

## Response to referee #1 comments

Authors gratefully thank to referee #1 for his/her thorough reviews and valuable comments which would contribute to improve the manuscript. Authors have revised the manuscript to respond the referee's comments. Major changes are marked in **red** in the revised manuscript. I hope that this manuscript will be accepted for the publication in AMT.

### General Comments:

The authors were asked to compare with previous studies of individual methodologies. Although more references were added, the fact that previous studies of individual methodologies show better results those than shown in the manuscript was not addressed.

> Most conventional methodologies have been verified for daytime clear skies during the several days. While this study tried to attempt to consider all available weather conditions including cloudy sky during the 4 years. Therefore, the verification values were lower than those for the previous studies. The above expression is added in L515-517.

The methodology presented here does not address the impact of clouds/precipitation

and individual methodologies do not use post-processing application which is likely the reason why individual methodologies perform worse than those in literature. By consider the previous literature, the authors should note that ISABLE uses existing methodologies and then applies a cluster technique to addresses layer attribution in post processing. As all other individual methodologies are not applied with any improvements for layer attribution, consider stating that ISABLE improved PBLH retrievals due to the layer attribution techniques. Stating the ISABLE was superior to individual methodologies is misleading as individual methodologies were not aided by post processing techniques as ISABLE was.

> Not cloud but precipitation is removed before data processing in this studies. The effects on clouds were explained in **Fig. 11** and **Table 3**.

> The core post-processing technique of this study is to cluster for multiple layers. The ABLH calculated by the methodology used in the previous studies are not multi-layer but single layer. On the other hand, ISABLE can cluster by calculating several ABLHs using four methodologies, and improved ABLH calculation performance through statistical post-processing.

> A description of multiple ABLH estimation using four different methodologies and performance improvement through statistical post-processing is given in **L474-475**: The ISABLE developed in this study integrated the conventional ABLH estimation methodologies to produce optimal ABLH and applied statistical post-processing techniques to improve accuracy.

### Specific Comments:

L26: consider the term mixed layer. As a dynamic layer, the ML is not under a finite state.

> The term 'mixed layer (ML)' is changed to the term 'a well-mixed layer (ML) or a convective boundary layer (CBL)'. **(L26)**.

L53: Correct to "In a ML"

> The term 'In ML' is corrected to the term 'In a ML' in **L54**.

L66: Studies have shown that the EKF technique is sensitive to low SNR and to indeed require further longtime

averaging and range smoothing. Please revise this statement.

> The term 'except for low SNR' is added at the end of the statement in L68.

L180-185: It is stated that defining the RL using radiosonde profiles is difficult due to large variations in temperature and wind profiles. L184-185 later states that the top of the RL is determined as the SBLH due to large variations in temperature and turbulence. At first glance it seems contradicting, please expand on this. Were these occurrences during certain conditions, or errors from the retrieval method?

> The term 'RL' is changed to the term 'SBLH' in L181.

L367: Provide uncertainties here

> The uncertainty was analyzed in terms of IQR (interquartile range) and added in L368-371 such that: At night, the mean ABLHs were determined as around 500 m, and outliers appeared above 1 km, which were identified as the RL or clouds (Fig. 8). The interquartile range (IQR;  $Q3 - Q1$ ) showed the minimum value (268 m) at 0900 LST and the maximum (740 m) at 1800 LST. Overall, ABLHs were concentrated in the lower layer at night, and the IQR values increased as the ML developed after sunrise.

L368-371: Please revise, repetition and grammatical errors make these statements unclear. What are rural surfaces? What does 'that' refer to? What is defined as a compact urban surface?

> The repetition is removed, and ambiguous expression or grammatical errors are corrected in L372-375: The SBL over rural areas such as grass or cropfield is well developed due to active radiative cooling at night, especially under clear skies. On a while, the radiative cooling over urban areas was not always active because of heat storage by urban materials and anthropogenic heat by energy use (Hong et al., 2013; Park et al., 2014). As a result, formation and evolution of SBL were not active over dense urban areas such as Jungnang station.

L371: As previously stated "However, that over urban areas is not always developed" in L368-369, L371 should state that 'SBL were not always active'.

> The sentence is changed to 'As a result, formation and evolution of SBL were not active over dense urban areas such as Jungnang station' in L374-375.

L372: "which were determined using the residual layer or clouds" should this state "which were identified as the residual layer or clouds"?

> The sentence is changed to 'which were identified as the RL or clouds' in L369.

L380: R symbol needs to be corrected

> It is corrected as 'R' in L381.

L382: Is this still referring to ISABLE results or all ABLH results? I suspect this is referring to GM, CLST, and WAV results. Please clarify. If this refers to ISABLE please note that the residual layer contains the remains of the previous day's mixing layer at similar heights.

> The statement is clarified in L382-385 such that: The ABLHs from ISABLE as well as ceilometer-based methods (GM, WAV2, WAV3, and CLST) were similar to those by RS during the daytime, however, the ABLHs from the former appeared at higher levels than those from the latter during the nighttime.

With peak daytime mixing layers in Figure 9 at ~1600 (RS), residual layers signals would be expected in similar heights. Figure 9 shows ISABLE overestimations not exceeding ~500m therefore, how can the overestimation of ISABLE during nighttime be attributed to residual layer signals? Instead, it is likely due to additional stratification of the PBL and therefore additional lofted aerosol layer during nighttime.

> As can be seen in Fig.7, it is difficult to judge this as a residual layer signal. The difference between RS and ISABLE at 1500 LST on 23 is due to cumulus cloud. The following paragraph is added in **L385-395**: ISABLE tried to complement the shortcomings by integrating the four methodologies through considering the SBL using a vertical temperature from MWR at night. The maximum ABLHs during daytime appeared at 1600 LST on 23, the RS and the ISABLE estimated ABLHs of 1,620 m, 2,009 m, respectively. At this time, a cumulus cloud was formed over the top of ABL due to strong convection, and the cloud base height observed by the ceilometer was 1,910 m. The ABLHs estimation results showed that RS was below the cloud, while ISABLE and individual methodologies (GM: 2,080 m, WAV2: 2,060 m, WAV3: 2,050 m) detected ABLHs as the cloud. In the presence of clouds, the  $Ri_b$  method tends to detect the lower layer of the cloud, where the temperature profile changes rapidly. The GM and WAV2 methods using the ceilometer determine the ABLHs as the top of cloud layer because of strong negative gradient of backscattering coefficient, whereas the CLST can detect both the bottom and top of cloud layer. In ISABLE, the effect of clouds is compensated for averaging multiple heights determined by individual methodologies. However, the ISABLE still has limitations in the presence of thick clouds.

L391: MB was not defined in text

> The MB means mean bias. The definition is added in **L382**.

L393: Please clarify what “The scatter distribution of GM, WAV2, and CLST at sunrise, sunset, and nighttime could be fitted to two different linear functions” means

>> The sentence is clarified in 405 such as: The scatter distribution could be divided into two groups with different linear functions.

L394: “In cases where symbols were plotted below the trend line (dashed line), RLs during nighttime or cloud layers in daytime existed at the layer”. Please clarify the effect of “cloud layers [existing] at the layer”

> Effect of cloud is explained in **L390-395**: In the presence of clouds, the  $Ri_b$  method tends to detect the lower layer of the cloud, where the temperature profile changes rapidly, as the ABLH. The GM and WAV2 methods using the ceilometer detect the upper layer of the cloud because the height at which the backscattering coefficient rapidly decreases is determined by ABLH, whereas the CLST can detect both the lower and upper layer of the cloud. In ISABLE, the effect of clouds is compensated for averaging multiple heights determined by individual methodologies. However, the ISABLE still has limitations in the presence of thick clouds.

L399: As stated in L367-371, the urban effect will impact the SBL detection from RS. Consider clarifying this statement by adding this overestimation effect on RS heights that lead to an underestimation when compared to ABLHs.

> It is an underestimation of the frequency of occurrence, not the height. Since there is a possibility of misunderstanding, it has been modified to be more precise in **L410-411**: Anthropogenic heat release from urban materials could be one reason for detecting less number of SBLHs at night.

L401: Use WAV3

> It is modified to “WAV3” in **L412**.

L405-408: Similar as above, please clarify the statement “could be fitted to two different linear lines”

> The sentence is clarified in L419 such as: The scatter plots could be classified into two groups with different linear lines.

L419-420: “The ABLHs for clear skies were significantly higher than those for cloudy skies during the daytime, however, the difference was not as significant during the **daytime**” Is this repetitive or a typo?

> It is corrected as “nighttime” in **L432**.

L479-481: What was further verified?

> It is corrected to “. Furthermore, the ISABLE was verified through the separation of the data into four time zones” in **L492-493**.

L483-484: See general comments above. The ‘superior’ performance of ISABLE is largely due to the addition of post-processing techniques. As individual methodologies are not applied with any post-processing techniques, it is not correct to state a superior performance. Instead, the large improvement that was seen with the ISABLE post-processing technique should be highlighted.

> The individual methodologies of previous studies determine one layer, and even if multiple layers are additionally determined, the cluster classification of post-processing is not possible. On the other hand, ISABLE can cluster by calculating several ABLHs using four methodologies, and improved ABLH calculation performance through statistical post-processing.

> A description of multiple ABLH estimation using four different methodologies and performance improvement through statistical post-processing is given in **L474-475**: The ISABLE developed in this study integrated the conventional ABLH estimation methodologies to produce optimal ABLH and applied statistical post-processing techniques to improve accuracy.