JOURNAL OF GEOPHYSICAL RESEARCH, VOL. ???, XXXX, DOI:10.1029/,

Improved boundary layer depth retrievals from MPLNET

Jasper R. Lewis,[1,2] Ellsworth J. Welton,[2] Andrea M. Molod,[3,4] Everette Joseph[5]

Jasper R. Lewis, NASA Goddard Space Flight Center, Code 612, Greenbelt, MD 20771, USA jasper.r.lewis@nasa.gov

Andrea M. Molod, Earth System Science Interdisciplinary Center (ESSIC), University of Maryland College Park, College Park, MD, 20742, USA andrea.m.molod@nasa.gov

Ellsworth J. Welton, NASA Goddard Space Flight Center, Code 612, Greenbelt, MD 20771, USA ellsworth.j.welton@nasa.gov

Everette Joseph, Beltsville Center for Climate System Observation, Howard University, 2355 6th Street, NW Washington, DC 20059, USA ejoseph@howard.edu

[1]Joint Center for Earth Systems

Abstract. Continuous lidar observations of the planetary boundary layer (PBL) depth have been made at the Micropulse Lidar Network (MPLNET) site in Greenbelt, MD since April 2001. However, because of issues with the operational PBL depth algorithm, the data is not reliable for determining seasonal and diurnal trends. Therefore, an improved PBL depth algorithm has been developed which uses a combination of the wavelet technique and image processing. The new algorithm is less susceptible to contamination by clouds and residual layers, and in general, produces lower PBL depths. A 2010 comparison shows the operational algorithm overestimates the daily mean PBL depth when compared to the improved algorithm (1.85 and 1.07 km,

Technology, University of Maryland
Baltimore County, Baltimore, MD, USA

[2]NASA Goddard Space Flight Center,
Greenbelt, MD, USA

[3]University of Maryland College Park,
College Park, MD, USA

[4]Global Modeling and Assimilation Office,
Greenbelt, MD, USA

[5]Beltsville Center for Climate System
Observation, Howard University,
Washington, DC, USA

respectively). The improved MPLNET PBL depths are validated using radiosonde comparisons which suggests the algorithm performs well to determine the depth of a fully developed PBL. A comparison with the Goddard Earth Observing System-version 5 (GEOS-5) model suggests that the model may underestimate the maximum daytime PBL depth by ~ 410 m during the spring and summer. The best agreement between MPLNET and GEOS-5 occurred during the fall and they differed the most in the winter.

1. Introduction

The planetary boundary layer (PBL), also referred to as the atmospheric boundary layer (ABL) or simply boundary layer (BL), is the shallow layer of the troposphere nearest to the Earth's surface. The PBL is directly influenced by the surface and responds to surface forcings on the timescale of one hour or less [*Stull*, 1988]. Detailed descriptions of the vertical structure and evolution of the PBL are provided by *Stull* [1988] and *Emeis* [2011], so only a brief description is given here. The PBL (particularly over land surfaces) exhibits a diurnal variation due to the exchange of energy and momentum between the surface and the atmosphere. During the day, convective forces can induce turbulence which results in mixing of pollutants in the atmosphere, commonly referred to as a convective boundary layer (CBL) or mixing layer. At night, as the surface cools, convection ceases and a shallow stable boundary layer (SBL) or nocturnal boundary layer (NBL) develops with a nearly neutral residual layer above. It should be noted that mechanically induced turbulence is also capable of producing a mixed layer, in addition to thermally induced turbulence by convection. The top height (or depth) of the PBL can range from less than one hundred meters to several kilometers. Accurate measurements of the PBL depth with high spatial and temporal coverage are crucial to studies of air quality, weather, and climate.

Several operational methods exist for measuring the PBL depth, including the use of: meteorological masts [*Kaimal and Gaynor*, 1983; *van Ulden and Wieringa*, 1996], radiosondes [*Holzworth*, 1964, 1967], aircraft [*Spangler and Dirks*, 1974], sodar [*Melas*, 1990; *Beyrich*, 1997], wind profilers [*Ecklund et al.*, 1988; *Angevine et al.*, 1994], lidar [*Olsen* *et al.*, 1974; *Lammert and Bösenberg*, 2006], and Global Positioning System (GPS) radio occultation [*von Engeln et al.*, 2005; *Guo et al.*, 2011; *Ao et al.*, 2012]. Each method comes with its own advantages and limitations, so the best option is to use some combination of methods [*Seibert et al.*, 2000]. However, there is no universal definition to determine the PBL depth and the definition may vary depending on the measurement method. Even for a single instrument, there are multiple ways to determine the PBL depth. For example, lidar-derived PBL depths have been obtained from gradients or variance in the backscatter profile, wavelet covariance, and fits to idealized profiles [*Flamant et al.*, 1997; *Hooper and Eloranta*, 1986; *Davis et al.*, 2000; *Steyn et al.*, 1999]. The limitations, capabilities, and biases of several exisiting lidar and ceilometer mixing height retrieval algorithms have been discussed in recent literature [*Haeffelin et al.*, 2011; *Träumner et al.*, 2011; *Brooks and Fowler*, 2012].

Long-term, continuous PBL measurements from lidar are rare, but necessary to ascertain seasonal and diurnal variations in the PBL depth. With multiple continuously-running lidar sites located around the globe and a multiyear record of PBL depths, the MPLNET provides a valuable dataset for improving our understanding of the PBL. However, the current operational PBL algorithm has several problems which had to be addressed in order to make the dataset more useful. Therefore, an improved PBL algorithm, which uses a combination of wavelet covariance and image processing, was developed for this effort. Section 2 describes the methodology used to determine the PBL depth for the operational and improved algorithms. A comparison of PBL depth retrievals at Goddard Space Flight Center (GSFC) for the two algorithms is given in section 3. In section 4, the improved PBL depths from MPLNET are validated using radiosonde-derived PBL depths. The improved PBL depths are then compared to modeled GEOS-5 PBL depths in section 5. Finally, a summary and discussion of future plans are presented in section 6.

2. Methods

The Micropulse Lidar Network (MPLNET) [*Welton et al.*, 2001] is a federated network of micropulse lidar (MPL) systems [*Spinhirne et al.*, 1995], deployed worldwide in support of basic science and the National Aeronautics and Space Adminsitration (NASA) Earth Observing System (EOS) program [*Wielicki et al.*, 1995]. Most MPLNET sites are co-located with Aerosol Robotic Network (AERONET) sunphotometers [*Holben et al.*, 1998]. The operational MPLNET Level 1 data product contains real-time normalized relative backscatter [*Welton and Campbell*, 2002; *Campbell et al.*, 2002] which is used in all higher level products. Scene classification, including aerosol, cloud, and PBL top heights, is available from the Level 1.5b data product (http://mplnet.gsfc.nasa.gov).

The method of retrieving the PBL depth from the operational algorithm is based on the wavelet covariance transform (WCT) described by *Davis et al.* [2000] and *Brooks* [2003]. The convolution of a five-minute averaged scattering ratio profile and the Haar wavelet is used to produce the WCT given by

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

where z_b and z_t are the bottom and top altitudes in the scattering ratio profile, $f(z)$ is the scattering ratio as a function of altitude, z, and the Haar wavelet is defined as

$$h\left(\frac{z-b}{a}\right) = \begin{cases} -1, & \text{for} \quad b-\frac{a}{2} \le z \le b \\ 1, & \text{for} \quad b \le z \le b+\frac{a}{2} \\ 0, & \text{elsewhere} \end{cases} \tag{2}$$

where a and b describe the dilation and translation of the function, respectively. The altitude corresponding to the maximum value of the WCT is recorded as the initial estimate of the PBL top height, z_i. Additionally, a two-fold threshold is used to determine if a secondary z_i' at a lower-altitude peak in the WCT should replace the initial estimate of the PBL top height. In order for z_i' to replace z_i, (i) the value of the WCT at the lower-altitude peak must be within 75% of the maximum WCT value, and (ii) the gradient in the WCT located in-between z_i' and z_i must be large enough to distinguish the lower-altitude peak from uncorrelated noise in the lidar profile.

Three problems have been identified with this product: (1) the presence of low-level clouds can cause difficulty in properly retrieving the PBL depth and frequently produces incorrect, deeper PBL retrievals, (2) residual layers or aerosol layers aloft often mask the growth and collapse of the PBL, and (3) erratic and unphysical fluctuations in the PBL depth retrieved occur frequently. Furthermore, the algorithm must be robust enough to work for any site and meteorological condition within the network. All of these issues had to be addressed in the improved algorithm in order to investigate climatological trends. The improved PBL algorithm has three basic steps: feature identification, layer attribution, and continuity.

2.1. Feature Identification

As done in the operational PBL algorithm, the improved algorithm uses five-minute averages of the scattering ratio profile to calculate the WCT. However, in the improved routine, each lidar profile is screened to remove cases when clouds occur within 5 km of the site elevation and the first derivative of a Gaussian wavelet is used instead of the Haar wavelet because it more closely resembles the gradient in the lidar profile. In this study, cloud screening resulted in the removal of nearly 50% of lidar profiles and showed little seasonal dependence. At locations dominated by cloud cover; however, obtaining reliable PBL depth retrievals could be problematic. The use of a Gaussian wavelet reduces noise in the WCT which improves edge-detection results in subsequent stages of the algorithm.

Features are identified from the WCT using an image detection process similar to the method used to identify gradients in the Structure of the Atmosphere 2D (STRAT-2D) algorithm [*Morille et al.*, 2007; *Haeffelin et al.*, 2011]. The Canny edge-detection algorithm [*Canny*, 1986] is used to identify the upper and lower bounds of features in the WCT image, as shown in Figure 1. The altitude of the maximum WCT value within the extracted feature corresponds to a peak in the gradient in the lidar profile and is recorded as the possible PBL depth. For each time-step, up to three feature altitudes are retained: the altitude of the lowest feature and the altitudes of the two largest peaks in the WCT.

2.2. Layer Attribution

The method used to select an appropriate PBL depth from the retained feature altitudes is based upon the local time of day, altitudes of the extracted features, magnitude of the WCT, variance in the lidar profiles, and the mean altitude of the most recent PBL depth retrievals. Fuzzy logic [*Klir and Yuan*, 1997; *Bianco and Wilczak*, 2002] is used to determine a quality score for each of the retained feature altitudes based on six membership functions (see Appendix A). The feature with the highest quality score is selected as the best estimate of the PBL depth. In most cases, the feature with the lowest altitude is chosen at night and a choice between the higher-altitude features is made between sunrise and sunset.

Because the choice of PBL depth depends partially on the most recent retrieval, the processing direction of the algorithm matters. For example, Figure 2 shows the PBL depth at GSFC for 6 July 2010 when processed in the forward ($0 \rightarrow 24$ UTC) and reverse ($0 \leftarrow 24$ UTC) directions. In the present analysis, retrievals using both processing directions are combined and the lowest altitude for each profile has been selected as the final PBL depth. While this selection may not always result in the correct choice of the PBL depth, it should be noted that in the great majority of cases, both processing directions give the same result. For example, less than 5% of the PBL depth retrievals in 2010 gave different results for the forward and reverse processing directions; and of those, nearly 70% resulted in the selection of the forward-direction PBL depth.

While only one of the extracted feature altitudes is selected as the best estimate of the PBL depth, all feature altitudes are recorded in the final data product for possible future use. The full set of feature altitudes will be useful for studies of the residual layer, identifying smoke and dust layers, and development of a quality assured PBL depth product.

2.3. Continuity

Finally, a continuity scheme is employed to reduce sudden changes in the retrieved PBL depth. Each five-minute averaged PBL depth is compared to a baseline determined by the nearest four (two preceeding and two succeeding) PBL depth retrievals. If the PBL depth for the five-minute average exceeds the average of the other 20-minutes by more than 150 meters, then the PBL depth is set equal to the baseline PBL depth. The process is repeated for the entire day until no further changes can be made.

3. Comparison of PBL depth retrievals

A visual comparison of the operational and improved PBL depth retrievals is provided in Figure 3. A cross-section of the normalized relative backscatter for 5 July 2010 at GSFC is shown with the operational PBL depth represented by black triangles and the improved PBL depth represented by red squares.

At night, the operational algorithm reports the residual layer (~ 2 km) while the improved algorithm generally gives a much lower altitude. However, the improved PBL retrieval should not be interpreted as the true depth of the NBL. Due to instrument limitations in the near-field caused by afterpulsing [*Campbell et al.*, 2002], the MPL has a minimum detectable gradient altitude of approximately 500 m, but the NBL can collapse to altitudes less than 100 m. It is worth mentioning that newer model MPLs do not exhibit the same near-field behavior which will reduce the range cutoff to ~ 200 m in the future. The PBL growth can be seen from sunrise until it stabilizes at approximately 15 UTC. The operational PBL retrieval detects the residual layer at 12 UTC, while the improved algorithm continues to follow the growing PBL. The growth and collapse of the PBL are the most difficult to detect because the gradient at the top of the residual layer can be much larger than at the true PBL top height. From 18 UTC until the end of the day, the improved algorithm stays at the top of the PBL while the operational PBL retrieval fluctuates erratically between 2 km and below 1 km because the two-fold threshold described in section 2 was exceeded.

The monthly means of the daily maximum PBL depth, annual diurnal cycles, and daily mean probability distributions for the two algorithms are compared in Figure 4 for the year 2010 at GSFC. The monthly means from the improved algorithm show that the daily maximum PBL depth at GSFC is highest in the spring/summer and lowest during winter. However, the operational retrieval shows only a weak trend with significant month-to-month oscillation and has higher PBL depths due to the influence of residual aerosol layers and cloud contamination. The diurnal cycles show the largest differences between the improved and operational algorithms occur at night when the the improved PBL retrieval is set to the altitude of the lowest detected feature. The growth of the PBL can be clearly seen in the improved PBL retrieval starting after sunrise, but it is largely hidden by the residual layer in the operational retrieval, resulting in a physically unrealistic reduction in PBL depth after sunrise with a minimum at 1000–1100 local time. From the probability distrubutions, we see that the operational PBL retrieval not only has a larger mean PBL depth (operational: 1.85 km, improved: 1.07 km) but also a broader distribution (operational: $\sigma = 0.58$ km, improved: $\sigma = 0.36$ km). It should be noted that the daily mean PBL depth derived from MPLNET will have a high bias due to instrument limitations that prevent measurements below 500 m.

Seasonal comparisons of the mean diurnal cycles and daily mean probability distributions for 2010 are shown in Figures 5 and 6, respectively. With the exception of the spring diurnal cycle, the improved PBL retrieval is less than the operational retrieval in all cases. This exception is attributed to a high occurrence of cases when the two-fold threshold was exceeded during the spring, producing spurious low PBL depths in the operational retrieval similar to those seen in Figure 3.

The growth of the PBL is visible during all seasons in the improved algorithm; however, it is only seen in part during the spring and summer in the operational retrievals and is completely hidden by residual layers in the fall and winter. From Figure 6 we note that both algorithms show the largest daily mean PBL depths occurring during the summer and the lowest occurring in the winter. The winter probability distribution is very broad for the operational PBL retrieval ($\sigma = 0.60$ km). Because there is less energy available for convection, the distribution is expected to be narrower during winter as seen in the improved PBL retrieval ($\sigma = 0.27$ km). From this point forward, all lidar-derived PBL depths will be calculated using the improved PBL retrieval.

4. Validation of the Improved Algorithm

Estimates of the PBL depth can be derived from radiosondes, launched twice-daily at 0000 and 1200 UTC. However, these standard times occur in the early morning and evening in the eastern United States, which are not adequate for observing the diurnal variation of the PBL or maximum daytime PBL depth [*Liu and Liang*, 2010; *Seidel et al.*, 2012; *McGrath-Spangler and Denning*, 2012]. Furthermore, at these times, the PBL has not fully developed (early morning) or has started to collapse (evening) which the MPL is less likely to detect due to instrument limitations. Therefore, attempts to validate the improved PBL algorithm are limited to periods when radiosonde measurements can be made at non-standard times.

One such opportunity occured when radiosondes were launched from the Howard University Beltsville Center for Climate System Observation as part of the July 2011 DISCOVER-AQ field campaign (http://www.nasa.gov/discover-aq). The Beltsville Center for Climate System Observation (39.05°N, 76.88°W, 52-m site elevation) is located 7 km from the GSFC MPLNET site (38.99°N, 76.84°W, 50-m site elevation). The MPLNET PBL depths were averaged to 20-minute temporal resolution centered around the time of the radiosonde launch for this comparison. Lidar-derived retrievals of the PBL depth were possible during 23 of the 25 radiosonde launches which took place between 1357 and 2134 UTC. Radiosonde data are originally sampled at 2-second intervals and interpolated to 1-second intervals, which results in a nominal vertical resolution of 5-m compared to the 75-m vertical resolution of the lidar. The radiosonde-derived PBL depths were determined using the parcel method [*Holzworth*, 1964, 1967].

Figure 7 shows a cross-section of the normalized relative backscatter at GSFC, the potential temperature profiles from the radiosonde launches at Beltsville, and the PBL depth retrievals from both sources for 1–2 July 2011. In Figure 8, the correlation between the lidar and radiosonde measurements is shown for the entire field campaign. The MPLNET algorithm underestimated the PBL depth with a mean difference of 119-m for the 23 observations. The lidar-derived PBL depths compare well with the radiosonde measurements, suggesting the algorithm performs well for detecting the maximum daytime PBL depth during the summer. Due to the limited availability of radiosondes at times when the PBL has fully developed, it is unknown how this performance varies throughout the year.

5. GEOS-5 Comparison

There are limited observational datasets with which to compare long-term, continuous PBL depth measurements like those obtained from MPLNET [*Liu and Liang*, 2010; *Seidel et al.*, 2012]. Furthermore, the lack of observational datasets makes the validation of modeled PBL depths difficult. Therefore, in this section we compare lidar-derived PBL depths with results from the GEOS-5 model. The GEOS-5 Atmospheric General Circulation Model (AGCM) was developed at NASA's Global Modeling and Assimilation Office (GMAO) as the single AGCM for use in a wide range of applications at a wide range of resolutions. The current version of the AGCM, documented in *Rienecker et al.* [2008] and *Molod et al.* [2012] was used for the GMAO coupled atmosphere/ocean and atmosphere-only simulations at 2.0° resolution submitted to the Coupled Model Intercomparison Project Phase 5 (CMIP5), is part of the GMAO's operational data assimilation system run at 0.25° resolution, and is used regularly for atmosphere-only coupled chemistry climate simulations.

A previous version of the GEOS-5 AGCM was used as part of the Modern Era Reanalysis for Research and Applications (MERRA). Direct comparisons between MERRA and Cloud Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) PBL depths perfomed by *Jordan et al.* [2010] resulted in correlation coefficients between 0.47 and 0.73 in the Western Hemisphere. However, their comparison included PBL depths derived using aerosol as well as cloud layers and contained a majority of data over the ocean. *McGrath-Spangler and Denning* [2012] showed that over much of the United States and portions of the subtropical oceans, the MERRA PBL depths are within 25% of the estimates derived from CALIPSO. The turbulence parameterization underwent substantial change in behavior between the previous and current versions of the GEOS-5 AGCM (documented in *Molod et al.* [2012]), resulting, in general, in larger PBL depths in the current simulations.

The full suite of GEOS-5 AGCM physical parameterizations is described in the references mentioned, but a brief description of the turbulence parameterization is warranted here. The turbulence parameterization in the GEOS-5 AGCM is a combination of the non-local scheme of *Lock et al.* [2000] and the local diffusion scheme of *Louis et al.* [1982]. At any model time step, the larger of the eddy diffusion coefficients computed by the two schemes are used for turbulent diffusion. The AGCM's estimate of PBL depth is based on vertical profiles of K_h, the eddy exchange coefficient for the vertical diffusion of heat. The first level above the ground at which K_h descends to below 2 m^2s^{-1} is designated as the PBL depth, and is used by the turbulence parameterization as an estimate of the turbulent length scale for use in the *Louis et al.* [1982] scheme. PBL depth estimates from a single atmospheric simulation at 0.5° horizontal resolution and 72 vertical levels (approximately 8 of them in the boundary layer) are used here for comparison against MPLNET PBL depths. The AGCM simulation is not expected to follow the synoptic evolution of the atmosphere, and so monthly mean diurnal cycles are used.

For this comparison, the lidar-derived PBL depths are averaged to the three-hour temporal resolution of the monthly mean diurnal cycle from the model. The comparison is limited to years when data was available from both GEOS-5 and MPLNET (2001–2008) and only includes months when at least 20 days of lidar measurements were made at GSFC. In total, 58 months met these requirements. Due to the aforementioned instrument limitations, the discussion is limited to daytime measurements when the PBL has fully developed.

Figure 9 shows a comparison of the annual mean diurnal cycles from the GEOS-5 model and MPLNET derived from the monthly mean diurnal cycles. The vertical bars indicate the standard deviation of the monthly means. Although the PBL appears to rise faster in the model, both the modeled and measured diurnal cycles peak at the same time. It should be noted that 34% of the monthly diurnal cycles from GEOS-5 peak one timestep before MPLNET; nearly all occurring between the months of April and August. However, since this comparison is performed at a coarse 3-hour resolution, the difference may be somewhat exaggerated.

Figure 10 shows the mean diurnal cycles for each season. In the spring and summer, when aerosol loading is highest, the lidar-derived PBL remains elevated late into the afternoon while it collapses sooner in the model. The most significant disagreement occurs during the winter, when the maximum daytime PBL depth from the model is nearly half the lidar-derived value. One possible explanation for these disagreements is the difference in criteria used to define the PBL depth (turbulence in the case of the GEOS-5 AGCM and aerosol gradients in the case of MPLNET) which can lead to different estimations of the PBL depth [*Seibert et al.*, 2000; *Tucker et al.*, 2009].

Figure 11 shows a comparison of the monthly mean daytime maximum PBL depths and the correlation plot between GEOS-5 versus MPLNET. The seasonal differences between the modeled and measured PBL depths are given in Table 1. The best agreement between the GEOS-5 and MPLNET PBL depths occurs in the fall and the largest differences occur during winter. During the spring and summer, it is believed that the modeled PBL depths are underestimated due to an overestimation in soil moisture in the Mid-Atlantic region based on a comparison of GEOS-5 precipitation to the Global Precipitation Climatology Project (not shown).

6. Summary and Future Work

An improved PBL depth algorithm has been developed for use in the MPLNET which uses a combination of the wavelet technique and image processing. A fuzzy logic routine is used to select the best estimate of the PBL depth from three extracted features using six membership functions. The improved algorithm reveals seasonal and diurnal trends undetected by the current operational routine. The improved algorithm has the advantage of being influenced less by clouds and residual layers. Instrument limitations make nighttime retrievals unreliable; therefore, MPLNET PBL depths are best suited for daytime retrievals under convective situations.

A July 2011 comparison with radiosonde observations suggests that the algorithm performs well for determining the maximum daytime PBL depth in the summer. Additional radiosonde data at non-standard times are needed to evaluate the algorithm performance at other times during the year. Comparisons with the GEOS-5 AGCM show the model may underestimate the maximum daytime PBL depth in the spring and summer by $\sim 22\%$. The largest differences between the model and lidar-derived PBL depths occur during the winter, when the GEOS-5 PBL depths are nearly half the values obtained from MPLNET.

Testing is being performed to evaluate the performance of the improved PBL depth algorithm at other sites in the MPLNET. Once finalized, the improved algorithm will be incorporated into regular processing and made available for public use. Further research is planned to fully explain and resolve differences between the MPLNET and GEOS-5 PBL depths and will be the topic of a future study. Comparisons with PBL retrievals from CALIPSO as demonstrated by *McGrath-Spangler and Denning* [2012] will also be investigated. While not explored in this study, the improved algorithm can be adapted to provide an estimate of the entrainment zone thickness, and will be researched at a later time.

Appendix A: Fuzzy Logic Membership Functions

The fuzzy logic algorithm used to select the PBL depth from the extracted feature altitudes calculates a quality score based on six membership functions. The feature altitude with the highest quality score is selected as the best estimate of the PBL depth. Each membership function, f_i, has a maximum value of unity and the quality score, Q, is the product of the individual membership functions.

$$Q = \prod_{i=1}^{6} f_i \tag{A1}$$

In this sense, the value of a membership function represents the likelihood that the extracted feature is the actual PBL depth based on that particular parameter. The membership functions have been developed through a trial-and-error process until they worked well to identify the PBL depth. Three distinct membership function types are used: Gaussian,

$$f(x; \sigma, c) = \exp[-\frac{(x-c)^2}{2\sigma^2}] \tag{A2}$$

Decaying exponential,

$$f(t; t_o) = \exp[-(t - t_o)]^4 \leq 1 \tag{A3}$$

and Absolute value

$$f(z; \bar{z}) = 1 - \left|\frac{z - \bar{z}}{\bar{z}}\right| \geq \frac{1}{3}. \tag{A4}$$

A summary of the six membership functions along with nominal parameter values is given in Table 2.

A1. Artifact Membership Function

The first membership function, f_1, accounts for an artifact in the WCT image that is related to the choice of dilation. This artifact is visible in the latter part of the day in Figure 1 as the lightly shaded area ~ 500 m, just above the minimum detectable gradient. When the PBL is low (e.g. near sunrise), real features can be detected at this altitude; however later in the day, false PBL depths can occur similar to the spurious low PBL depths in Figure 3 (black triangles from 18–24 UTC). To account for this artifact, a decaying exponential membership function is applied to features occurring within three range bins of the minimum detectable gradient. The parameter t_o is chosen as the time for the membership function to start decaying (e.g. sunrise). Therefore, feature altitudes occurring near the minimum detectable gradient are less likely to be chosen later in the day.

A2. Residual Layer Membership Function

The growth of the PBL in the morning is difficult to detect with lidar because stronger gradients can exist in the overlying residual layer. The second membership function, f_2, is used during early morning retrievals to reduce the probabilty of selecting the residual layer in the PBL depth algorithm. The mean altitude of the strongest gradients at nighttime is used to define the residual layer altitude, z_R. Then the value of the membership function is determined using a dimensionless parameter, x, given by

$$x = 1 - \frac{z}{z_R} \tag{A5}$$

where z represents the altitude of the extracted feature. A lower value of x is less likely to represent the true PBL depth.

A3. Elevated Layer Membership Function

Aerosol layers aloft in the atmosphere can produce false elevated PBL depths. In order to identify these elevated layers, the minimum altitude, z_{min}, where the scattering ratio falls below a certain threshold (e.g. the mean scattering ratio) is calculated for each five-minute averaged lidar profile. A dimensionless parameter, x, given by

$$x = 1 - \frac{z_{min}}{z} > 0 \tag{A6}$$

is used to determine the value of the third membership function, f_3. Features with a higher value of x are more likely to represent layers aloft and less likely to represent the actual PBL depth.

A4. WCT Membership Function

The PBL depth can be identified by the maximum value in the WCT. In the fourth membership founction, f_4, the WCT is normalized by the maximum value for each five-minute averaged lidar profile. The value of the normalized WCT at each extracted feature altitude is then used to calculate this membership function.

A5. Variance Membership Function

The altitude where the maximum variance in the lidar profile occurs can also be used to identify the PBL depth. Therefore, variance analysis at 20-minute intervals is used to calculate the fifth membership function. Similar to the WCT membership function, the value of the normalized variance at each extracted feature altitude is used to determine the value of f_5.

A6. Recent Retrieval Membership Function

The final membership function, f_6, uses a 20-minute average of the most recent PBL depth retrievals, $\bar{z}$, in the forward or reverse processing directions. Thus a higher probability of selection exists when the altitude, z, is closer to the mean. Because the membership function is used to determine the best choice between three feature altitudes, the minimum value of f_6 is set to one-third.

A7. Implementation of Membership Functions

Due to the empirical nature of the individual membership functions, the PBL algorithm will need to be parameterized for each site based on the meteorological conditions. For example, f_1 is applied at GSFC only when a feature is detected within three range bins of the minimum detectable gradient. At other sites within the network, the altitude at which this membership function is applied may differ. Similary, f_2 is only applied during the first three hours after sunrise at GSFC, but this time interval may differ depending on the expected rate of PBL growth at a particular site. The sensitivity of each of the Gaussian membership functions depends on the parameter σ. Smaller values of σ produce more sensitivity in the PBL retrieval. Because f_6 depends only upon the most recently retrieval, it can be implemented unchanged at every site in the network.

Acknowledgments. The authors would like to acknowledge the support of Larry Belcher for processing the lidar data, Lawrence Takacs who ran the model simulations, and the MPLNET staff for their efforts in establishing and maintaining the GSFC site. This research was supported by an appointment to the NASA Postdoctoral Program at the Goddard Space Flight Center, administered by Oak Ridge Associated Universities through a contract with NASA. The NASA Micro-Pulse Lidar Network is funded by the NASA Earth Observing System and Radiation Sciences Program.

References

Angevine, W.M., A.B. White, and S.K. Avery (1994). Boundary-layer depth and entrainment zone characterization with a boundary-layer profiler, *Bound. Layer Meteor.* 68: 375–385.

Ao, C.O., D.E. Waliser, S.K. Chan, J.-L. Li, B. Tian, F. Xie, and A. J. Mannucci (2012). Planetary boundary layer heights from GPS radio occultation refractivity and humidity profiles,*J. Geophys. Res.*, 117, D16117, doi:10.1029/2012JD017598.

Beyrich, F. (1997). Mixing height estimation from sodar data - A critical discussion, *Atmos. Environ.* 31 (23), 3941–3953.

Bianco, L. and J.M. Wilczak (2002). Convective boundary layer depth: Improved measurements by doppler radar wind profiler using fuzzy logic methods, *J. Atmos. Ocean. Tech.*, 19, 1745–1758.

Brooks, I.M. (2003). Finding boundary layer top: Application of a wavelet covariance transform to lidar backscatter profiles, *J. Atmos. Ocean. Tech.*, 20, 1092–1105.

Brooks, I.M. and A.M. Fowler (2012). An evaluation of boundary-layer depth, inversion, and entrainment parameters by large-eddy simulation, *Bound.-Layer Meteor.*, 142:245-263, doi:10.1007/s10546-011-9668-3.

Campbell, J.R., D.L. Hlavka, E.J. Welton, C.J. Flynn, D.D. Turner, J.D. Spinhirne, V.S. Scott, and I.H. Hwang (2002). Full-time, eye-safe cloud and aerosol lidar observation at Atmospheric Radiation Measurement program sites: Instrument and data processing,*J.*

Atmos. Oceanic Technol., 19, 431–442.

Canny, J. (1986). A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8 (6), 679-698.

Davis, K.J., N. Gamage, C.R. Hagelberg, C. Kiemle, D.H. Lenschow, and P.P. Sullivan (2000). An objective method for deriving atmospheric structure from airborne lidar observations. *J. Atmos. Oceanic Technol.*, 17, 1455-1468.

Ecklund, W.L., D.A. Carter, and B.B. Balsley (1988). A UHF wind profiler for the boundary layer: Brief description and initial results, *J. Atmos. Ocean. Technol.* 5: 432–441.

Emeis, S. (2011). Surface-based remote sensing of the atmospheric boundary layer, Series: Atmospheric and oceanographic sciences library, Vol.40. *Springer*.

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

Guo, P., Kuo, Y.-H., Sokolovskiy, S. V., and Lenschow, D. H. (2011). Estimating atmospheric boundary layer depth using COSMIC radio occultation data, *J. Atmos. Sci.*, 68, 1703-1713.

Haeffelin, M., F. Angelini, Y. Morille, G. Martucci, S. Frey, G. P. Gobbi, S. Lolli, C.D. O'Dowd, L. Sauvage, I. Xueref-Rémy, B. Wastine, and D.G. Feist (2011). Evaluation of mixing-height retrievals from automatic profiling lidars and ceilometers in view of future integrated networks in Europe, *Bound. Layer Meteor.* vol, pp.

Holben, B. N., T. F. Eck, I. Slutsker, D. Tanré, J. P. Buis, A. Setzer, E. Vermote, J. A. Reagan, Y. J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov (1998), AERONET - A federated instrument network and data archive for aerosol characterization. *Remote Sens. Environ.,* 66: 1–16.

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

Holzworth, G.C. (1967). Mixing depths, wind speeds and air pollution potential for selected locations in the United States, *J. Appl. Meteor.* 6: 1039-1044.

Hooper, W.P. and E.W. Eloranta (1986). Lidar measurements of wind in the planetary boundary layer: the method, accuracy and results from joint measurements with radiosonde and kytoon. *J. Clim. Appl. Meteor.* 25: 990-1001.

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

Kaimal, J.C. and J.E. Gaynor (1983). The boulder atmospheric observatory, *J. Climate Appl. Met.* 22: 863-880.

Klir, G.J. and B. Yuan (1997). Fuzzy set theory foundations and applications. *Prentice Hall*, 245 pp.

Lammert, A. and J. Bösenberg (2006). Determination of the convective boundary-layer height with laser remote sensing, *Bound. Layer Meteor.* 119: 159-170.

Liu, S., and X.-Z. Liang (2010), Observed diurnal cycle climatology of planetary boundary layer height, *J. Clim.*, 23, 5790-5809, doi:10.1175/2010JCLI3552.1.

Lock, A.P., A.R. Brown, M.R. Bush, G.M. Martin, and R.N.B. Smith (2000). A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Mon. Wea. Rev.*, 138, 31873199.

Louis, J.F., M. Tiedtke, and J.F. Geleyn (1982). A short history of the operational PBL parameterization at ECMWF, in *Proceedings of ECMWF Workshop on PBL Parameterization*, ECMWF, Reading, UK.

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.

Melas, D. (1990). Sodar estimates of surface heat flux and mixed layer depth compared with direct measurements, *Atmos. Environ.* 24A, (11), 2847–2853.

Molod, A., L. Takacs, M. Suarez, J. Bacmeister, I.-S. Song, and A. Eichmann (2012). The GEOS-5 Atmospheric General Circulation Model: Mean Climate and Development from MERRA to Fortuna. Technical Report Series on Global Modeling and Data Assimilation, Vol. 28.

Morille, Y., M. Haeffelin, P. Drobinski, J. Pelon (2007). STRAT: an automated algorithm to retrieve the vertical structure of the atmosphere from single-channel lidar data. *J. Atmos. Ocean Tech.* 24, 761-775.

Olsen, L.E., M.P. McCormick, W.P. Elliott, S.H. Melfi (1974). An observational study of the mixing layer in western Oregon, *Atmos. Environ.* 8, 214–252.

Rienecker, M.M., M.J. Suarez, R. Todling, J. Bacmeister, L. Takacs, H.-C. Liu, W. Gu, M. Sienkiewicz, R.D. Koster, R. Gelaro, I. Stajner, and J.E. Nielsen (2008). The GEOS-5 Data Assimilation System - Documentation of Versions 5.0.1, 5.1.0, and 5.2.0. Technical Report Series on Global Modeling and Data Assimilation, Vol. 27.

Seibert P., F. Beyrich, S. Gryning,S. Joffre, A. Rasmussen, P. Tercier (2000). Review and intercomparison of operational methods for the determination of the mixing height.

Atmos. Environ. 34: 1001-1027.

Seidel, D.J., Y. Zhang, A. Beljaars, J.-C. Golaz, A. R. Jacobson, and B. Medeiros (2012). Climatology of the planetary boundary layer over the continental United States and Europe. *J. Geophys. Res.*, 117, D17106, doi:10.1029/2012JD018143.

Spangler, T.C. and R.A. Dirks (1974). Meso-scale variations of the urban mixing height, *Bound.-Layer Meteor.* 6: 423–441.

Spinhirne, J.D., J.A.R. Rall, and V.S. Scott (1995), Compact eye safe lidar systems, *Rev. Laser Eng.*, 23, 112–118.

Stull, R.B. (1988). An introduction to boundary-layer meteorology, *Kluwer Acad. Publ., Dordrecht*, 666 pp.

Steyn, D.G., M. Baldi, and R.M. Hoff (1999). The detection of mixed layer depth and entrainment zone thickness from lidar backscatter profiles, *J. Atmos. Ocean Technol.* 16: 953-959.

Träumner K., Ch. Kottmeier, U. Corsmeier, A. Wieser (2011). Convective boundary-layer entrainment: short review and progress using doppler lidar. *Bound.-Layer Meteor.* 141:369-391, doi:10.1007/s10546-011-9657-6.

Tucker, S.C., W.A. Brewer, R.M. Banta, C.J. Senff, S.P. Sandberg, D.C. Law, A.M. Weickmann, and R.M. Hardesty (2009). Doppler lidar estimation of mixing height using turbulence, shear, and aerosol profiles, *J. Atmos. Ocean. Tech.*, 26 (4), 673–688.

van Ulden, A.P., and J. Wieringa (1996). Atmospheric boundary layer research at Cabauw, *Bound.-Layer Meteor.* 78: 39-69.

von Engeln, A., J. Teixeira, J. Wickert, and S.A. Buehler (2005). Using CHAMP radio occultation data to determine the top altitude of the planetary boundary layer, *Geophys.*

Res. Lett., 32, L06815.

Welton, E. J., J. R. Campbell, J. D. Spinhirne, and V. S. Scott (2001), Global monitoring of clouds and aerosols using a network of micro-pulse lidar systems. *In: Lidar Remote Sensing for Industry and Environmental Monitoring* [U. N. Singh, T. Itabe, N. Sugimoto (eds.)]. Proc. SPIE, 4153, 151–158.

Welton, E.J., and J.R. Campbell (2002). Micropulse lidar signals: Uncertainty analysis. *J. Atmos. Oceanic Technol.*, 19, 2089–2094.

Wielicki, B. A., R. D. Cess, M. D. King, D. A. Randall, and E. F. Harrison (1995), Mission to Planet Earth: Role of clouds and radiation in climate. *Bull. Amer. Meteor. Soc., 76*, 2125–2153.

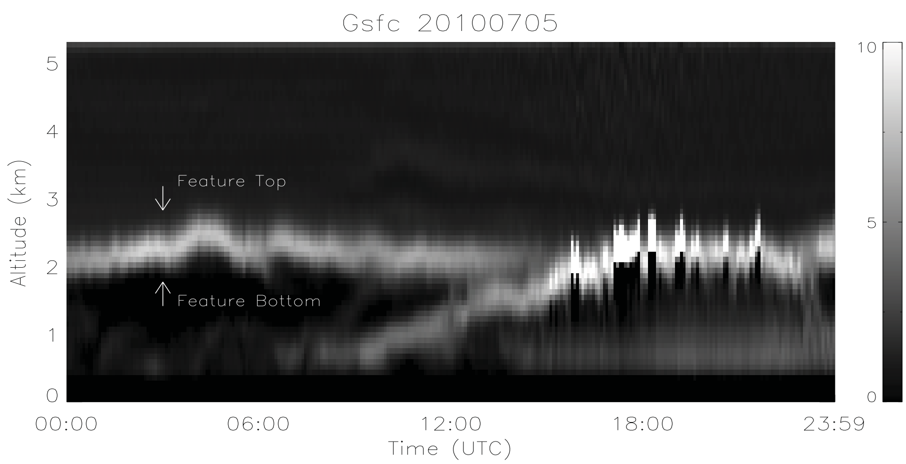

Figure 1. WCT image (arbitrary units) at GSFC on 5 July 2010. Gradients in the lidar profile are not detectable below ~ 500 m.

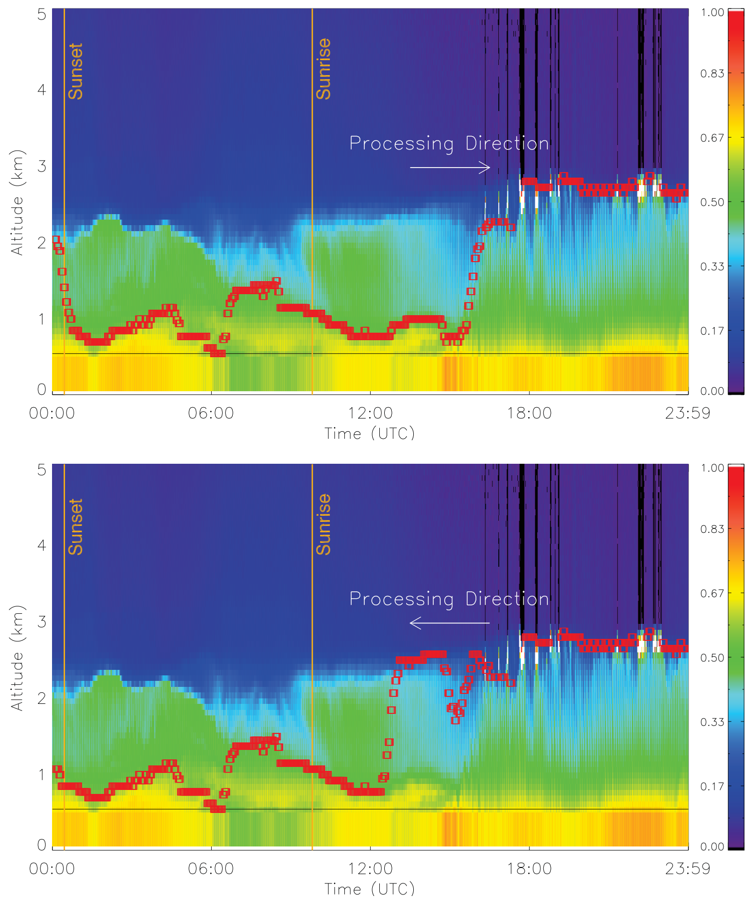

Figure 2. Normalized relative backscatter at GSFC on 6 July 2010 showing a comparison of the improved PBL depth algorithm in the forward (top) and reverse (bottom) processing directions. The best estimate of the PBL depth is indicated by red squares. The vertical orange lines indicate the mean times for sunrise (SR) and sunset (SS) and the horizontal black line indicates the altitude of the minimum detectable gradient.

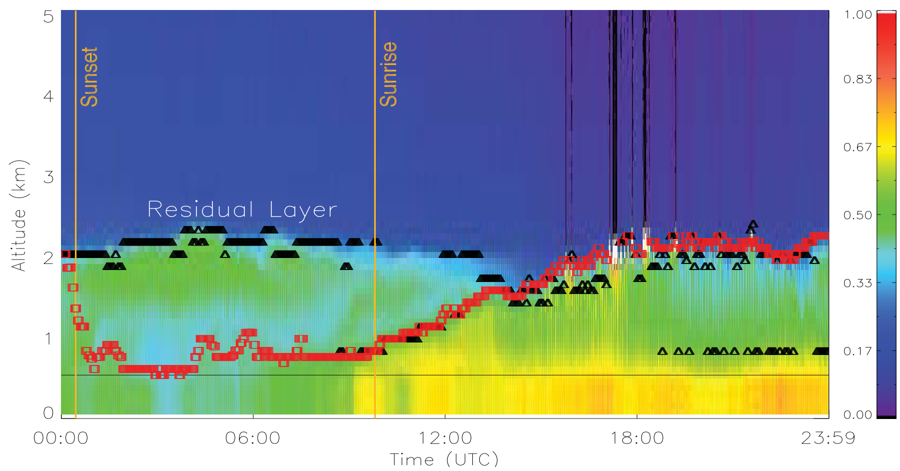

Figure 3. Normalized relative backscatter at GSFC on 5 July 2010. The black triangles and red squares are the operational and improved PBL depths, respectively. The horizontal black line indicates the altitude of the minimum detectable gradient.

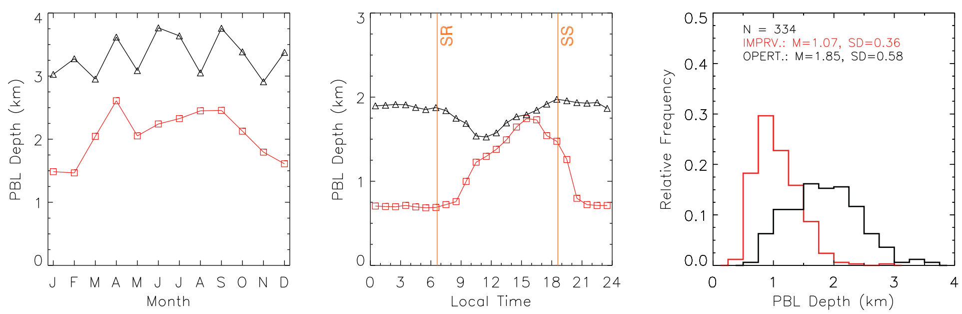

Figure 4. Comparisons of (left) monthly means of the daily maximum PBL height, (center) annual diurnal cycles, and (right) daily mean probability distributions at GSFC for the 2010 operational PBL retrieval (black triangles) and improved PBL algorithm (red squares). The vertical orange lines in the diurnal cycle indicate the mean times for sunrise (SR) and sunset (SS).

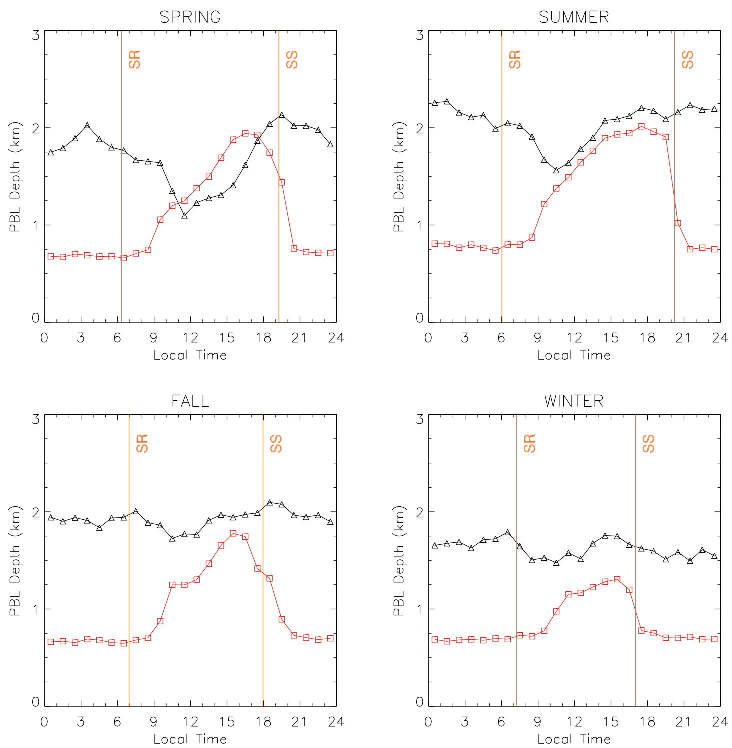

Figure 5. Seasonal diurnal cycle of the PBL depth at GSFC for 2010 with the operational retrieval represented by black triangles and the improved retrieval represented by red squares.

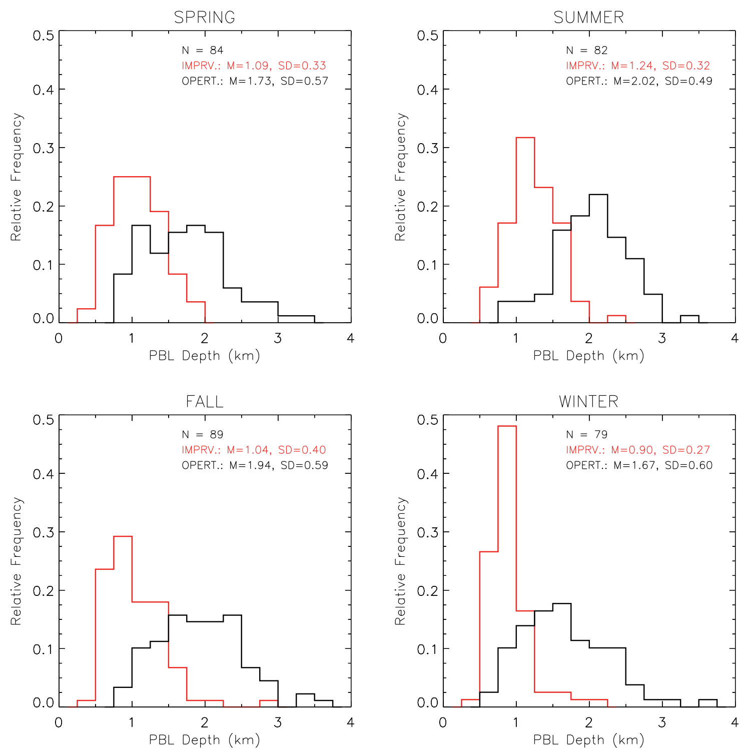

Figure 6. Seasonal probability distribution of the daily mean PBL depth at GSFC for 2010 with the operational retrieval in black and the improved retrieval in red.

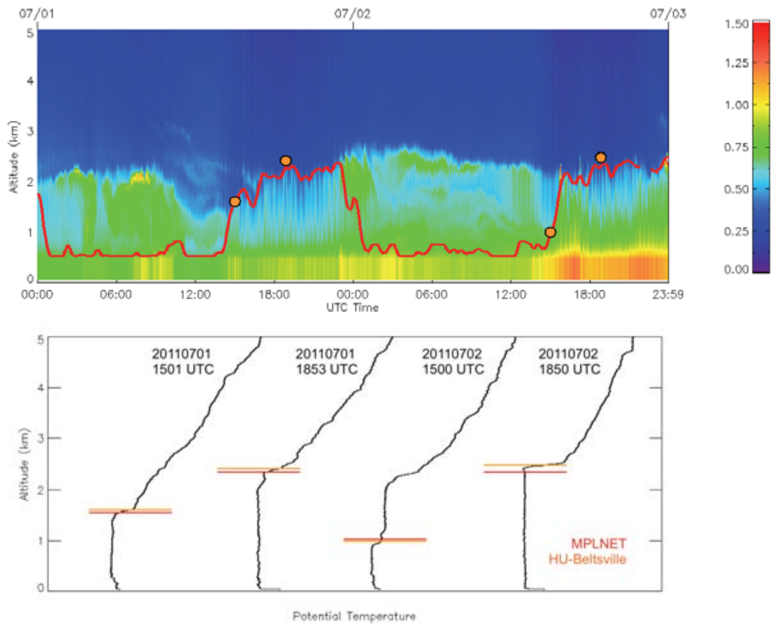

Figure 7. (Top) Normalized relative backscatter at GSFC on 1–2 July 2011 with the PBL depths from MPLNET (red line) and radiosondes (orange filled circles). (Bottom) The potential temperature profiles from the the radiosonde profiles with the PBL depths from MPLNET (red) and radiosondes (orange).

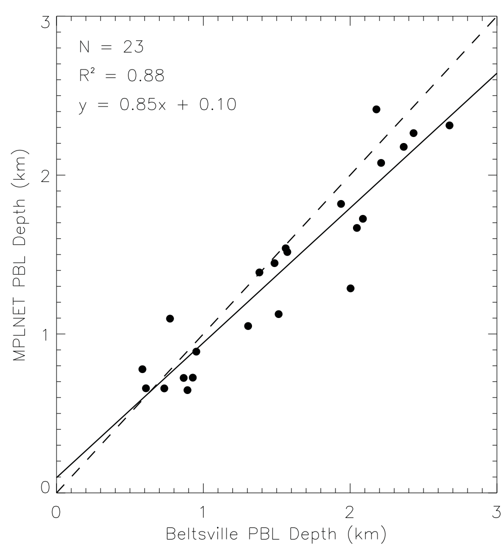

Figure 8. Correlation of radiosonde-derived PBL depths at Beltsville and lidar-derived PBL depths from MPLNET. The dashed line is the unity line and the solid line is the best-fit line.

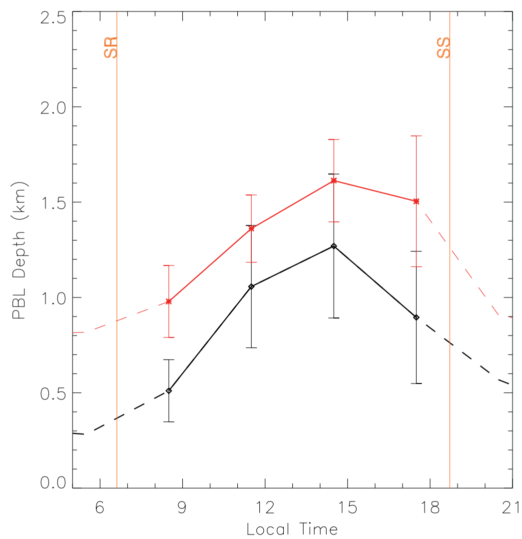

Figure 9. Comparison of the annual mean diurnal cycles from the GEOS-5 model (black diamonds) and MPLNET (red squares) derived from the monthly mean diurnal cycles from 2001–2008. Daytime (nighttime) retrievals are symbolized using solid (dashed) lines. The vertical bars indicate the standard deviation of the monthly means.

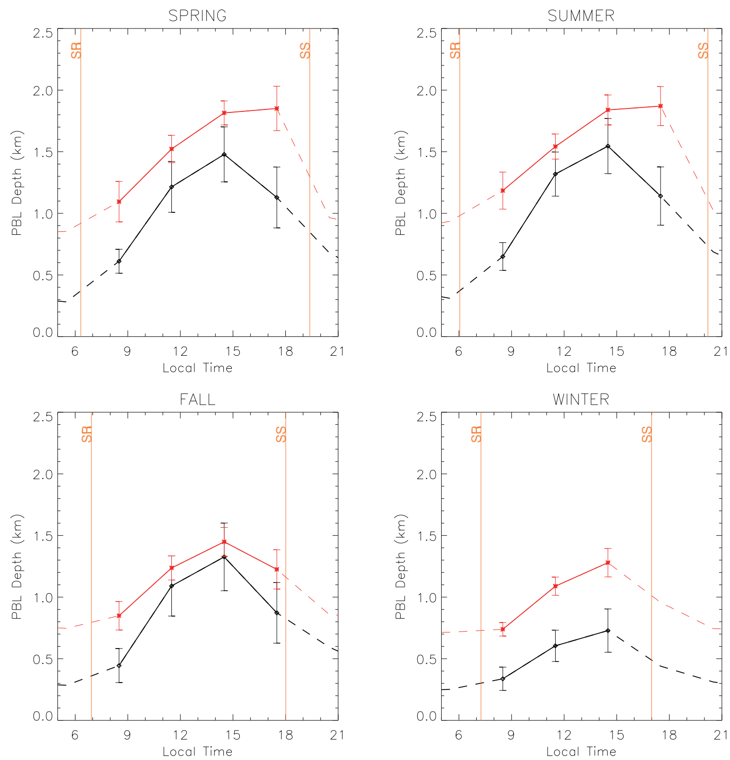

Figure 10. Comparison of seasonal diurnal cycles of the PBL at GSFC for 2001–2008 from MPLNET (red squares) and GEOS-5 (black diamonds).

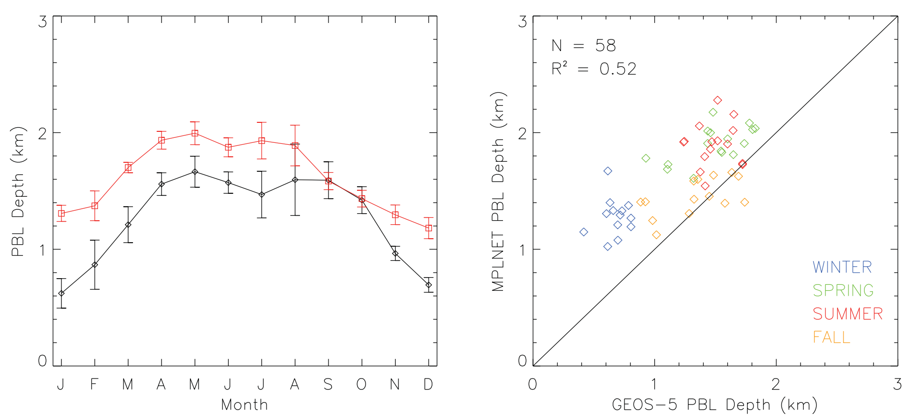

Figure 11. (Left) Comparison of monthly mean daytime maximum PBL depths for MPLNET (red squares) and GEOS-5 (black diamonds) from 2001–2008. (Right) Correlation plot between GEOS-5 and MPLNET for each month.

Table 1. Seasonal difference between MPLNET and GEOS-5 PBL depths

Season	h_{MPLNET} (km)	h_{GEOS5} (km)	Δh (km)	$\sigma_{\Delta h}$(km)	Months
Winter	1.28	0.68	0.60	0.19	13
Spring	1.90	1.49	0.41	0.21	16
Summer	1.90	1.49	0.41	0.24	15
Fall	1.45	1.33	0.12	0.23	14
All	1.65	1.27	0.38	0.27	58

Table 2. Fuzzy Logic Membership Functions

f_i	Type	Parameter	Parameter
f_1	Decaying exponetial	t_o = sunrise	-
f_2	Gaussian	c = 1	$\sigma = 0.4$
f_3	Gaussian	c = 0	$\sigma = 0.1\bar{6}$
f_4	Gaussian	c = 1	$\sigma = 0.68$
f_5	Gaussian	c = 1	$\sigma = 0.68$
f_6	Absolute value	$\bar{z}$ = mean PBL height	-