present, abundant aerosols near the surface generate a stronger aerosol heating rate in the lower PBL than in the upper PBL. Such aerosol radiative forcing lowers the potential

temperature gradient ($d\theta/dz$) in the middle and upper PBL and can further strengthen vertical convection in the middle and upper PBL. The above-mentioned two opposite aerosol effects on PBL stability lead to a relatively weak aerosol feedback and a relatively weak aerosol–PBL interaction. When an inverse aerosol structure is present, the significant aerosol heating effect on the upper PBL facilitates the formation of temperature inversion and further increases the stability and suppresses the PBLH. The increase in stability and decrease in PBLH both favor the accumulation of aerosols.

Highly variable aerosol vertical distributions cause large variations in the impact of aerosol on stability and thus exert important and highly variable influences on the aerosol–PBL interactions. Although aerosol stabilizes the PBL for the majority of cases (Figure 6.12), aerosol can also suppress stability in the lower atmosphere, when the aerosol heating effect is much stronger on the near surface than the upper PBL, and further lead to a potential negative feedback loop. The positive feedback loop leads to strong aerosol–PBL interactions, whereas the negative feedback loop leads to weak aerosol–PBL interactions. It explains the paradox of the different correlations between PBLH and surface pollutants since its magnitude, significance, and even sign reportedly vary or even reverse (Quan et al., 2013; Tang et al., 2016; Geiß et al., 2017; Su et al., 2018).

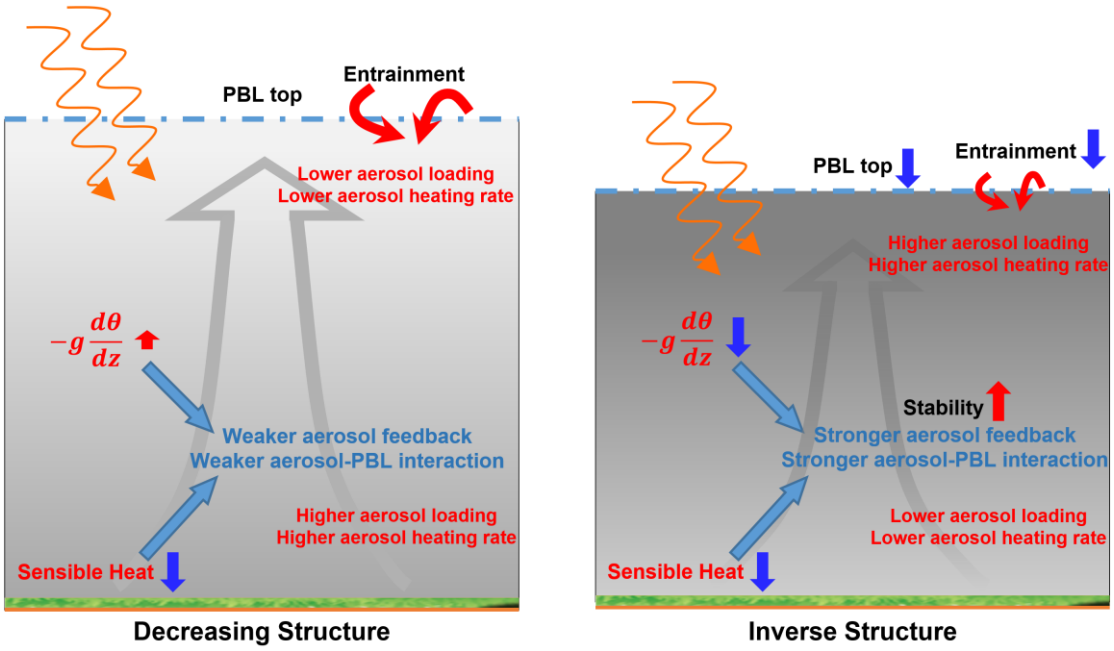


***Figure 6.13**: Schematic diagrams describing aerosol-PBL interactions (API) when decreasing and inverse aerosol structures are present. The blue dash-dotted line indicates the top of the PBL.*

Orange curved arrows indicate solar radiation. The background grey arrow sketches the vertical transport of humidity, aerosols, and heat. The background greyscale indicates the pollution level.

6.3 The association between PBL and air pollution during the COVID-19 lockdown in China

The outbreak of the COVID-19 in early 2020 had tremendous social and economic impacts on China (Zu et al., 2020; Tian et al., 2020; C. Wang et al., 2020). Despite the abrupt decreases in gaseous emissions caused by record-low anthropogenic activity, severe haze pollution occurred in northern China during the COVID-19 lockdown (CLD). The incident serves as a natural testbed for understanding the role of meteorology in air pollution formation, in particular, API in affecting air quality. To quantify the role of the PBL on the severe haze episode that occurred during the CLD, we employed a nationwide PBLH database along with other observations of pollutants and surface meteorological parameters. The abnormally shallow PBL occurring in northern China could account for the unexpected severe air pollution episode, which likely offset the reduction in primary emissions (Su et al., 2020d).

6.3.1 Nationwide changes in air pollution and meteorology

We focus on the CLD, defined as the period between 26 January and 17 February 2020, when anthropogenic emissions dramatically declined over China (Huang et al., 2020; Liu et al., 2020). We compared the surface pollutant concentrations with climatological means for the same period. Hereafter, "climatological means" refer to the same period in the Chinese lunar calendar but from 2016 to 2019, to potentially account for the holiday effect (Zhang et al., 2010; Le et al., 2020). Figure 6.14 presents the ratios of changes in NO_2, SO_2, CO, and $PM_{2.5.}$ On average, surface concentrations of NO_2, SO_2, and CO decreased by 45%, 47%, and 20%, respectively. The reduced primary emissions directly lead to a dramatic decrease in gaseous pollution. Moreover, despite the considerable reduction in emissions during the CLD, $PM_{2.5}$ increased by 19% in the northern part of China, with large regional discrepancies (Figure 6.14d).

Figure 6.14e-h shows simultaneous changes in meteorological variables (i.e., WS, relative humidity (RH), rainfall amount, and PBLH). In general, WS exhibited a 0–10% decrease during the CLD period, except for a few increasing trends. RH did not notably change in southern China, but considerably increased in northern China on the order of a mean of 25% and a range of 10–

70%. A moist environment would facilitate the formation of secondary aerosols (Wang et al., 2016). Rainfall is unlikely a major driving force behind the spatial pattern of air pollutants because both southern and northern China generally experienced more precipitation during the CLD to varying degrees. By contrast, the PBLH showed dramatic changes: significant decreases during the CLD in northern China but increases in central China. Meanwhile, the changes in surface temperature are similar to the changes in PBL to some extent (Figure D.3). Shallow PBL is generally associated with low surface temperature.

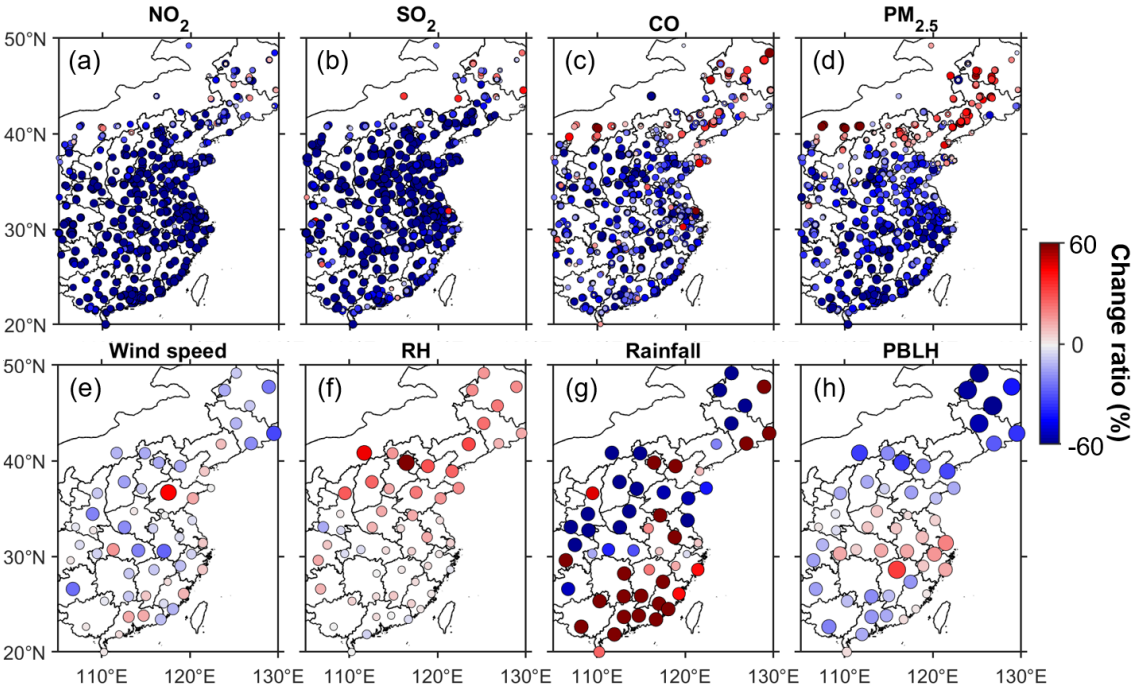


***Figure 6.14**: Percentage differences in (a) NO_2, (b) SO_2, (c) CO, (d) $PM_{2.5}$, (e) wind speed (WS), (f) relative humidity (RH), (g) rainfall amount, and (h) planetary boundary layer height (PBLH) between mean values during the COVID-19 lockdown (26 January to 17 February 2020) and the climatological mean during the same period of the years 2016 to 2019. The PBLH is retrieved during noontime (1100–1500 Beijing Time, or BJT) due to the diurnal transition, while other parameters are averaged during daytime (0800–1900 BJT).*

Regional differences in the ratios of changes in pollutants and meteorology are notable, especially the sharp contrast between northern China (latitudes above 38°N) and central/southern China (latitudes below 38°N). We averaged the $PM_{2.5}$, NO_2, PBLH, WS, and RH anomalies during the CLD within each region, and all anomalies are normalized by climatological means: (values

in 2020 - Climatology)/Climatology. Figure 6.15a reveals that the gaseous pollution diminished considerably during the CLD. As mentioned in previous studies (Huang et al., 2020; Liu et al., 2020), the variation of gaseous pollution is presumably as a result of the lockdown, and the gaseous pollution gradually recovered to normal levels after the CLD. Near the end of the CLD, NO_2 and $PM_{2.5}$ both decreased by more than 50% in different parts of China. In particular, over central and southern China, $PM_{2.5}$ generally decreased by 30%, and NO_2 decreased by 50% during the CLD. However, this was not the case in northern China where $PM_{2.5}$ concentrations were higher than normal. This is linked to meteorological conditions (Figure 6.15b). In northern China, both the humid environment and low wind speeds during the CLD favored the formation and accumulation of aerosols. The PBLH in northern China had the most significant decrease, i.e., a 45% decrease. In central/southern China, changes in these meteorological variables were in similar directions but with lesser magnitudes. Comparing the change ratios in northern China to those in central/southern China, WS was 9.7% lower, RH was 19.8% higher, and the PBLH was 39.1% lower. The PBLH could play a more dominant role in enhancing air pollution during the CLD, discussed next.

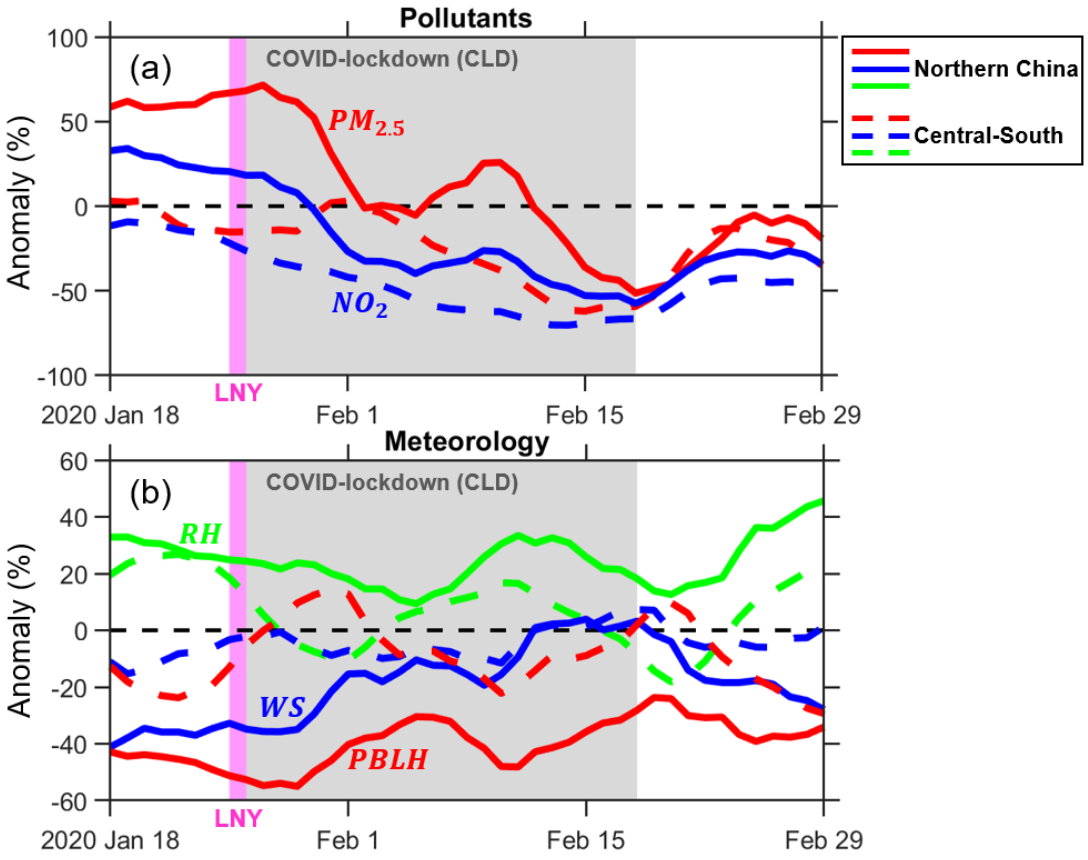


Figure 6.15: *(a) $PM_{2.5}$ (red line) and NO_2 (blue line) anomalies, and (b) planetary boundary layer height (PBLH, red line), wind speed (WS, blue line), and relative humidity (RH, green line) anomalies over China. In (a) and (b), a nine-day smoothing window was applied to all anomalies. Colored solid lines represent results averaged over northern China (latitudes above 38°N), and colored dashed lines represent results averaged over central and southern China (latitudes below 38°N). Zero lines represent the climatology averaged over the same COVID lockdown (CLD) period of the years 2016 to 2019. Pink areas represent the Lunar New Year (LNY), and grey areas represent the CLD period.*

6.3.2 Synergistic changes in the PBLH and $PM_{2.5}$ concentration

Because northern China experienced an unusual increase in $PM_{2.5}$ levels during the CLD, we investigated two regions of interests: Beijing and northeast China. Figure 6.16 shows the general relationships between PBLH and $PM_{2.5}$ during wintertime in Beijing and northeast China. Following Su et al. (2018), an inverse function [$i.e., f(x) = {}^{A}/_{x} + B$] describes the relationship between PBLH and $PM_{2.5}$ (more details are given in the supplement). During wintertime, the relationship was highly nonlinear. The $PM_{2.5}$ concentration increased rapidly with decreasing PBLH for PBLHs lower than 1 km. This dependence weakened for higher PBLHs. Although it is not an absolute value, 1 km appears to be the turning point dictating the $PM_{2.5}$-PBLH relationship.

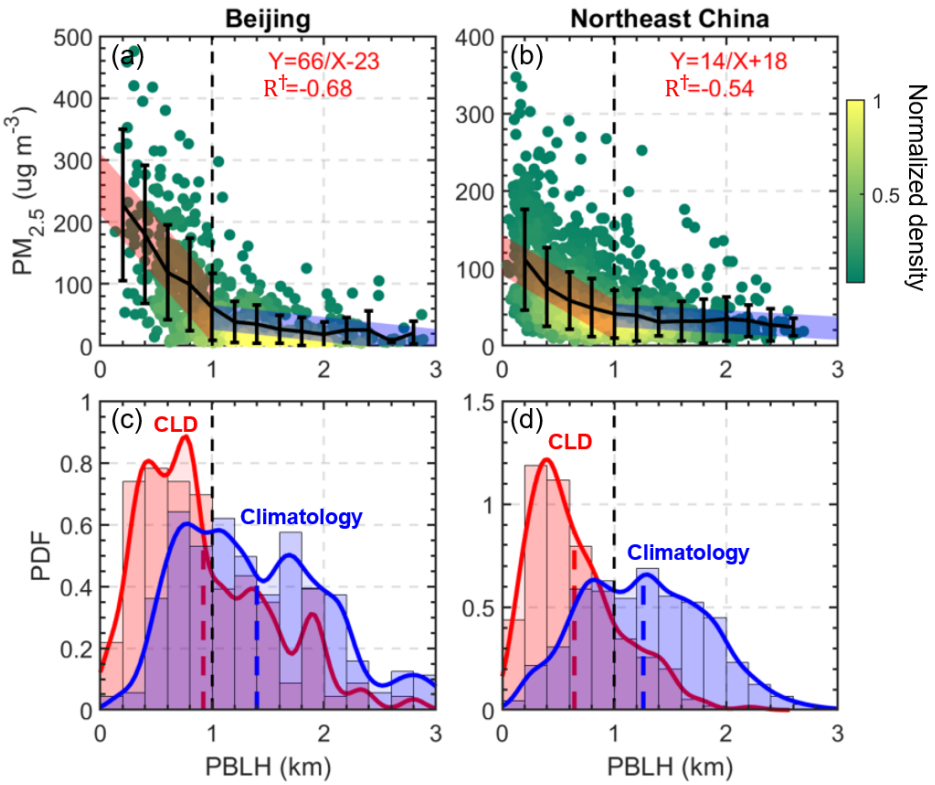


***Figure 6.16**: The relationship between the planetary boundary layer height (PBLH) and $PM_{2.5}$ over (a) Beijing and (b) northeast China during wintertime. The black dots and whiskers represent average values and standard deviations in each bin, respectively. The vertical black dashed line shows the turning point of the PBLH (1 km). The thick red lines and thick blue lines indicate regressions before and after the turning point, respectively. The fitting functions and coefficient correlations of the inverse fittings are given at the top of each panel. Red and blue areas in panels (c) and (d) represent the probability density functions (PDFs) of the PBLH during the COVID lockdown (CLD) and the climatology, respectively. The red and blue dash lines represent the mean PBLH during the CLD and the climatological mean, respectively.*

To gain further insight into the potential role of the PBL in regulating air pollution during the CLD, we compared the probability density functions of the PBLH during the CLD with that during the same period in the lunar calendar during the previous four years. The PBLH was much lower in both Beijing and northeast China during the CLD compared with the mean of the previous four years. The frequency of PBLHs lower than the turning point was significantly higher during the CLD. According to the climatology, the PBLH was mostly above the turning point,

corresponding to a weak interaction scenario. The shallow-PBL cases during the CLD appear to have triggered a strong API, leading to excessively high near-surface aerosol loadings. Due to the positive feedback loop (Bond et al., 2013; Petäjä et al., 2016; Ding et al., 2016; Z. Li et al., 2017b; Su et al., 2020c), the low-PBL effect amplified, contributing to the drastic changes in $PM_{2.5}$ when the PBLH was below the turning point.

Figure 6.17 shows the time series of normalized $PM_{2.5}$ and 1/PBLH in Beijing and northeast China. The normalization is done by subtracting the monthly mean, removing the trend, and subtracting the mean value in the lunar calendar. Highly consistent changes in 1/PBLH and $PM_{2.5}$ are found in the two regions. The $PM_{2.5}$ concentration was higher than normal during the CLD in both Beijing and northeast China, which is associated with the increase in 1/PBLH. Note that there are some discrepancies between the time series of $PM_{2.5}$ and PBLH. The precipitation effectively reduced the pollution level at the end of CLD haze event. Meanwhile, PBLH still maintains relatively low values due to the suppressed surface fluxes. Therefore, a decrease in $PM_{2.5}$ is ahead of the decrease in 1/PBLH during the CLD haze event. In addition, the relationships between meteorological factors and $PM_{2.5}$ are established, following standardized multiple regressions whose coefficients denote the relative importance of the individual factors (Table 6.1). PBLH turns out to have the strongest partial correlation with the daily $PM_{2.5}$ concentration. The other factors also matter.

Table 6.1. *Statistics of the partial correlation coefficients of the relationships between $PM_{2.5}$ and multiple meteorological parameters (PBLH, WS, RH, and rainfall amount) during CLD. Also shown are standardized multiple regression equations of $PM_{2.5}$ onto the meteorological parameters. Asterisks (*) denote the correlations that are statistically significant at the 95% confidence level. PBLH shows significant partial correlations with $PM_{2.5}$ over both Beijing and Northeast China.*

ROIs	**Partial Correlation Coefficients with $PM_{2.5}$ (y)**			
	PBLH (x_1)	**WS (x_2)**	**RH (x_3)**	**Rainfall (x_4)**
Beijing	-0.50*	-0.05	0.18	-0.42*
	Standardized multiple regression: $y = -0.59x_1 - 0.05x_2 + 0.18x_3 - 0.45x_4$			
	-0.44*	-0.41*	0.14	-0.19

Northeast China	Standardized multiple regression: $y = -0.40x_1 - 0.41x_2 + 0.09x_3 - 0.13x_4$

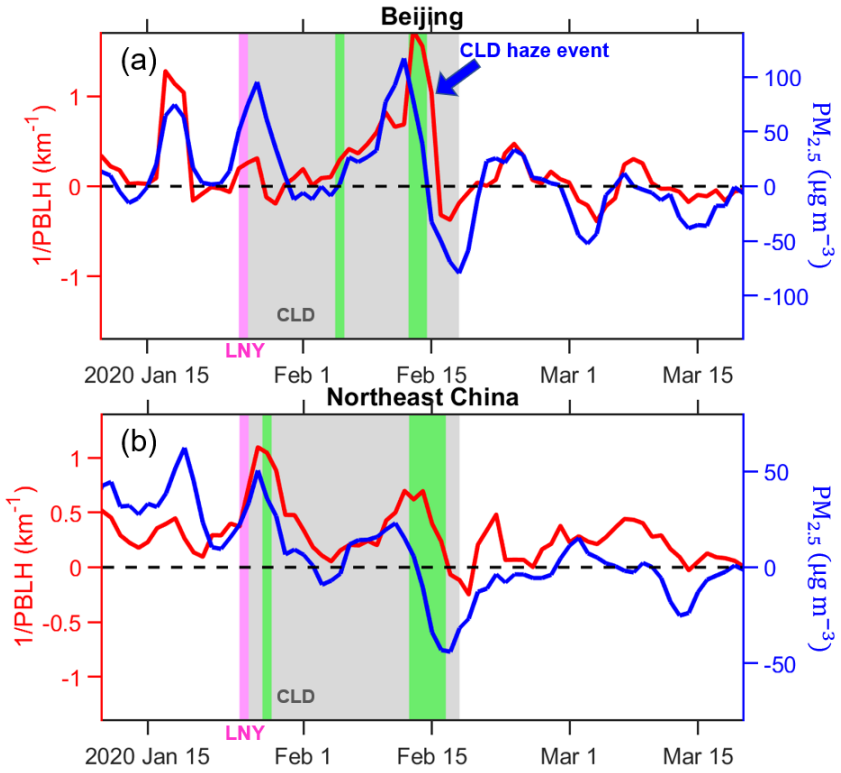


Figure 6.17: *Time series of 1/PBLH (planetary boundary layer height, red lines) and $PM_{2.5}$ (blue lines) over (a) Beijing and (b) northeast China, with a three-day smoothing window. Seasonal cycles, trends, and climatology have been removed in the time series. Pink areas represent the Lunar New Year (LNY), and grey areas represent the COVID lockdown (CLD) period. The green areas indicate periods with daily rainfall above 1 mm. A notable haze event that occurred in Beijing during the CLD is pointed out in (a).*

In the middle of the CLD, Beijing experienced a heavy haze episode. Table D.1 lists the daily air pollution and meteorological parameters during the CLD. During the CLD haze event, there were five consecutive days with noontime PBLHs lower than 0.7 km, which is rare, having only occurred three times during winter from 2013 to 2019 (detailed in Table D.1). For other shallow PBL periods in the previous seven years, the mean $PM_{2.5}$ is 270 μg m^{-3}, with a standard

deviation of 125 μg m^{-3}. A continuously low PBLH can generate an unfavorable environment for the dissipation of pollutants, thus compounding severe air pollution in Beijing.

Further investigated were historical cases with similar meteorological conditions in Beijing, as detailed in the supplement. Figure 6.18 presents the distribution of daytime $PM_{2.5}$ under these similar conditions. Even though WS and RH during the CLD favored the accumulation of near-surface pollutants, these conditions did not necessarily lead to a high surface $PM_{2.5}$ concentration. Due to the unprecedented reduction of various emissions during the CLD, the $PM_{2.5}$ level appears to be the lowest case, compared with the previous level under similar meteorological conditions.

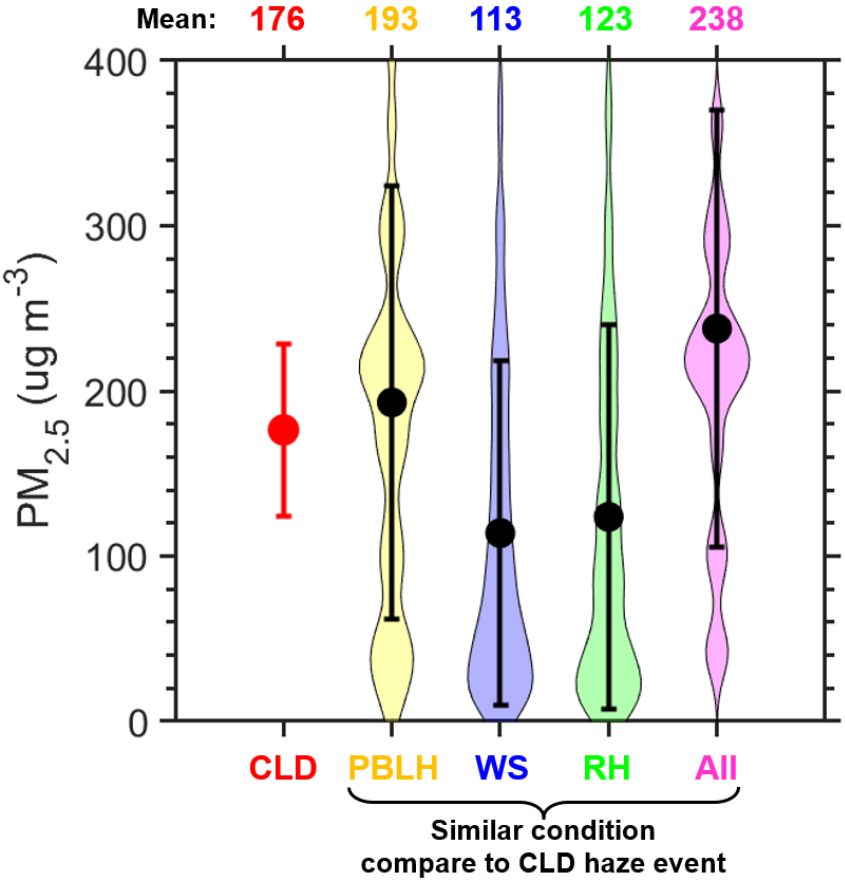


Figure 6.18: *The red dot and whisker represent the average value and standard deviation of daytime $PM_{2.5}$ during the COVID lockdown (CLD) haze event in Beijing. Based on wintertime data from 2013 to 2019, days with a similar planetary boundary layer height (PBLH), wind speed (WS), relative humidity (RH), and all three are selected (less than 20% difference). The distribution of daytime $PM_{2.5}$ is presented under these similar conditions. The black dot and whisker represent the average value and standard deviation, while the width of the color-shaded areas represents*

the smoothed distribution of daytime $PM_{2.5}$. The mean $PM_{2.5}$ values for these categories are given at the top of the figure (unit: $\mu g\ m^{-3}$).

The abnormally shallow PBL during the CLD haze event is attributed to atmospheric dynamics. Using Modern-Era Retrospective analysis for Research and Applications Version 2 reanalysis data (Gelaro et al., 2017), we examine the composite mean meteorological field during the CLD haze event in Figure D.4. During this period, Beijing is located near the southeast flank of a low-pressure system. The southwesterly winds advected warm air masses over the cold surface, which stabilized the PBL. This process is commonly known as the mechanism behind formations of frontal inversion (Zhang et al., 2009). The dynamic processes caused a heating effect on the upper PBL and a cooling effect near the surface during this period (Figure D.5), notably reducing the PBLH. This is also consistent with previous studies showing the frequent occurrence of more stable PBLs under similar synoptic patterns (e.g., Miao et al., 2017; Wu et al., 2017).

Chapter 7 Summary, future work, and perspective

7.1 Summary

To address the scientific questions concerning the API, we developed a set of algorithms for retrieving aerosols (AOD) and PBL from satellite sensors and ground-based lidar. The remote sensing algorithms have been validated against comprehensive ground-based measurements. In general, these methods demonstrate improved performance for retrieving AOD and PBLH. By virtue of the enhanced PBL retrievals, we shall develop the climatology of PBLH and access its relationships with air pollution over a large scale. By making use of the rich dataset of the atmospheric compositions, aerosol loading, vertical distribution, and optical properties under a wide range of conditions, the API has been discussed at the process level. A radiative transfer model is used to access the aerosol radiative forcing and its potential impacts on PBL thermodynamics.

This dissertation contributes to several aspects of remote sensing techniques and understanding of API:

(1) We developed a new method (DTDL) that combines the traditional DT approach, as applied in the NOAA/STAR algorithm for AHI, with a deep learning technique to improve the retrievals of AOD. The DTDL algorithm demonstrates better performance over the study region, with a ~ 30% reduction in random noise. Four independent tests are carried out to train and test in different regions to ensure the applicability of the DTDL algorithm. The DTDL algorithm consistently produces better retrievals than the traditional DT algorithm over untrained regions, with a ~ 20% reduction in uncertainties.

(2) By using long-term datasets of MPL and radiosonde, PBLH retrieved from lidar is investigated systematically under different thermodynamic stability conditions (i.e., SBL, NBL, and CBL). We present a novel algorithm aimed at accurately retrieving the PBLH under Different Thermo-Dynamic Stabilities (DTDS). The DTDS method can assure the temporal continuity of the PBLH in response to the natural variation of atmospheric thermodynamics. The algorithm is tailored to cope with the PBLH changes during its growing, decaying, and maintenance periods, based on a physical model of PBL development. The DTDS method performs better than the widely used gradient and wavelet methods both in terms of the PBLH magnitude and its diurnal variation.

(3) We further developed a lidar-based method to determine the surface-cloud coupling over the land. The surface-cloud coupling identified by the DTDS method shows better consistency with the results derived from radiosondes, with about 10% differences between the lidar-based retrievals and radiosonde results. With knowledge of cloud coupling state, the PBLH can be retrieved under cloudy conditions, which has been a problem with lidar-retrieval of PBLH that is either infeasible or has large uncertainties induced by the occurrence of low clouds.

(4) As an application of the PBLH retrievals, we investigated large-scale relationships between PBLH and $PM_{2.5}$ over China. Although the PBLH-$PM_{2.5}$ correlations are generally nonlinearly negative for the majority of conditions, the magnitude, significance, and even sign vary greatly with location, season, and meteorological conditions. The regional differences in the PBLH-$PM_{2.5}$ relationships are correlated with topography. Related to the more frequent air stagnation and strong local emission, the PBLH-$PM_{2.5}$ correlations are found to be more significant in low-altitude regions. Horizontal transport also shows a significant inverse correlation with PM2.5 concentrations. The mountain breezes and a larger fraction of transported aerosol above the PBL weaken the PBLH-$PM_{2.5}$ correlation over highland areas.

(5) The API is assessed for different aerosol vertical structures, i.e., decreasing, well-mixed, and inversely increasing with height. The aerosol radiative forcing in the vertical for decreasing, well-mixed, and inverse aerosol structures differ drastically, with strong heating in the lower, middle, and upper PBL, respectively. Such a difference in the ARI affects the atmospheric buoyancy and PBL thermodynamics differently for three distinct aerosol vertical structures. For aerosols decreasing with height, $PM_{2.5}$ typically decreases in the morning with relatively large PBLH growth rates. In this situation, absorbing aerosols are not very helpful in stabilizing the lower atmosphere. Meanwhile, the API can be further strengthened by the inverse aerosol structure. It reveals the important roles of the aerosol vertical distribution in API, which should be carefully taken into account in both observational analyses and model simulations.

(6) Using comprehensive ground-based observations of air pollutants, PBLH, and other meteorological variables, we revealed that the abnormally shallow PBL during the COVID-19 lockdown (CLD) was likely a key player in dictating $PM_{2.5}$ in northern China. Compared with the climatology during the same period in the previous four years, the noontime PBLH in central China increased during the CLD, but decreased by more than 40% in northern China. As the most prominent city in northern China, Beijing experienced a persistent low PBLH during the CLD,

leading to a severely polluted episode relative to historical data. This analysis presented here can help explain the high-profile haze event that occurred in Beijing during the CLD.

7.2 Limitations of the current study

Despite the general soundness of our remote sensing methods, we note their limitations in the current study. For satellite retrievals of AOD, we only used a limited area, i.e., a portion of the AHI full disk, to validate the DTDL algorithm. Regarding ground measurements of AOD used in the study, not all underlying surface conditions were considered. Applying the DTDL algorithm to a larger area is thus warranted so that a full understanding of the representativeness and adaptability of this method can be gained. Also, the DNN-constrained SRR is only designed for the DT method and cannot be applied to the other two major AOD retrieval methods (i.e., DB and MAIAC). The DT method has a low accuracy under low NDVI and high albedo conditions. Although this problem is lessened, AOD retrievals from the DTDL algorithm still have a relatively low accuracy under low NDVI and high albedo conditions. Another limitation is that the visualization and interpretation of the deep learning model are still insufficient. Interpretability and causality are the major weaknesses of the DNN. It requires us to have a physical understanding of the parameters and feature rankings in the neural networks. Meanwhile, spatial and temporal independent tests should be carried out to ensure the robustness of deep learning techniques.

For PBLH remote sensing, the accuracy of lidar retrievals can vary considerably under different stability conditions and for different aerosol levels, incurring uncertainties when studying the PBL using lidar observations. This problem is caused by the inherent limitations of lidar techniques and is not completely solved in our method. Specifically, a lidar detects the PBLH from the aerosol layer, and a radiosonde detects the PBLH based on the stratification of potential temperature. When the PBL is well developed, the two approaches tend to agree. However, when the PBL is stable, aerosol stratification is not driven by surface forcing so cannot represent well the vertical structure of potential temperature. Hence, a considerable discrepancy is incurred in PBLH retrievals between lidar- and radiosonde-based approaches. In addition, our PBLH detection method for the MPL has only been applied to two sites (SGP and Beijing). It is still an open question whether the method is applicable to different underlying surfaces and meteorological backgrounds.

For investigating API, this study is largely based on observational data. Observational analyses have limitations in identifying the causality in complex interactions related to aerosols, making it challenging to interpret aerosol-meteorology relations and the confounding impacts induced by aerosols. Moreover, despite aerosol vertical loading having been intensively analyzed, vertical variations in SSA, hygroscopicity, and aerosol type remain unknown, inducing uncertainties in the estimation of aerosol effects. The vertical variations in aerosol properties are expected to critically impact ARI and API. However, this has not been well accounted for in this study. In addition, API is much more complicated for cloud-topped boundary layers. The role of aerosol vertical structure in API is only investigated under cloud-free conditions. Although the processes behind API have been tackled, the joint effects and coupled interactions of ACI and API have not yet been examined thoroughly.

7.3 Future work

More investigations are suggested to address some of the limitations of this study mentioned in Section 7.2. To ensure the adaptability of the method, we plan to apply our PBLH detection algorithm to ample ARM observations at multiple permanent sites and mobile facilities (e.g., SGP, the North Slope of Alaska, the GOAmazon Mobile Facility, the Steamboat Springs Mobile Facility, and the ARM Cloud Aerosol Precipitation Experiment). In addition, selecting appropriate values for empirical parameters in the algorithm is a critical step. Following the sensitivity tests in Section 4.3, other empirical parameters can also be determined using the same principle for different underlying surfaces and meteorological backgrounds. Using ample ARM data, these sensitivity tests will demonstrate the dependence of PBLH biases on different parameters. This will help us understand the importance of each parameter and determine their optimal values.

The long-term goal is to advance our understanding of the integrated interactions between aerosols, the PBL, and clouds under various aerosol and meteorological conditions. This task requires a systematic consideration of API, ARI, and ACI. The impacts of aerosol vertical variability and clouds on API are the primary focus of our study. The vertical variations in aerosol amount and properties will be jointly considered. We argue that this problem also is contingent upon surface-cloud coupling. Surface-cloud coupling not only dictates directly the development of the PBL, but it also affects the transport of near-surface aerosols to the cloud base. Taking advantage of our new methodology for determining surface-cloud coupling, we propose to

examine the impacts of surface-cloud coupling on aerosol-PBL-cloud interactions over land. Moreover, the individual and joint effects of API and ACI under different aerosol conditions and for different surface-cloud coupling regimes merit further investigation. For this purpose, our enhanced datasets of AOD, PBLH, and surface-cloud coupling will be especially useful. This ambitious project for the long run may be beneficial for fundamentally improving our understanding of surface-PBL-cloud coupling and its interaction with aerosols at the process level. There are several tasks worth pursuing:

1) The aerosol radiative effects on PBL thermodynamics by accounting vertical variabilities of aerosol properties.
2) The impacts of API on the vertical distribution of heat fluxes, entrainment, and convection initiation within the PBL;
3) The impacts of API on surface-cloud coupling and cloud regime transitions;
4) The impact of surface-cloud coupling on ACI and API and vice versa.

Appendices

Appendix A: Miscellaneous supplementary information for Chapter 3

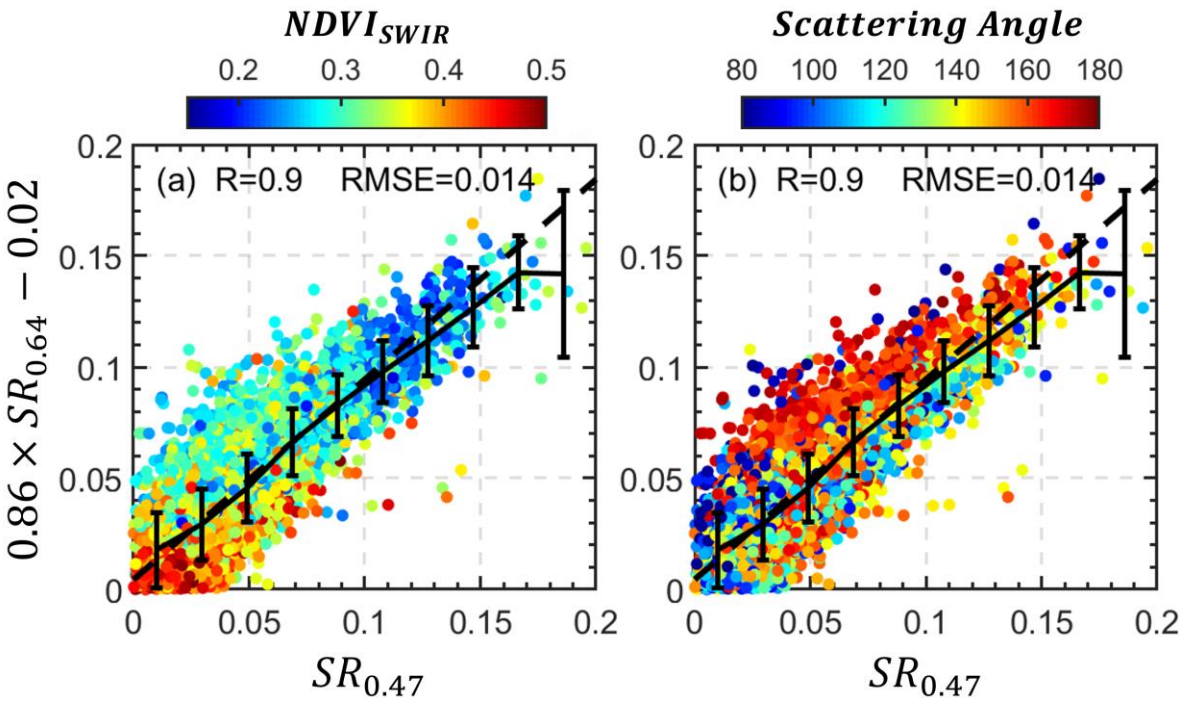


Figure A.1: The comparison between "true" $SR_{0.47}$ and $SR_{0.47}$ predicted by the relationship $0.86 \times SR_{0.64} - 0.02$. In (a), the colored dots indicate the corresponding $NDVI_{SWIR}$. In (b), the colored dots indicate the corresponding scattering angle. The correlation coefficients and RMSE are given at the top of each panel. The linear fitting is reasonably good between $SR_{0.47}$ and $SR_{0.64}$.

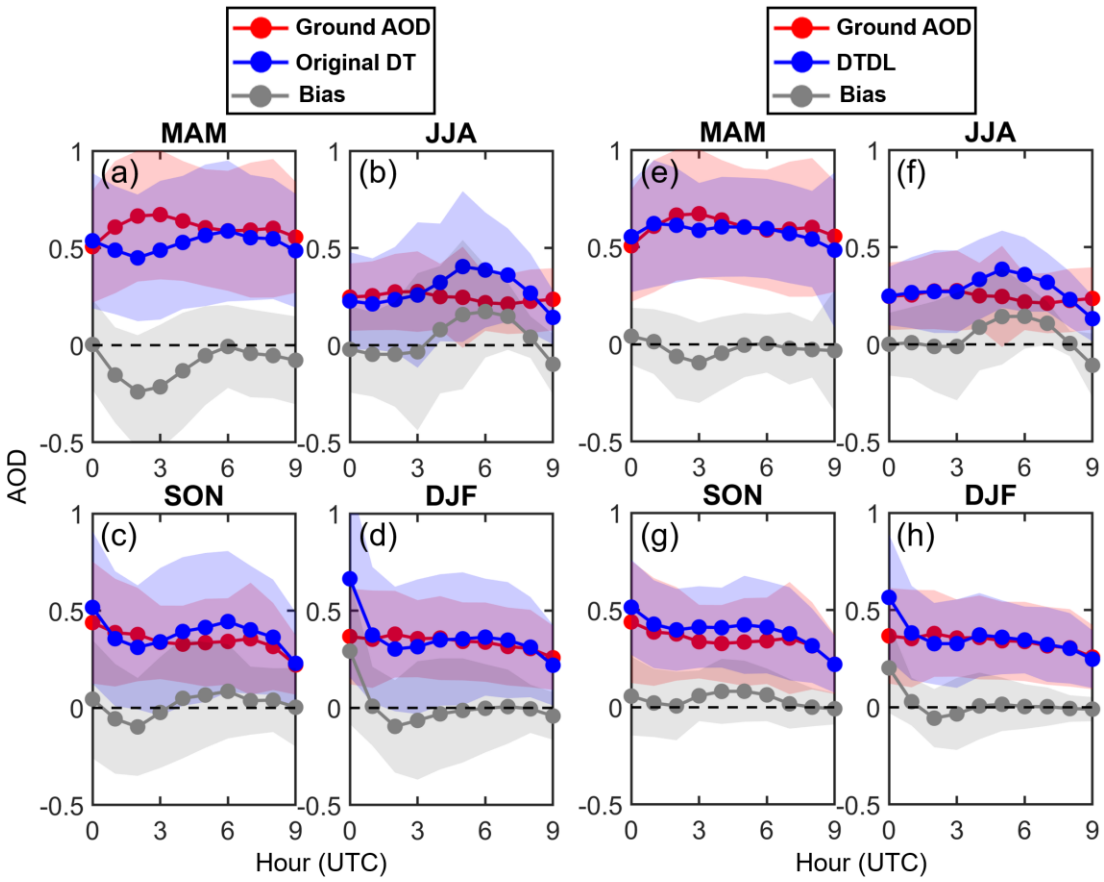


Figure A.2: Diurnal variation of Himawari-8/AHI AOD retrievals derived from (a-d) the original DT and (e-h) DTDL for different seasons (i.e., MAM, JJA, SON, and DJF). Here, all matched pairs of AHI retrievals and ground measurements are used. The diurnal variation derived from ground-based measurements is shown in red, and the bias between AHI and ground AOD is shown in grey. The shaded areas represent the standard deviations.

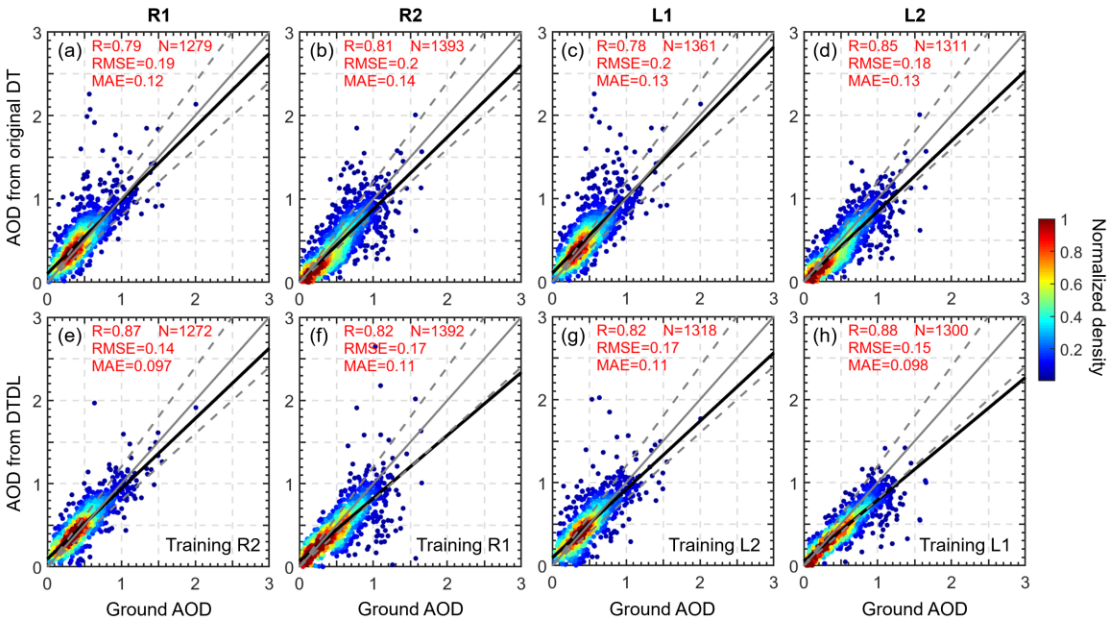


Figure A.3: Comparisons between ground AOD and AOD derived from original DT method over regions of (a) R1, (b) R2, (c) L1, (d) L2. Comparisons between ground AOD and AOD derived from DTDL over regions of (e) R1, (f) R2, (g) L1, (h) L2, while the training datasets are obtained over the regions of (e) R2, (f) R1, (g) L2, (h) L1 respectively. Only high-quality retrievals are used here. This process ensures the independence of the validation dataset. The correlation coefficients, RMSE, sampling numbers, and MAE are given in each panel.

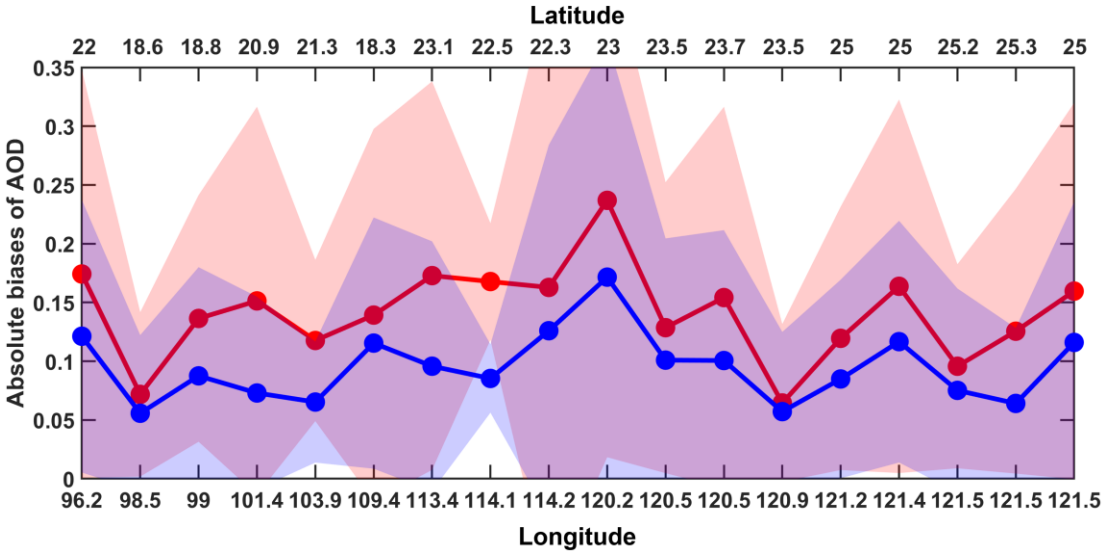


Figure A.4: Absolute biases between AOD derived from ground AOD and AHI at 18 sites. The original DT (red lines) and DTDL (blue lines) are used, respectively. The latitude and longitude for each site are marked in the x-axis. The shaded areas indicate the standard deviations.

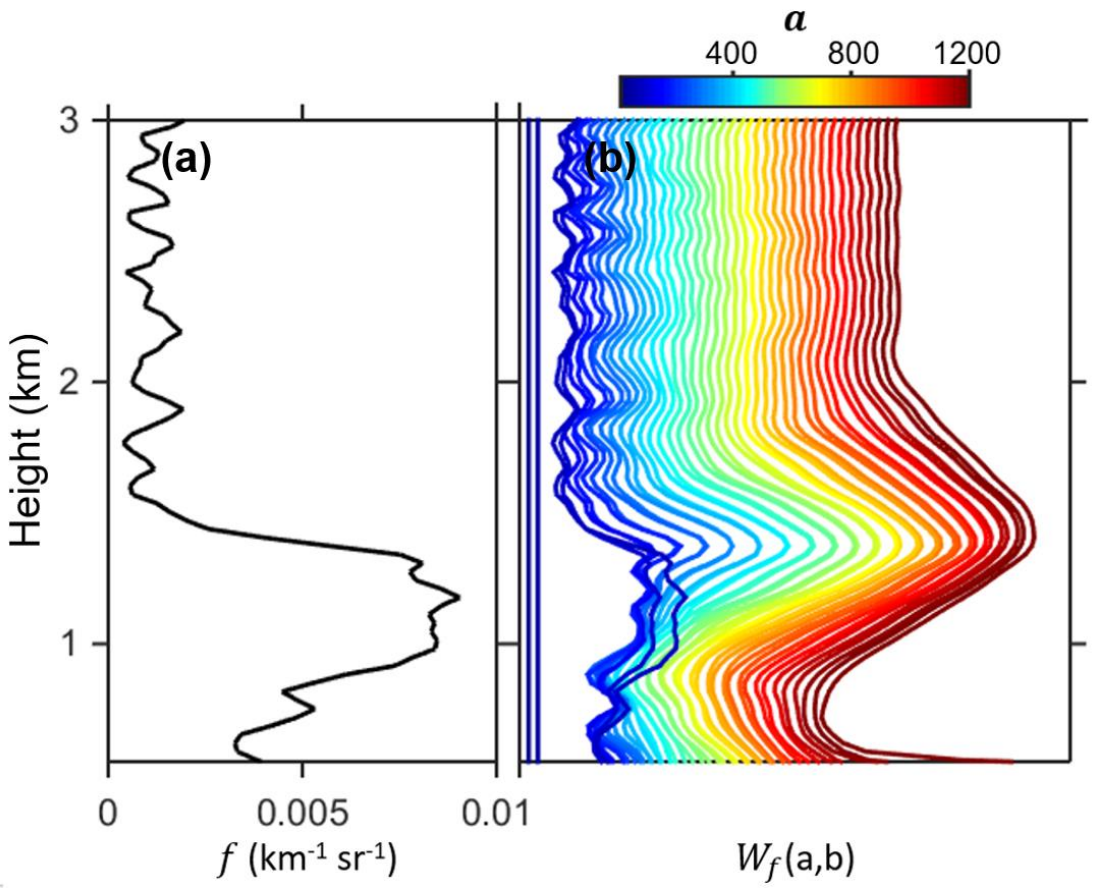


Figure B.1: (a) An example of a lidar backscatter profile and (b) corresponding Haar function $W_f(a,b)$ at various values of the dilation a. Different colors represent different values of the dilation a. The profile of $W_f(a,b)$ become smoother with increased dilation a.

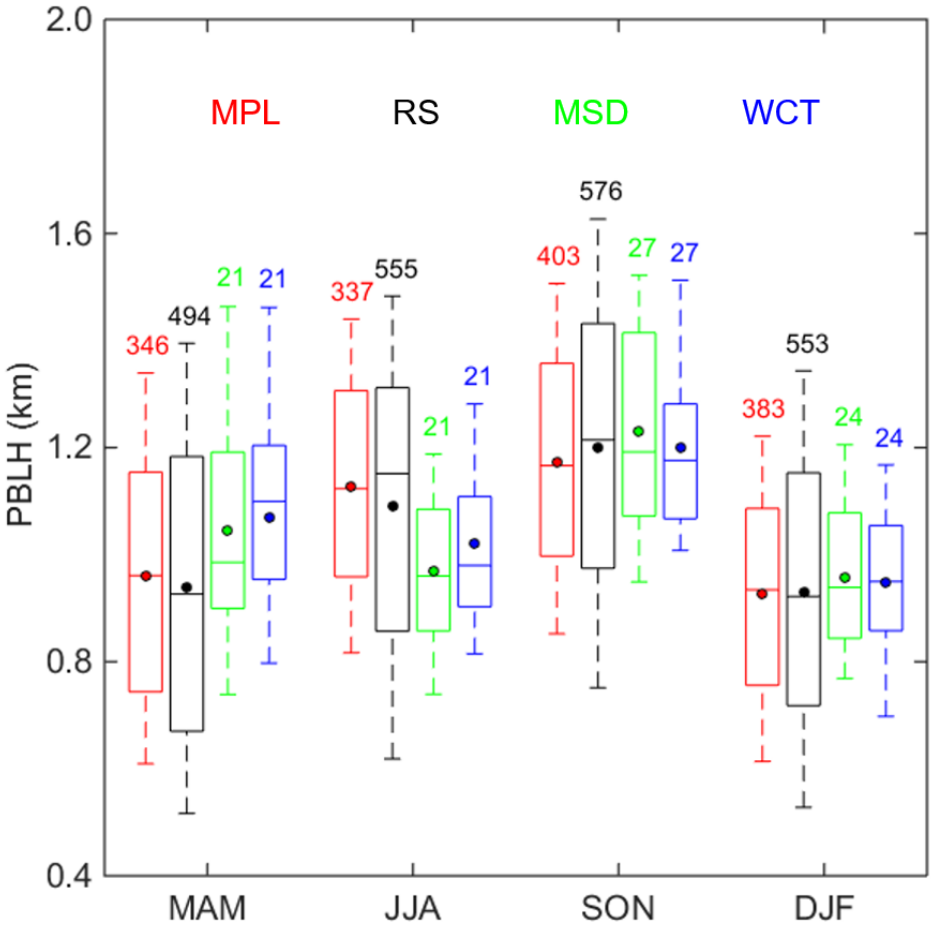


Figure B.2: Seasonal box-and-whisker plots showing 10th, 25th, 50th, 75th, and 90th percentile values of PBLH derived from MPL (red), radiosonde (black), MSD (green) and WCT (blue) for MAM (spring), JJA (summer), SON (autumn) and DJF (winter) respectively. The dots indicate the mean PBLH for each method. The numbers above each box refer to the number of samples.

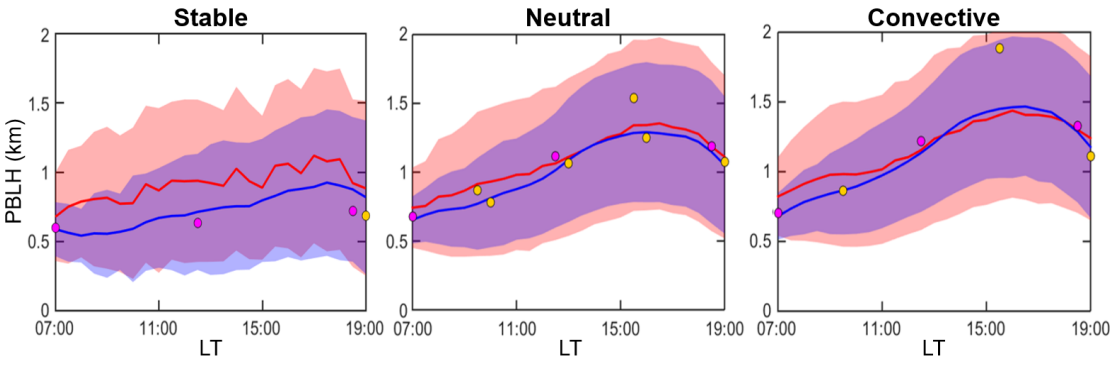


Figure B.3: Diurnal cycles of PBLH under (a) SBL, (b) NBL, and (c) CBL conditions, which are determined based on RS at 1230LT. The blue lines indicate the PBLH retrieved by the DTDS, and the red lines represent the averaged results derived from the gradient and wavelet methods. The shaded areas show the standard deviation of PBLH. Pink dots indicate the PBLH averaged from routine RS data, and yellow dots indicate the PBLH averaged from RS with limited sampling.

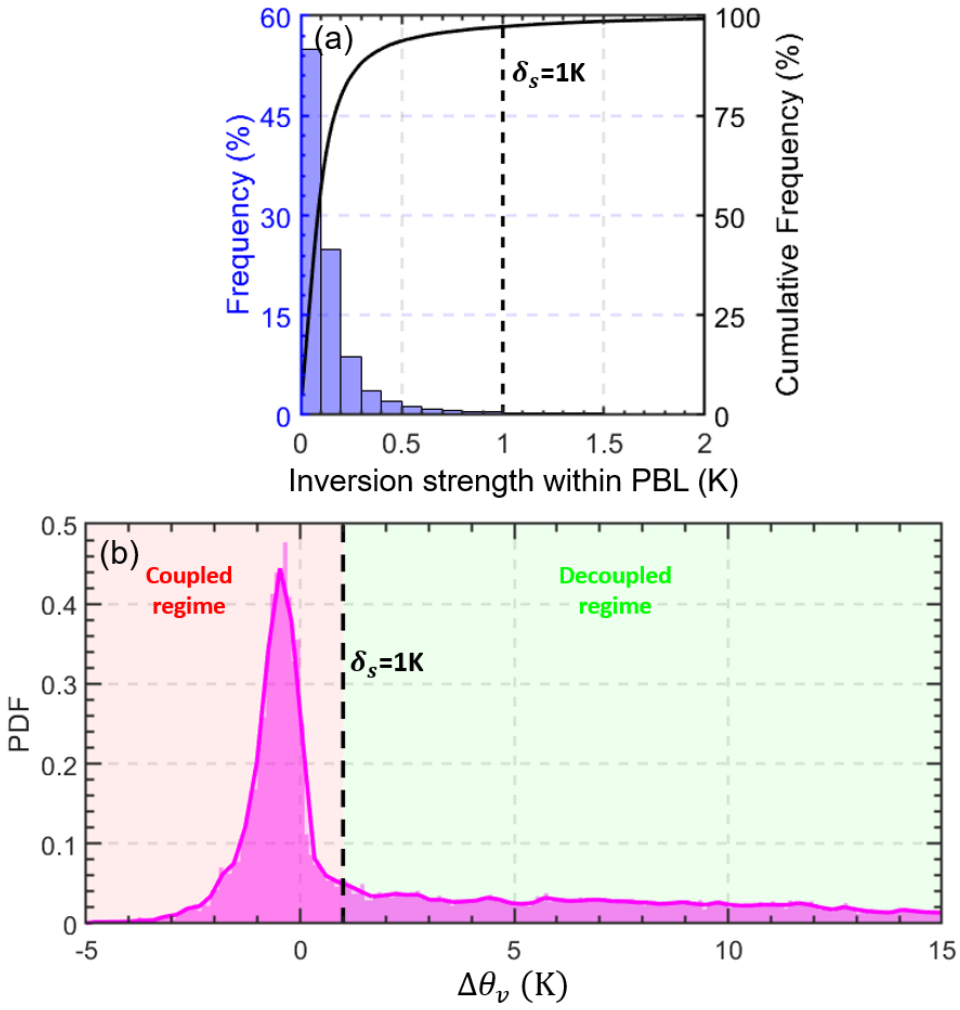


Figure C.1: (a) Blue bars represent the inversion strength of θ_v within the PBL. The inversion strength is derived from the radiosonde during daytime (0800-1900LT). The inversions near surface or across PBL top are excluded. The black solid line represents cumulative frequency. (b) Pink area represents the probability density function (PDF) of the differences in the virtual potential temperature between cloud-base height (CBH) and PBLH ($\Delta\theta_v = \theta_v^{CBH} - \theta_v^{PBLH}$). By using a threshold of δ_s (1 K), coupled and decoupled regimes are classified.

Appendix D. Miscellaneous supplementary information for Chapter 6

Table D.1. $PM_{2.5}$ and meteorology during the CLD haze event and other three shallow PBL periods during 2013–2019.

	Date	$PM_{2.5}$ ($\mu g\ m^{-3}$)	PBLH (m)	WS (m/s)	RH (%)
CLD haze event	20200213	212	425	2.3	84
	20200212	195	648	1.2	60
	20200211	232	437	1.6	60
	20200210	116	429	1.2	47
	20200209	127	507	1.6	58
Period I	20170104	287	237	1.3	78
	20170103	287	332	1.7	72
	20170102	231	133	2.3	67
	20170101	497	301	1.5	88
	20161231	312	545	1.4	70
	20161230	166	647	1.3	67
Period II	20161221	407	378	1.1	78
	20161220	362	241	1.3	89
	20161219	205	558	1.2	56
	20161218	213	512	1.1	61
	20161217	206	320	1.7	58
	20161216	108	522	1.2	47
Period III	20151226	239	245	1.9	79
	20151225	553	238	1.5	97
	20151224	58	296	1.5	74
	20151223	211	539	1.3	73
	20151222	306	168	1.9	86
	20151221	215	420	1.3	68

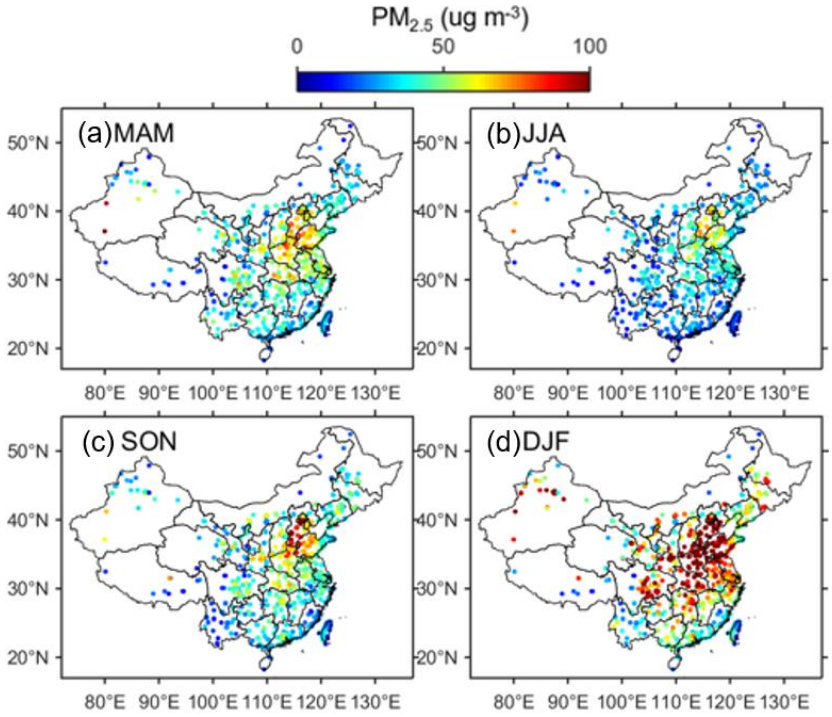


Figure D.1: Spatial distributions of climatological mean of early-afternoon $PM_{2.5}$ concentration (in μg m^{-3}) for (a) MAM, (b) JJA, (c) SON, and (d) DJF during the period 2012–2017.

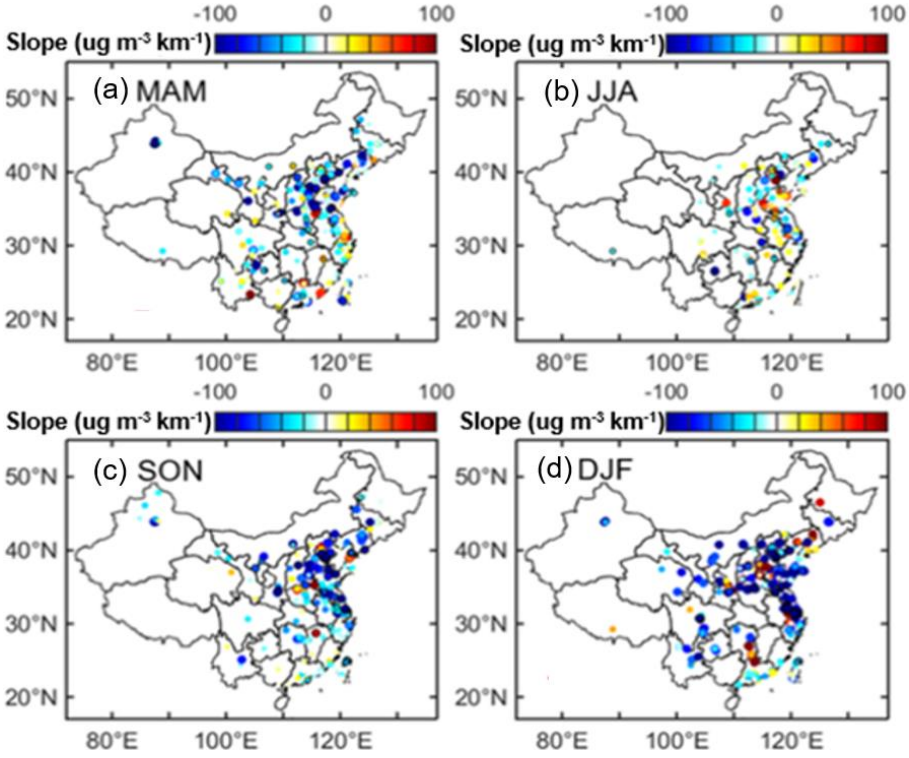


Figure D.2: Spatial distributions of the PBLH-$PM_{2.5}$ relationship slopes derived from the linear fitting for (a) MAM, (b) JJA, (c) SON, and (d) DJF.

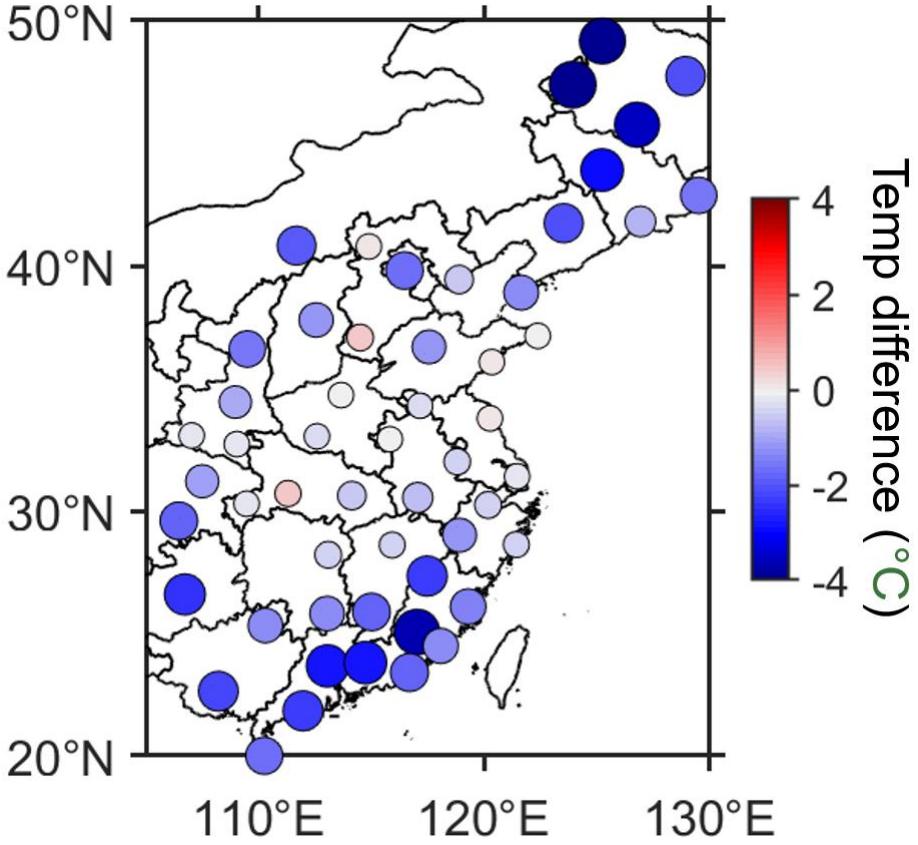


Figure D.3: The differences in daytime temperature between mean values during the COVID-19 lockdown and the climatological mean during the same period of the years 2016 to 2019.

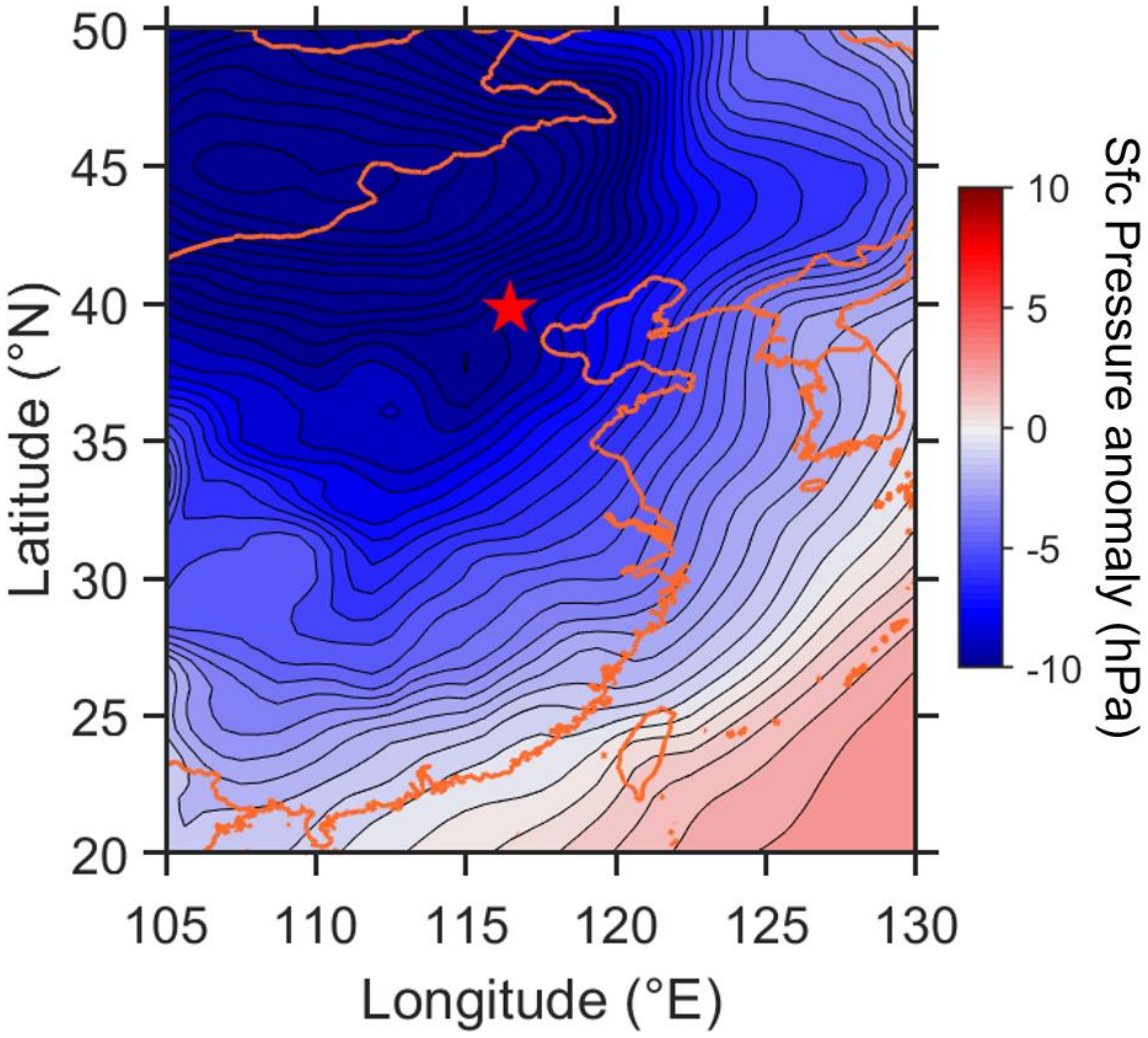


Figure D.4: The surface pressure anomalies (relative to the monthly mean) during the CLD haze event. The location of Beijing is marked as the red star. The orange line indicates the coastline/border. The data are obtained from MERRA-2 reanalysis data.

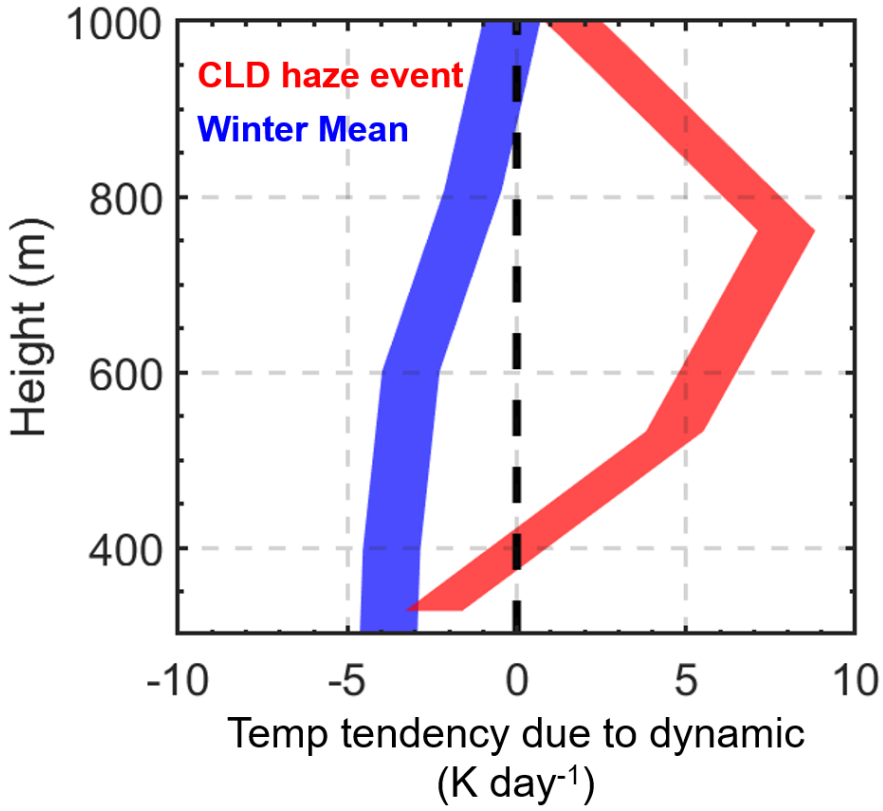


Figure D.5: Profiles of temperature tendency due to the dynamic processes from MERRA-2 reanalysis data. Red line represents the value at 1230 BJT during the CLD haze event, and blue line represents the corresponding mean value during winter. The structure of temperature tendency during the CLD haze event facilitate the temperature inversion, and thus increase the lower-atmosphere stability and reduce the PBLH.

Appendix E. Statement of Originality

This dissertation is largely based upon the following seven papers I lead:

Su, T., Laszlo, I., Li, Z., Wei, J. and Kalluri, S., 2020. Refining aerosol optical depth retrievals over land by constructing the relationship of spectral surface reflectances through deep learning: Application to Himawari-8. *Remote Sensing of Environment*, 251, p.112093.

© 2020 Elsevier.

Su, T., Li, J., Li, C., Xiang, P., Lau, A.K.H., Guo, J., Yang, D. and Miao, Y., 2017. An intercomparison of long-term planetary boundary layer heights retrieved from CALIPSO, ground-based lidar, and radiosonde measurements over Hong Kong. *Journal of Geophysical Research: Atmospheres*, 122(7), pp.3929-3943.

© 2017 American Geophysical Union.

Su, T., Li, Z. and Kahn, R., 2020. 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, p.111519.

© 2020 Elsevier.

Su, T., Zheng, Y. and Li, Z., 2021. Methodology to determine the coupling of continental clouds with surface from lidar and meteorological data. *Atmospheric Chemistry and Physics Discussions*, pp.1-44. In review.

© 2021 European Geosciences Union.

Su, T., Li, Z. and Kahn, R., 2018. Relationships between the planetary boundary layer height and surface pollutants derived from lidar observations over China: regional pattern and influencing factors. *Atmospheric Chemistry and Physics*, 18(21), pp.15921-15935.

© 2018 European Geosciences Union.

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

© 2020 European Geosciences Union.

Su, T., Li, Z., Zheng, Y., Luan, Q. and Guo, J., 2020. Abnormally shallow boundary layer associated with severe air pollution during the COVID-19 lockdown in China. *Geophysical Research Letters*, 47(20), p.e2020GL090041.

© 2020 American Geophysical Union.

Bibliography

Ackerman, A. S., Kirkpatrick, M. P., Stevens, D. E., and Toon, O. B., 2004. The impact of humidity above stratiform clouds on indirect aerosol climate forcing. *Nature*, 432, 1014–1017. https://doi.org/10.1038/nature03174.

Amiridis, V., Melas, D., Balis, D.S., Papayannis, A., Founda, D., Katragkou, E., Giannakaki, E., Mamouri, R.E., Gerasopoulos, E., & Zerefos, C., 2007. Aerosol lidar observations and model calculations of the planetary boundary layer evolution over greece, during the march 2006 total solar eclipse. *Atmospheric Chemistry and Physics, 7*, 6181-6189

Davies, F., Middleton, D.R. and Bozier, K.E., 2007. Urban air pollution modelling and measurements of boundary layer height. *Atmospheric environment*, 41(19), pp.4040-4049.

Dubovik, O. and King, M.D., 2000. A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements. *Journal of Geophysical Research: Atmospheres*, 105(D16), pp.20673-20696.

Atwater, M. A., 1971. The radiation budget for polluted layers of the urban environment. *Journal of Applied Meteorology*, 10(2), 205–214.

Baars, H., Ansmann, A., Engelmann, R., & Althausen, D., (2008. Continuous monitoring of the boundary-layer top with lidar. *Atmospheric Chemistry and Physics, 8*, 7281-7296

Boers, R., Eloranta, E.W., & Coulter, R.L., 1984. Lidar Observations of Mixed Layer Dynamics - Tests of Parameterized Entrainment Models of Mixed Layer Growth-Rate. *Journal of Climate and Applied Meteorology, 23*, 247-266

Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Berntsen, T., DeAngelo, B. J., Flanner, M. G., Ghan, S., Kärcher, B., Koch, D. and Kinne, S., 2013. Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, 118(11), 5380–5552.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V. M., Kondo, Y., Liao, H., Lohmann, U., and Rasch, P., 2013. Clouds and aerosols, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 571–657, *Cambridge Univ. Press*, Cambridge, UK and New York, NY, USA.

Bravo-Aranda, J.A., Moreira, G.D., Navas-Guzman, F., Granados-Munoz, M.J., Guerrero-Rascado, J.L., Pozo-Vazquez, D., Arbizu-Barrena, C., Reyes, F.J.O., Mallet, M., & Arboledas, L.A. (2017). A new methodology for PBL height estimations based on lidar depolarization measurements: analysis and comparison against MWR and WRF model-based results. *Atmospheric Chemistry and Physics, 17*, 6839-6851

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

Bühlmann, P. and Van De Geer, S., 2011. Statistics for high-dimensional data: methods, theory and applications. Springer Science & Business Media.

Burden, F. and Winkler, D., 2008. Bayesian regularization of neural networks. In Artificial Neural Networks, 23–42.

Campbell, J.R., Hlavka, D.L., Welton, E.J., Flynn, C.J., Turner, D.D., Spinhirne, J.D., Scott, V.S., & Hwang, I.H., 2002. 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

Campbell, J.R., Welton, E.J., Spinhirne, J.D., Ji, Q., Tsay, S.C., Piketh, S.J., Barenbrug, M., & Holben, B.N., 2003. Micropulse lidar observations of tropospheric aerosols over northeastern South Africa during the ARREX and SAFARI 2000 dry season experiments. *Journal of Geophysical Research-Atmospheres, 108*

Carslaw, K.S., Lee, L.A., Reddington, C.L., Pringle, K.J., Rap, A., Forster, P.M., Mann, G.W., Spracklen, D.V., Woodhouse, M.T., Regayre, L.A. and Pierce, J.R., 2013. Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature*, 503(7474), p.67.

Carson, D.J., 1973. The development of a dry inversion-capped convectively unstable boundary layer. *Quarterly Journal of the Royal Meteorological Society, 99*, 450-467

Charlson, R.J., Schwartz, S.E., Hales, J.M., Cess, R.D., Coakley, J.J., Hansen, J.E. and Hofmann, D.J., 1992. Climate forcing by anthropogenic aerosols. *Sc*ience, 255(5043), pp.423-430.

Chu, Y., Li, J., Li, C., Tan, W., Su, T. and Li, J., 2019. Seasonal and diurnal variability of planetary boundary layer height in Beijing: Intercomparison between MPL and WRF results. *Atmospheric research*, 227, pp.1-13.

Cireşan, D., Meier, U., Masci, J., Schmidhuber, J., 2012. Multi-column deep neural network for traffic sign classification. Neural Networks 32, 333–338.

Clothiaux, E. E., Ackerman, T. P., Mace, G. G., Moran, K. P., Marchand, R. T., Miller, M. A., & Martner, B. E. (2000). Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites. *Journal of Applied Meteorology*, *39*(5), 645–665. https://doi.org/10.1175/1520-0450(2000)039<0645:ODOCHA>2.0.CO;2

Cohn, S. A. and Angevine, W. M., 2000. Boundary layer height and entrainment zone thickness measured by lidars and wind-profiling radars, *J. Appl. Meteorol.*, 39, 1233–1247.

Compton, J.C., Delgado, R., Berkoff, T.A., & Hoff, R.M., 2013. Determination of Planetary Boundary Layer Height on Short Spatial and Temporal Scales: A Demonstration of the Covariance Wavelet Transform in Ground-Based Wind Profiler and Lidar Measurements. *Journal of Atmospheric and Oceanic Technology, 30*, 1566-1575

Davis, K.J., Gamage, N., Hagelberg, C.R., Kiemle, C., Lenschow, D.H., & Sullivan, P.P., 2000. An objective method for deriving atmospheric structure from airborne lidar observations. *Journal of Atmospheric and Oceanic Technology, 17*, 1455-1468

Deng, L. and Yu, D., 2014. Deep learning: methods and applications. Foundations and Trends® in Signal Processing 7(3–4), 197–387.

Ding, A.J., Huang, X., Nie, W., Sun, J.N., Kerminen, V.M., Petäjä, T., Su, H., Cheng, Y.F., Yang, X.Q., Wang, M.H. and Chi, X.G., 2016. Enhanced haze pollution by black carbon in megacities in China. *Geophysical Research Letters*, 43(6), pp.2873-2879.

Dong, X., Schwantes, A. C., Xi, B., & Wu, P., 2015. Investigation of the marine boundary layer cloud and CCN properties under coupled and decoupled conditions over the Azores. *Journal of Geophysical Research: Atmospheres*, *120*, 6179–6191. https://doi.org/10.1002/2014JD022939

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, Atmos. Chem. Phys., 17, 7997–8009, https://doi.org/10.5194/acp-17- 7997-2017, 2017.

Driedonks, A.G.M., 1982. Models and Observations of the Growth of the Atmospheric Boundary-Layer. *Boundary-Layer Meteorology, 23*, 283-306

Eck, T.F., Holben, B.N., Reid, J.S., Dubovik, O., Smirnov, A., O'neill, N.T., Slutsker, I., Kinne, S., 1999. Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. *J. Geophys. Res. Atmos*. 104(D24), 31,333–31,349.

Fernald, F.G., 1984. Analysis of atmospheric lidar observations: some comments. *Appl. Opt.* 23, 652e653.

Ferrero, L., Castelli, M., Ferrini, B.S., Moscatelli, M., Perrone, M.G., Sangiorgi, G., D'Angelo, L., Rovelli, G., Moroni, B., Scardazza, F. and Močnik, G., 2014. Impact of black carbon aerosol over Italian basin valleys: high-resolution measurements along vertical profiles, radiative forcing and heating rate. *Atmospheric Chemistry and Physics*, 14(18), pp.9641-9664.

Flamant, C., Pelon, J., Flamant, P.H. and Durand, P., 1997. Lidar determination of the entrainment zone thickness at the top of the unstable marine atmospheric boundary layer. *Boundary-Layer Meteorology*, 83(2), pp.247-284.

Flynn, D., Shi, Y., Lim, K., & Riihimaki, L., 2017. Cloud Type Classification (cldtype) Value-Added Product. Ed. by Robert Stafford, ARM Research Facility. DOE/SC-ARM-TR-200.

Gamage, N., & Hagelberg, C., 1993. Detection and Analysis of Microfronts and Associated Coherent Events Using Localized Transforms. *Journal of the Atmospheric Sciences, 50*, 750-756

Geiß, A., Wiegner, M., Bonn, B., Schäfer, K., Forkel, R., Schneidemesser, E.V., Münkel, C., Chan, K.L. and Nothard, R., 2017. Mixing layer height as an indicator for urban air quality?. *Atmospheric Measurement Techniques*, 10(8), pp.2969-2988.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R. and Wargan, K., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). Journal of Climate, 30(14), pp.5419-5454.

Giles, D., Sinyuk, A., Sorokin, M., Schafer, J., Smirnov, A., Slutsker, I., Eck, T., Holben, B., Lewis, J., Campbell, J., Welton, E., Korkin, S., Lyapustin, A., 2019. Advancements in the Aerosol Robotic Network (AERONET) Version 3 database – automated near real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. *Atmos. Meas. Tech.* 12, 169–209. https://doi.org/10.5194/amt-12-169-2019

Guo, J., Su, T., Chen, D., Wang, J., Li, Z., Lv, Y., Guo, X., Liu, H., Cribb, M. and Zhai, P., 2019. Declining summertime local-scale precipitation frequency over China and the United States, 1981–2012: The disparate roles of aerosols. *Geophysical Research Letters*, 46(22), pp.13281-13289.

Guo, J., Su, T., Li, Z., Miao, Y., Li, J., Liu, H., Xu, H., Cribb, M. and Zhai, P., 2017. Declining frequency of summertime local-scale precipitation over eastern China from 1970 to 2010 and its potential link to aerosols. *Geophysical Research Letters*, 44(11), pp.5700-5708.

Gupta, P., Levy, R.C., Mattoo, S., Remer, L.A., Munchak, L.A., 2016. A surface reflectance scheme for retrieving aerosol optical depth over urban surfaces in MODIS Dark Target retrieval algorithm. *Atmos. Meas. Tech*. 9, 3293–3308. https://doi.org/10.5194/amt-9-3293-2016

Hageli, P., Steyn, D.G., & Strawbridge, K.B., 2000. Spatial and temporal variability of mixed-layer depth and entrainment zone thickness. *Boundary-Layer Meteorology, 97*, 47-71

Haykin, S., 2009. Neural Networks and Learning Machines. 3rd edition, Pearson Education, Inc., New Jersey.

Haywood, J. and Boucher, O., 2000. Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: A review. *Reviews of geophysics*, 38(4), pp.513-543.

He, Q.S., Li, C.C., Mao, J.T., Lau, A.K.H. and Li, P.R., 2006. A study on the aerosol extinction-to-backscatter ratio with combination of micro-pulse LIDAR and MODIS over Hong Kong. *Atmospheric Chemistry and Physics*, 6(11), pp.3243-3256.

Holben B. N., T. F. Eck, and co-authors, 1998, AERONET - A federated instrument network and data archive for aerosol characterization, *Rem. Sens. Environ*., 66, 1-16.

Holzworth, G.C., 1964. Estimates of mean maximum mixing depths in the contiguous United States. Monthly Weather Review, 92(5), pp.235-242.

Hooper, W.P., & Eloranta, E.W., 1986. 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*, 990-1001

Hsu, N.C., Jeong, M.J., Bettenhausen, C., Sayer, A.M., Hansell, R., Seftor, C.S., Huang, J., Tsay, S.C., 2013. Enhanced Deep Blue aerosol retrieval algorithm: The second generation. J. Geophys. Res. Atmos. 118(16), 9296–9315.

Hsu, N.C., Tsay, S.C., King, M.D., Herman, J.R., 2006. Deep blue retrievals of Asian aerosol properties during ACE-Asia. IEEE Trans. Geosci. Remote Sens. 44(11), 3180–3195.

Hu, X.M., Nielsen-Gammon, J.W. and Zhang, F., 2010. Evaluation of three planetary boundary layer schemes in the WRF model. Journal of Applied Meteorology and Climatology, 49(9), pp.1831-1844.

Huang, X., Ding, A., Gao, J., Zheng, B., Zhou, D., Qi, X., Tang, R., Wang, J., Ren, C., Nie, W., and Chi, X., 2020. Enhanced secondary pollution offset reduction of primary emissions during COVID-19 lockdown in China. National Science Review, https://doi.org/10.1093/nsr/nwaa137. Zhang, M., Wang, X., Chen, J., Cheng, T.,

Huang, X., Wang, Z. and Ding, A., 2018. Impact of Aerosol-PBL Interaction on Haze Pollution: Multiyear Observational Evidences in North China. *Geophysical Research Letters*, 45(16), pp.8596-8603.

Jackson, J.M., Liu, H., Laszlo, I., Kondragunta, S., Remer, L.A., Huang, J., Huang, H.C., 2013. Suomi-NPP VIIRS aerosol algorithms and data products. *J. Geophys. Res. Atmos*. 118(22), 12–673.

Jeong, M.-J., Li, Z., Chu, D.A., Tsay, S.-C., 2005. Quality and compatibility analyses of global aerosol products derived from the advanced very high resolution radiometer and Moderate Resolution Imaging Spectroradiometer. J. Geophys. Res. Atmos. 110, D10S09. https://doi.org/10.1029/2004JD004648

Jones, C., Bretherton, C., & Leon, D., 2011. Coupled vs. decoupled boundary layers in VOCALS-REx. *Atmospheric Chemistry and Physics*, *11*(14), 7143–7153. https://doi.org/10.5194/acp-11-7143-2011

Jordan, N. S., R. M. Hoff, and J. T. Bacmeister (2010), Validation of Goddard Earth Observing System-version 5 MERRA planetary boundary layer heights using CALIPSO, J. Geophys. Res., 115, D24218, doi:10.1029/2009JD013777.

Karimian, H., Li, Q., Li, C., Jin, L., Fan, J. and Li, Y., 2016. An improved method for monitoring fine particulate matter mass concentrations via satellite remote sensing. Aerosol and Air Quality Research, 16(4), pp.1081-1092.

Kaufman, Y. J., and Coauthors, 1997a. Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer. J. Geophys. Res 102(D14), 17,051–17,067.

Kaufman, Y.J. and Remer, L.A., 1994. Detection of forests using mid-IR reflectance: an application for aerosol studies. IEEE Trans. Geosci. Remote Sens. 32(3), 672–683.

Kaufman, Y.J., Wald, A.E., Remer, L.A., Gao, B.C., Li, R.R., Flynn, L., 1997b. The MODIS 2.1 μm channel-correlation with visible reflectance for use in remote sensing of aerosol. IEEE Trans. Geosci. Remote Sens. 35, 1286–1298.

Klett, J.D., 1985. Lidar inversion with variable backscatter/extinction ratios. *Applied optics*, 24(11), pp.1638-1643.

Kobayashi, H., Suzuki, R., Kobayashi, S., 2007. Reflectance seasonality and its relation to the canopy leaf area index in an eastern Siberian larch forest: multi-satellite data and radiative transfer analyses. *Remote Sens. Environ*. 106(2), 238–252.

Kuang, Y., Zhao, C.S., Tao, J.C., Bian, Y.X. and Ma, N., 2016. Impact of aerosol hygroscopic growth on the direct aerosol radiative effect in summer on North China Plain. *Atmospheric Environment*, 147, pp.224-233.

Laszlo, I. 2018. Remote Sensing of Tropospheric Aerosol Optical Depth From Multispectral Monodirectional Space-Based Observations, in *Comprehensive Remote Sensing*, edited by S. Liang, pp. 137–196, Elsevier, Oxford, https://doi.org/10.1016/B978-0-12-409548-9.10389-6.

Laszlo, I., Ciren, P., Liu, H.Q., Kondragunta, S., Tarpley, J.D., Goldberg, M.D. 2008. Remote sensing of aerosol and radiation from geostationary satellites. *Adv. Space Res.* 41(11), 1882–1893. https://doi.org/10.1016/j.asr.2007.06.047

Laszlo, I., Liu, H., Kim, H.Y., Pinker, R.T., 2020. Shortwave Radiation from ABI on the GOES-R Series. In The GOES-R Series (pp. 179–191). Elsevier.

Le, T., Wang, Y., Liu, L., Yang, J., Yung, Y. L., Li, G., and Seinfeld, J. H., 2020. Unexpected air pollution with marked emission reductions during the COVID-19 outbreak in China. *Science*, https://doi.org/10.1126/science.abb7431.

Levy, R. C., and Coauthors, 2007a. Global aerosol optical properties and application to MODIS aerosol retrieval over land. J. Geophys. Res. Atmos. 112, D13210. https://doi.org/10.1029/2006JD007815

Levy, R., Mattoo, S., Munchak, L.A., Remer, L.A., Sayer, A.M., Patadia, F., Hsu, N.C., 2013. The Collection 6 MODIS aerosol products over land and ocean. Atmos. Meas. Tech. 6, 2989–3034. https://doi.org/10.5194/amt-6-2989-2013

Levy, R.C., Remer, L.A., Mattoo, S., Vermote, E.F., Kaufman, Y.J., 2007b. Second-generation operational algorithm: retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance. J. Geophys. Res. Atmos. 112(D13).

Levy, R.C., Remer, L.A., Kleidman, R.G., Mattoo, S., Ichoku, C., Kahn, R. and Eck, T.F., 2010. Global evaluation of the Collection 5 MODIS dark-target aerosol products over land. *Atmospheric Chemistry and Physics*, 10(21), pp.10399-10420.

Lewis, J. R., Welton, E. J., Molod, A. M., & Joseph, E. (2013). Improved boundary layer depth retrievals from MPLNET. *Journal of Geophysical Research: Atmospheres*, 118(17), 9870-9879.

Li, J., Chu, Y., Li, X. and Dong, Y., 2020. Long-term trends of global maximum atmospheric mixed layer heights derived from radiosonde measurements. Environmental Research Letters, 15(3), p.034054.

Li, Z., et al., 2016. Aerosol and monsoon climate interactions over Asia. *Rev. of Geophys.*, 54, 866–929. https://doi.org/10.1002/2015RG000500.

Li, Z., J. Guo, A. Ding, H. Liao, J. Liu, Y. Sun, T. Wang, H. Xue, H. Zhang, B. Zhu, 2017b. Aerosol and boundary-layer interactions and impact on air quality. National Science Review, nwx117. https://doi.org/10.1093/nsr/nwx117.

Li, Z., Niu, F., Fan, J., Liu, Y., Rosenfeld, D. and Ding, Y., 2011. Long-term impacts of aerosols on the vertical development of clouds and precipitation. *Nature Geoscience*, 4(12), p.888.

Li, Z., Rosenfeld, D., and Fan, J., 2017a. Aerosols and their Impact on Radiation, Clouds, Precipitation and Severe Weather Events, *Oxford Encyclopedia in Environmental Sciences*, 10.1093/acrefore/9780199389414.013.126.

Li, Z., Wang, Y., Guo, J., Zhao, C., Cribb, M.C., Dong, X., Fan, J., Gong, D., Huang, J., Jiang, M. and Jiang, Y., 2019. East Asian Study of Tropospheric Aerosols and their Impact on Regional Clouds, Precipitation, and Climate (EAST-AIRCPC). *Journal of Geophysical Research: Atmospheres*, 124, 13,026-13,054.

Li, Z.Q., Xu, H., Li, K.T., Li, D.H., Xie, Y.S., Li, L., Zhang, Y., Gu, X.F., Zhao, W., Tian, Q.J. and Deng, R.R., 2018. Comprehensive study of optical, physical, chemical, and radiative properties of total columnar atmospheric aerosols over China: An overview of Sun–Sky Radiometer Observation Network (SONET) measurements. *Bulletin of the American Meteorological Society*, 99(4), pp.739-755.

Liang, X., Li, S., Zhang, S., Huang, H. and Chen, S.X., 2016. PM2. 5 data reliability, consistency, and air quality assessment in five Chinese cities. *Journal of Geophysical Research: Atmospheres*, 121(17), pp.10-220.

Lin, C., Li, Y., Lau, A.K., Deng, X., Tim, K.T., Fung, J.C., Li, C., Li, Z., Lu, X., Zhang, X., Yu, Q., 2016. Estimation of long-term population exposure to $PM_{2.5}$ for dense urban areas using 1-km MODIS data. *Remote Sens. Environ.* 179, 13–22.

Liu, F., Page, A., Strode, S. A., Yoshida, Y., Choi, S., Zheng, B., Lamsal, L. N., Li, C., Krotkov, N. A., Eskes, H., and Veefkind, P., 2020. Abrupt decline in tropospheric nitrogen dioxide over China after the outbreak of COVID-19. Science Advances, 6(28), eabc2992

Liu, J., J. Huang, B. Chen, T. Zhou, H. Yan, H. Jin, Z. Huang, and B. Zhang., 2015, Comparisons of PBL heights derived from CALIPSO and ECMWF reanalysis data over China, J. Quant. Spectrosc. Radiat. Transfer, 153, 102– 112.

Liu, J., Zheng, Y., Li, Z., Flynn, C. and Cribb, M., 2012. Seasonal variations of aerosol optical properties, vertical distribution and associated radiative effects in the Yangtze Delta region of China. *Journal of Geophysical Research: Atmospheres*, 117(D16).

Liu, N., Zou, B., Feng, H., Wang, W., Tang, Y., Liang, Y., 2019. Evaluation and comparison of multiangle implementation of the atmospheric correction algorithm, Dark Target, and Deep Blue aerosol products over China. Atmos. Chem. Phys. 19, 8243–8268.

Liu, S. and Liang, X.-Z.: Observed diurnal cycle climatology of planetary boundary layer height, *J. Climate*, 22, 5790–5809, https://doi.org/10.1175/2010JCLI3552.1, 2010.

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., Alsaadi, F.E., 2017. A survey of deep neural network architectures and their applications. *Neurocomputing* 234, 11–26.

Lou, M., J. Guo, L. Wang, H. Xu, D. Chen, Y. Miao, Y. Lv, Y. Li, X. Guo, S. Ma, and J. Li., 2019. On the relationship between aerosol and boundary layer height in summer in China under different thermodynamic conditions. *Earth and Space Science*, 6(5), pp.887-901.

Lyapustin, A., Martonchik, J., Wang, Y., Laszlo, I., Korkin, S., 2011. Multiangle implementation of atmospheric correction (MAIAC): 1. Radiative transfer basis and look-up tables. *J. Geophys. Res. Atmos*. 116(D3).

Lyapustin, A., Wang, Y., Korkin, S., Huang, D., 2018. MODIS Collection 6 MAIAC algorithm. *Atmos. Meas. Tech*. 11(10).

Ma, R., Letu, H., Yang, K., Wang, T., Shi, C., Xu, J., Shi, J., Shi, C., Chen, L., 2020. Estimation of surface shortwave radiation from Himawari-8 satellite data based on a combination of radiative transfer and deep neural network. IEEE Trans. Geosci. Remote Sens. 58(8), 5304–5316.

Martin, C.L., Fitzjarrald, D., Garstang, M., Oliveira, A.P., Greco, S., & Browell, E., 1988. Structure and Growth of the Mixing Layer over the Amazonian Rain-Forest. *Journal of Geophysical Research-Atmospheres, 93*, 1361-1375

Medeiros, B., Hall, A., & Stevens, B., 2005. What controls the mean depth of the PBL? *Journal of Climate, 18*, 3157-3172

Melfi, S.H., Spinhirne, J.D., Chou, S.H. and Palm, S.P., 1985. Lidar observations of vertically organized convection in the planetary boundary layer over the ocean. *Journal of climate and applied meteorology*, 24(8), pp.806-821.

Menon, S., Hansen, J., Nazarenko, L. and Luo, Y., 2002. Climate effects of black carbon aerosols in China and India. Science, 297(5590), pp.2250-2253.

Mhawish, A., Banerjee, T., Sorek-Hamer, M., Lyapustin, A., Broday, D.M., Chatfield, R., 2019. Comparison and evaluation of MODIS multi-angle implementation of atmospheric correction (MAIAC) aerosol product over South Asia. *Remote Sens. Environ*. 224, 12–28.

Miao, Y., Guo, J., Liu, S., Liu, H., Li, Z., Zhang, W., and Zhai, P., 2017. Classification of summertime synoptic patterns in Beijing and their associations with boundary layer structure affecting aerosol pollution, *Atmos. Chem. Phys.*, 17, 3097–3110, https://doi.org/10.5194/acp-17-3097-2017.

Miao, Y., Guo, J., Liu, S., Wei, W., Zhang, G., Lin, Y. and Zhai, P., 2018. The climatology of low-level jet in Beijing and Guangzhou, China. *Journal of Geophysical Research: Atmospheres*, 123(5), pp.2816-2830.

Mielonen, T., Levy, R.C., Aaltonen, V., Komppula, M., De Leeuw, G., Huttunen, J., Lihavainen, H., Kolmonen, P., Lehtinen, K.E.J., Arola, A., 2011. Evaluating the assumptions of surface reflectance and aerosol type selection within the MODIS aerosol retrieval over land: the problem of dust type selection. *Atmos. Meas. Tech*. 4, 201–214.

Montavon, G., Samek, W., Müller, K.-R., 2017. Methods for interpreting and understanding deep neural networks. *Digit. Signal Process*. 73, 1–15.

Nieuwstadt, F.T.M., 1984. The Turbulent Structure of the Stable, Nocturnal Boundary-Layer. *Journal of the Atmospheric Sciences, 41*, 2202-2216

Okamoto, H., et al., 2007, Vertical cloud structure observed from shipborne radar and lidar: Midlatitude case study during the MR01/K02 cruise of the research vessel Mirai, J. Geophys. Res., 112, D08216, doi:10.1029/2006JD007628. McGrath-Spangler, E. L., and A. S. Denning (2012), Estimates of North American summertime planetary boundary layer depths derived from space-borne lidar, *J. Geophys. Res.*, 117, D15101, doi:10.1029/2012JD017615.

Petäjä, T., Järvi, L., Kerminen, V.M., Ding, A.J., Sun, J.N., Nie, W., Kujansuu, J., Virkkula, A., Yang, X., Fu, C.B., Zilitinkevich, S., and M. Kulmala., 2016. Enhanced air pollution via aerosol-boundary layer feedback in China. *Scientific Reports*, 6. https://doi.org/10.1038/srep18998.

Platt, C.M., Young, S.A., Carswell, A.I., Pal, S.R., McCormick, M.P., Winker, D.M., DelGuasta, M., Stefanutti, L., Eberhard, W.L., Hardesty, M. and Flamant, P.H., 1994. The experimental cloud lidar pilot study (ECLIPS) for cloud–radiation research. *Bulletin of the American Meteorological Society*, 75(9), pp.1635-1654.

Poltera, Y., Martucci, G., Coen, M.C., Hervo, M., Emmenegger, L., Henne, S., Brunner, D., & Haefele, A., 2017. 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*, 10051-10070

Poulos, G.S., Blumen, W., Fritts, D.C., Lundquist, J.K., Sun, J., Burns, S.P., Nappo, C., Banta, R., Newsom, R., Cuxart, J., Terradellas, E., Balsley, B., & Jensen, M., 2002. CASES-99: A comprehensive investigation of the stable nocturnal boundary layer. *Bulletin of the American Meteorological Society, 83*, 555-581

Quan, J., Gao, Y., Zhang, Q., Tie, X., Cao, J., Han, S., Meng, J., Chen, P. and Zhao, D.: Evolution of planetary boundary layer under different weather conditions, and its impact on aerosol concentrations. Particuology, 11(1), pp.34-40, 2013.

Ramanathan, V.C.P.J., Crutzen, P.J., Kiehl, J.T. and Rosenfeld, D., 2001. Aerosols, climate, and the hydrological cycle. *Science*, 294(5549), pp.2119-2124.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., 2019. Deep learning and process understanding for data-driven Earth system science. Nature 566(7743), 195–204.

Remer, L. A., and Coauthors., 2001. Angular and seasonal variation of spectral surface reflectance ratios: implications for the remote sensing of aerosol over land. IEEE Trans. *Geosci. Remote Sens.* 39(2), 275–283.

Remer, L.A., Kaufman, Y.J., Tanré, D., Mattoo, S., Chu, D.A., Martins, J.V., Li, R.R., Ichoku, C., Levy, R.C., Kleidman, R.G., Eck, T.F., Vermote, E., Holben, B.N., 2005. The MODIS aerosol algorithm, products, and validation. *J. Atmos. Sci.* 62, 947–973.

Remer, L.A., Mattoo, S., Levy, R.C., Munchak, L.A., 2013. MODIS 3-km aerosol product: algorithm and global perspective. *Atmos. Meas. Tech.* 6, 1829–1844.

Remer, L.A., Tanré, D., Kaufman, Y.J., Levy, R., Mattoo, S., 2006. Algorithm for remote sensing of tropospheric aerosol from MODIS: Collection 5, Product ID MOD04/MYD04.

Ricchiazzi, P., Yang, S., Gautier, C., and Sowle, D., 1998. SBDART: A research and teaching software tool for planeparallel radiative transfer in the Earth's atmosphere, *B. Am. Meteorol. Soc.*, 79, 2101–2114, https://doi.org/10.1175/1520-0477(1998)079<2101:SARATS>2.0.CO;2.

Rodriguez, J.D., Perez, A., Lozano, J.A., 2010. Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE Trans. *Pattern Anal. Mach. Intell.* 32(3), 569–575.

Romps, D. M., 2017. Exact expression for the lifting condensation level. *Journal of the Atmospheric Sciences*, *74*(12), 3891–3900. https://doi.org/10.1175/JAS-D-17-0102.1

Runge, J., Petoukhov, V., Donges, J.F., Hlinka, J., Jajcay, N., Vejmelka, M., Hartman, D., Marwan, N., Paluš, M., Kurths, J., 2015. Identifying causal gateways and mediators in complex spatio-temporal systems. *Nature Commun*. 6(1), 1–10. https://doi.org/10.1038/ncomms9502

Sarle, Warren S., 1994. Neural networks and statistical models, Proc. 19th Annual SAS Users Group Int. Conf., Cary, NC, April 1994, 1538–1550.

Sawyer, V., & Li, Z.Q., 2013. Detection, variations and intercomparison of the planetary boundary layer depth from radiosonde, lidar and infrared spectrometer. *Atmospheric Environment, 79*, 518-528

Schmidhuber, J., 2015. Deep learning in neural networks: an overview. Neural Netw. 61, 85–117.

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

Seide, F., Li, G. Yu, D., 2011. Conversational speech transcription using context-dependent deep neural networks. In: Twelfth annual conference of the international speech communication association.

Sicard, M., Perez, C., Rocadenbosch, F., Baldasano, J.M., & Garcia-Vizcaino, D., 2006. Mixed-layer depth determination in the Barcelona coastal area from regular lidar measurements: Methods, results and limitations. *Boundary-Layer Meteorology, 119*, 135-157

Steyn, D.G., Baldi, M., & Hoff, R.M., 1999. The detection of mixed layer depth and entrainment zone thickness from lidar backscatter profiles. *Journal of Atmospheric and Oceanic Technology, 16*, 953-959

Stull, R.B., 1988. *An Introduction to Boundary Layer Meteorology*. Dordrecht: Springer Netherlands

Su, T., Laszlo, I., Li, Z., Wei, J. and Kalluri, S., 2020b. Refining aerosol optical depth retrievals over land by constructing the relationship of spectral surface reflectances through deep learning: Application to Himawari-8. Remote Sensing of Environment, 251, p.112093.

Su, T., Li, J., Li, C., Lau, A.K.H., Yang, D., Shen, C., 2017c. An intercomparison of AOD-converted PM2.5 concentrations using different approaches for estimating aerosol vertical distribution. *Atmospheric Environment*, 166, pp.531-542.

Su, T., Li, J., Li, C., Xiang, P., Lau, A.K.H., Guo, J., Yang, D. and Miao, Y., 2017a. 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.*, 122(7), 3929- 3943.

Su, T., Li, J., Li, J., Li, C., Chu, Y., Zhao, Y., Guo, J., Yu, Y. and Wang, L., 2017b. The Evolution of Springtime Water Vapor Over Beijing Observed by a High Dynamic Raman Lidar System: Case Studies. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(5), pp.1715-1726.

Su, T., Li, Z. and Kahn, R., 2018. Relationships between the planetary boundary layer height and surface pollutants derived from lidar observations over China: regional pattern and influencing factors. *Atmospheric Chemistry and Physics*, 18(21), pp.15921-15935.

Su, T., Li, Z. and Kahn, R., 2020a. 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, p.111519.

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

Su, T., Li, Z., Zheng, Y., Luan, Q. and Guo, J., 2020d. Abnormally shallow boundary layer associated with severe air pollution during the COVID-19 lockdown in China. Geophysical Research Letters, 47(20), p.e2020GL090041.

Su, T., Zheng, Y. and Li, Z., 2021. Methodology to determine the coupling of continental clouds with surface from lidar and meteorological data. Atmospheric Chemistry and Physics Discussions, pp.1-44. https://doi.org/10.5194/acp-2021-302

Summa, D., Di Girolamo, P., Stelitano, D., & Cacciani, M., 2013. Characterization of the planetary boundary layer height and structure by Raman lidar: comparison of different approaches. *Atmospheric Measurement Techniques, 6*, 3515-3525

Tang, G., Zhang, J., Zhu, X., Song, T., Münkel, C., Hu, B., Schäfer, K., Liu, Z., Zhang, J., Wang, L. and Xin, J., 2016. Mixing layer height and its implications for air pollution over Beijing, China. *Atmospheric Chemistry and Physics*, 16(4), pp.2459-2475.

Tang, G., Zhu, X., Hu, B., Xin, J., Wang, L., Munkel, C., Mao, G., & Wang, Y. (2015). Impact of emission controls on air quality in Beijing during APEC 2014: lidar ceilometer observations. *Atmospheric Chemistry and Physics, 15*, 12667-12680

Tian, H., Liu, Y., Li, Y., Wu, C.H., Chen, B., Kraemer, M.U., Li, B., Cai, J., Xu, B., Yang, Q. and Wang, B., 2020. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science, 368(6491), pp.638-642.

Tie, X., Zhang, Q., He, H., Cao, J., Han, S., Gao, Y., Li, X., and Jia, X. C.: A budget analysis of the formation of haze in Beijing, Atmos. Environ., 100, 25–36, 2015.

Tirelli, C., Curci, G., Manzo, C.,Tuccella, P., Bassani, C., 2015. Effect of the aerosol model assumption on the atmospheric correction over land: case studies with CHRIS/PROBA hyperspectral images over Benelux. Remote Sens. 7, 8391–8415. https://doi.org/10.3390/rs70708391

Tokinaga, H., Tanimoto, Y., Nonaka, M., Taguchi, B., Fukamachi, T., Xie, S.P., Nakamura, H., Watanabe, T., & Yasuda, I., 2006. Atmospheric sounding over the winter Kuroshio Extension: Effect of surface stability on atmospheric boundary layer structure. *Geophysical Research Letters, 33*

Tong, W., Li, L., Zhou, X., Hamilton, A., Zhang, K., 2019. Deep learning $PM_{2.5}$ concentrations with bidirectional LSTM RNN. Air Qual. Atmos. Hlth. 12(4), 411–423.

Vermote, E.F., El Saleous, N., Justice, C.O., Kaufman, Y.J., Privette, J.L., Remer, L., Roger, J.C., Tanré, D., 1997b. Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: background, operational algorithm and validation. J. Geophys. Res. Atmos. 102(D14), 17,131–17,141.

Vermote, E.F., Tanré, D., Deuze, J.L., Herman, M., Morcette, J.J., 1997a. Second simulation of the satellite signal in the solar spectrum, 6S: an overview. IEEE Trans. Geosci. Remote Sens. 35(3), 675–686.

Wallace, J.M. and Hobbs, P.V., 2006. Atmospheric science: an introductory survey (Vol. 92). *Elsevier*.

Walton, C.C., Pichel, W.G., Sapper, J.F., May, D.A., 1998. The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites. J. Geophys. Res. Atmos. 103(C12), 27,999–28,012. https://doi.org/10.1029/98JC02370

Wang, C., Horby, P.W., Hayden, F.G. and Gao, G.F., 2020. A novel coronavirus outbreak of global health concern. *The lancet*, 395(10223), pp.470-473.

Wang, G., Zhang, R., Gomez, M.E., Yang, L., Zamora, M.L., Hu, M., Lin, Y., Peng, J., Guo, S., Meng, J. and Li, J., 2016. Persistent sulfate formation from London Fog to Chinese haze. *Proceedings of the National Academy of Sciences*, 113(48), pp.13630-13635.

Wang, H., Shi, G.Y., Zhang, X.Y., Gong, S.L., Tan, S.C., Chen, B., Che, H.Z. and Li, T., 2015 Mesoscale modelling study of the interactions between aerosols and PBL meteorology during a haze episode in China Jing–Jin–Ji and its near surrounding region–Part 2: Aerosols' radiative feedback effects. Atmospheric Chemistry and Physics, 15(6), pp.3277-3287.

Wang, J., Wang, S., Jiang, J., Ding, A., Zheng, M., Zhao, B., Wong, D.C., Zhou, W., Zheng, G., 2014. Wang, L. and Pleim, J.E.: Impact of aerosol–meteorology interactions on fine particle pollution during China's severe haze episode in January 2013. Environmental Research Letters, 9(9), p.094002.

Wang, X., Dickinson, R. E., Su, L., Zhou, C., and Wang, K., 2017. PM2.5 pollution in China and how it has been exacerbated by terrain and meteorological conditions. *Bulletin of the American Meteorological Society*. https://doi.org/10.1175/BAMS-D-16-0301.1.

Wang, Y., Li, Z., Zhang, Y., Du, W., Zhang, F., Tan, H., Xu, H., Fan, T., Jin, X., Fan, X. and Dong, Z., 2018. Characterization of aerosol hygroscopicity, mixing state, and CCN activity at a suburban site in the central North China Plain. Atmospheric Chemistry and Physics, 18(16), pp.11739-11752.

Wang, Z., Huang, X. and Ding, A., 2018. Dome effect of black carbon and its key influencing factors: a one-dimensional modelling study. Atmospheric Chemistry and Physics, 18(4), pp.2821-2834.

Welton, E. J., et al, 2002. Measurements of aerosol vertical profiles and otical properties during INDOEX 1999 using micropulse lidars, *J. Geophys. Res.*, 107, 8019, doi:10.1029/2000JD000038.

Winker, D. M., Hunt, W. H., and McGill, M. J., 2007. Initial performance assessment of CALIOP, Geophys. Res. Lett., 34, L19803, https://doi.org/10.1029/2007GL030135.

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A., 2009. Overview of the CALIPSO mission and CALIOP data processing algorithms, J. Atmos. Ocean. Tech., 26, 2310–2323, https://doi.org/10.1175/2009JTECHA1281.1.

Wu, P., Ding, Y. and Liu, Y., 2017. Atmospheric circulation and dynamic mechanism for persistent haze events in the Beijing–Tianjin–Hebei region. Advances in Atmospheric Sciences, 34(4), pp.429-440.

Wu, Y., de Graaf, M., Menenti, M., 2016. The sensitivity of AOD retrieval to aerosol type and vertical distribution over land with MODIS data. Remote Sens. 8, 765.

Yan, H.R., Li, Z.Q., Huang, J.P., Cribb, M., & Liu, J.J., 2014. Long-term aerosol-mediated changes in cloud radiative forcing of deep clouds at the top and bottom of the atmosphere over the Southern Great Plains. *Atmospheric Chemistry and Physics, 14*, 7113-7124

Yang, D.W., Li, C.C., Lau, A.K.H., & Li, Y., 2013. Long-term measurement of daytime atmospheric mixing layer height over Hong Kong. *Journal of Geophysical Research-Atmospheres, 118*, 2422-2433

Yang, T., Wang, Z.F., Zhang, W., Gbaguidi, A., Sugimoto, N., Wang, X.Q., Matsui, I., & Sun, Y.L., 2017. 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*, 6215-6225

Zhang, M., Wang, X., Chen, J., Cheng, T., Wang, T., Yang, X., Gong, Y., Geng, F. and Chen, C., 2010. Physical characterization of aerosol particles during the Chinese New Year's firework events. *Atmospheric Environment*, 44(39), pp.5191-5198.

Zhang, Q., Ma, X., Tie, X., Huang, M. and Zhao, C., 2009. Vertical distributions of aerosols under different weather conditions: Analysis of in-situ aircraft measurements in Beijing, China. *Atmospheric Environment*, 43(34), pp.5526-5535.

Zhang, W., J. Guo, Y. Miao, H. Liu, Y. Song, Z. Fang, J. He, M. Lou, Y. Yan, Y. Li, and P. Zhai., 2018. On the summertime planetary boundary layer with different thermodynamic stability in China: A radiosonde perspective. *Journal of Climate*, 31(4), pp. 1451–1465.

Zhang, Y., Gao, Z., Li, D., Li, Y., Zhang, N., Zhao, X. and Chen, J., 2014. On the computation of planetary boundary-layer height using the bulk Richardson number method. *Geoscientific Model Development*, 7(6), pp.2599-2611.

Zheng, Y., & Rosenfeld, D., 2015. Linear relation between convective cloud base height and updrafts and application to satellite retrievals. *Geophysical Research Letters*, 42(15), 6485–6491. https://doi.org/10.1002/2015GL064809

Zhou, M., Zhang, L., Chen, D., Gu, Y., Fu, T.M., Gao, M., Zhao, Y., Lu, X. and Zhao, B., 2019. The impact of aerosol–radiation interactions on the effectiveness of emission control measures. *Environmental Research Letters*, 14(2), p.024002.

Zu, Z.Y., Jiang, M.D., Xu, P.P., Chen, W., Ni, Q.Q., Lu, G.M. and Zhang, L.J., 2020. Coronavirus disease 2019 (COVID-19): a perspective from China. *Radiology*, p.200490.