Author's Response
to Reviewer Feedback on the Preprint
"Improved method for temporally interpolating radiosonde profiles"

Linus von Klitzing (linus.vonklitzing@uni-hohenheim.de)[1], David D. Turner[2], Diego Lange[1], and Volker Wulfmeyer[1]

[1]Institute of Physics and Meteorology, University of Hohenheim, Stuttgart, Germany
[2]Global Systems Laboratory, NOAA, Boulder, CO, USA

November 14, 2025

Contents

1 Introductory Statement **1**

2 Comments by Referee #1 **1**
2.1 General, Major Scientific, and Technical Comments 1
2.2 Specific Comments 10

3 Comments by Referee #2 **16**
3.1 General, Scientific, and Technical Comments 16
3.2 Minor Revisions 19

1 Introductory Statement

The authors thank the reviewers for their constructive and encouraging comments. These are printed in **bold face**, our responses in *italic face*, and our taken actions in normal face. If text passages are removed or replaced, these passages are recapitulated here as ~~crossed-out text~~, followed by the, if applicable, replacement in normal face.

2 Comments by Referee #1

2.1 General, Major Scientific, and Technical Comments

Comment: The manuscript describes a new method for interpolating radiosonde observations to provide a high resolution (in time) profile of potential temperature and water vapor mixing ratio for the convective boundary layer.

The technique employs a normalization of the height coordinate by the planetary boundary layer height before interpolating in time helping to preserve the inversion-topped structure. Interpolated profiles employing this technique are compared to a case with no height normalization showing marked improvement compared to independent soundings. The products developed using this new technique could have important used for those study boundary layer transport and cloud development at sites where more advanced remote sensing is unavailable. There are several important issues that need to be clarified before the manuscript is ready for publication in AMT. The most significant of these are outlined in my general comments and are related to more clearly defining the targets and the motivations for the products that would be developed using this technique and considering some of the uncertainties related to varying environmental conditions and the instrumentation used for PBLH estimates. With this in mind, I recommend the paper be considerd for publication following major revisions.

Reply: *We agree that our motivation and product descriptions need some refinements and clarifications. Our replies to all individual comments and corresponding improvements of the manuscript are listed below.*

Action: Actions are taken in response to the individual comments.

Comment: The methodology presented in the manuscript focuses entirely on the convective boundary layer. This should be reflected in the title. I would suggest adding "...in the convective boundary layer" to the end of the current title.

Reply: *This is correct and we agree that this change helps in describing the content of this study more precisely.*

Action: The previous title "~~Improved method for temporally interpolating radiosonde profiles~~" is replaced by "Improved method for temporally interpolating radiosonde profiles in the convective boundary layer".

Comment: Lines 21-27: In this introductory paragraph, one of the motivations is the accuracy of the interpolated sounding product as it is used in the ARM MICROBASE product. However, the main use of that thermodynamic profile is in determining the "mixed-phase" layer, i.e. where the temperature is between 0C and -16C The manuscript focuses on an improved methodology in the convective boundary layer as applied to a period from mid-April through mid-July, a time period when I expect the freezing level will generally (if not always) be above the boundary layer and so improvement of the boundary layer thermodynamic profile would have no impact on the MICROBASE product. It might be better to provide a more applicable motivation. Similarly, there is a reference provided to the "MERGED SOUNDING" product (Troyan, 2010) which incorporated model output to improve the interpolation between soundings and might provide improved boundary layer thermody-

namics over interpolated sonde. Since MERGED sounding is not available in recent years it may be best to not reference unless it is discussed.

Reply: *Thanks for pointing us to this flaw in our motivation. We decided to replace the part concerning the ARM MICROBASE (and therefore MERGED SOUNDING) product with a broader scientific motivation on the importance of the availability of high-quality profiles of potential temperature and humidity. Additionally, we added our own personal research on the quantification of entrainment fluxes as a concrete example for the use of the presented interpolation technique.*

Action: The paragraph (lines 19-27 of the Preprint) "~~Consequently, temporal interpolations of RS profiles with high accuracy are very important. To address this need for the SGP site, ARM routinely provides a simple interpolation product, which linearly averages the nearest soundings on a fixed height grid (Fairless et al., 2021). An example of the application of this data product is the multisensor cloud retrieval product MICROBASE, which aims to produce macro- and microphysical properties of all clouds over the site. Its strength is its ability to apply to all clouds and conditions, enabled by various retrieval techniques, but more importantly for this work, it includes temporally interpolated RS data (Shupe et al., 2016; Dunn et al., 2011; Troyan et al., 2010). This interpolated RS data is used right at the beginning of the corresponding algorithm to determine the cloud water phase and, with that, the radar reflectivity (Dunn et al., 2011). More accurate data can be expected to improve the accuracy of this step and, consequently, the accuracy of the data product.~~" is replaced by "The study of turbulent transport processes (e.g., Wulfmeyer et al., 2016; Gibert et al., 2025), the feedbacks between land and atmosphere (e.g., Santanello et al., 2018; Wulfmeyer et al., 2020), as well as many other fields of atmospheric research, rely on high-resolution vertical profiles of temperature and humidity. A concrete example of this need, and initial motivation behind the presented study, is the investigation of the latent heat entrainment flux at the top of the convective boundary layer (CBL). This requires high-resolution data of humidity and vertical wind (Wulfmeyer et al., 2016), both available as lidar data products at the ARM SGP site (Newsom and Sivaraman, 2018; Newsom and Krishnamurthy, 2022). What is also required, but not available, is a temporally and vertically high-resolution profile of potential temperature that can capture, especially, the strong gradient at the top of the CBL at an arbitrary time. The ARM SGP Raman lidar has too high noise levels for this purpose (Osman et al., 2019), and the simple RS interpolation product (Fairless et al., 2021) by ARM, which linearly averages the nearest soundings on a fixed height grid, fails in recreating the actual structure of the CBL (demonstrated later). To address this need, this study proposes an improved interpolation technique." (lines 19-30 of the revised manuscript).
As the acronym "CBL" is defined earlier through the above changes, the original definition is replaced by the acronym (line 38 of the revised manuscript).

Comment: Lines 34-36: Do you consider both well-mixed and decoupled boundary layers in your analysis? In the case of decoupled boundary layers, I believe that even with the normalized height grid there will still be a smoothing of the decoupling layer that may or may not be better captured with the new technique.

Reply: *To avoid any misunderstandings, our understanding of a decoupled boundary* *layer is that of a boundary layer in which turbulence generation at the surface is too weak to connect the lower and upper parts of the boundary layer. Our study solely considers boundary layer situations between the morning transition and the afternoon decay (radiosondes between 14:30 and 23:30 UTC) for which a well-mixed layer and an inversion layer were clearly present, as only this consistent structure, temporally stretching over at least two soundings, allows for the presented improvements to the interpolation at an in-between time for the majority of the boundary layer. The latest utilized sounding (at 23:30 UTC) takes place during the afternoon decay with weaker turbulence production at the surface, and therefore, in some cases, exhibits hints (smaller than daytime near-surface temperature gradients) of a forming stable boundary layer at the surface with a residual layer above. These included cases, however, are still very similar in their structure to the CBL, so we decided to include them, although it might lead to some smoothing close to the surface. We added a corresponding description to the paragraph in question to make the investigated atmospheric setting clearer.*

Action: We inserted into the corresponding introductory paragraph (lines 44-48 of the revised manuscript) "To be precise, this study utilizes soundings of the boundary layer that took place between the morning transition and the afternoon decay. Especially the included soundings at 23:30 UTC might therefore take place during an on-setting decoupling of the lower and upper parts of the boundary layer as the turbulence generation at the surface grows weaker. However, the corresponding profiles are very similar in their structure to clear CBL profiles with a well-mixed/residual layer and an inversion layer above.".

Comment: Lines 43-45: The estimate of the PBLH will have some dependence on the instrument used. Have you considered the impact of these different definitions/measurements of the PBLH? E.g., I am wondering if the increase in standard deviation of the water vapor mixing ratio in figure 3d near the top of the boundary layer is mostly due to varying definitions of the PBL depth? I also wonder, if the technique REQUIRES this separate measurement of PBLH, would there still be value in using the interpolated PBLH depth (as determined from the radiosonde profiles) to normalize the height? I would hypothesize that this would still provide an improvement over the standard grid interpolation.

Reply: *We absolutely agree that the PBLH estimate will depend on the quality of the used product. However, we did not test this dependence explicitly, as we wanted to focus solely on demonstrating the new interpolation technique and therefore deemed the quality assessment of individual data products out of scope. That said, we knew that every product we chose would have its own weaknesses. To address this, we implemented a bias correction, briefly described in Section 3.3. We added a clarifying description to the corresponding introductory paragraph.*
The increasing standard deviation at the top of the boundary layer of the bias for potential temperature and humidity was not explained by us in the manuscript, especially regarding the new interpolation technique, which is why we wanted to add this explanation now. Examining the individual interpolated and reference profiles reveals a significant spread in the realized shapes of the gradient at the top of the boundary layer (see our more detailed description below). The reason should not be, however, a varying definition of the PBLH,

as we consistently defined this as the height of the maximum gradient, although, admittedly, the automatic detection of this height might not be correct in every single profile. A good example of this for potential temperature is shown in Figure 1 (not included in the manuscript). You can clearly see that the HN interpolation and the reference sounding have a comparable PBLH, but the shapes of the corresponding gradients differ, leading to a more variable bias at these heights compared to the rather constant potential temperature values in the mixed layer. Just as a side remark, here you can also see a good example of how much worse the structural representation of the reference profile by the non-HN interpolation is.
The answer to the question of whether the presented technique requires a separate measurement of the PBLH depends on the goal one wants to reach. We agree that one would already improve the interpolation with an interpolated PBLH, as this would result in a physically meaningful profile, i.e., a profile that does not include non-physical artifacts from e.g., averaging a mixing layer with a free atmosphere. So this could be a possible compromise or "light" version of the presented technique if no external PBLH information is available. However, it will be a compromise as this will assume a linearly developing CBL, which is usually not true. As we aimed for the greatest possible improvement, we included this extra information. However, the test of this suggested "light" version would be interesting for potential future work. We added this suggestion and our discussion of it to the manuscript.

Action: Regarding the bias correction, we added a brief description to the corresponding introductory paragraph (lines 62-64 of the revised manuscript): "The chosen PBLH product depends, of course, on the underlying instrumentation and subsequent processing, and will not be a perfect representation of the actual PBLH. To determine and correct for the relative difference in PBLH between the radiosondes and the "external" product, a bias correction is applied."
To the description of the standard deviation profile of the bias we added for the new interpolation technique (lines 348-355 of the revised manuscript), "From about 0.8 times the PBLH, the standard deviation of the bias increases drastically with height for the HN interpolation. The reason for this is the insufficient representation of *how* the gradients that characterize the interfacial layer are shaped. The HN interpolation can interpolate the information that there *is* a gradient (which defines the PBLH), but there is no information about the shape of this gradient at the time of interest. As this shape can vary greatly, e.g., the magnitude of the curvature, the onset of the gradient, or its steepness, the chance is high that the shapes of the interpolated and reference gradient are not very comparable, which translates into a broader spread in the biases at these heights. It must be stressed here that, although this may be one of the weaker aspects of the new technique, already the fact that the PBLH *is* incorporated allows for this much more meaningful interpolation."
We added "It is conceivable that a physically meaningful interpolation is already possible when the external PBLH estimate is replaced by an interpolation of the two PBLHs of the input soundings, as this would avoid the production of non-physical artifacts by averaging different atmospheric regimes. However, it can be expected that this approach can be a compromise at best, as this does not consider the non-linearity of the temporal CBL development, as it can be seen, e.g., in the Raman lidar observed humidity cross-section in Figure 1 (g)." at a suitable position in the introduction (lines 49-53 of the revised manuscript).

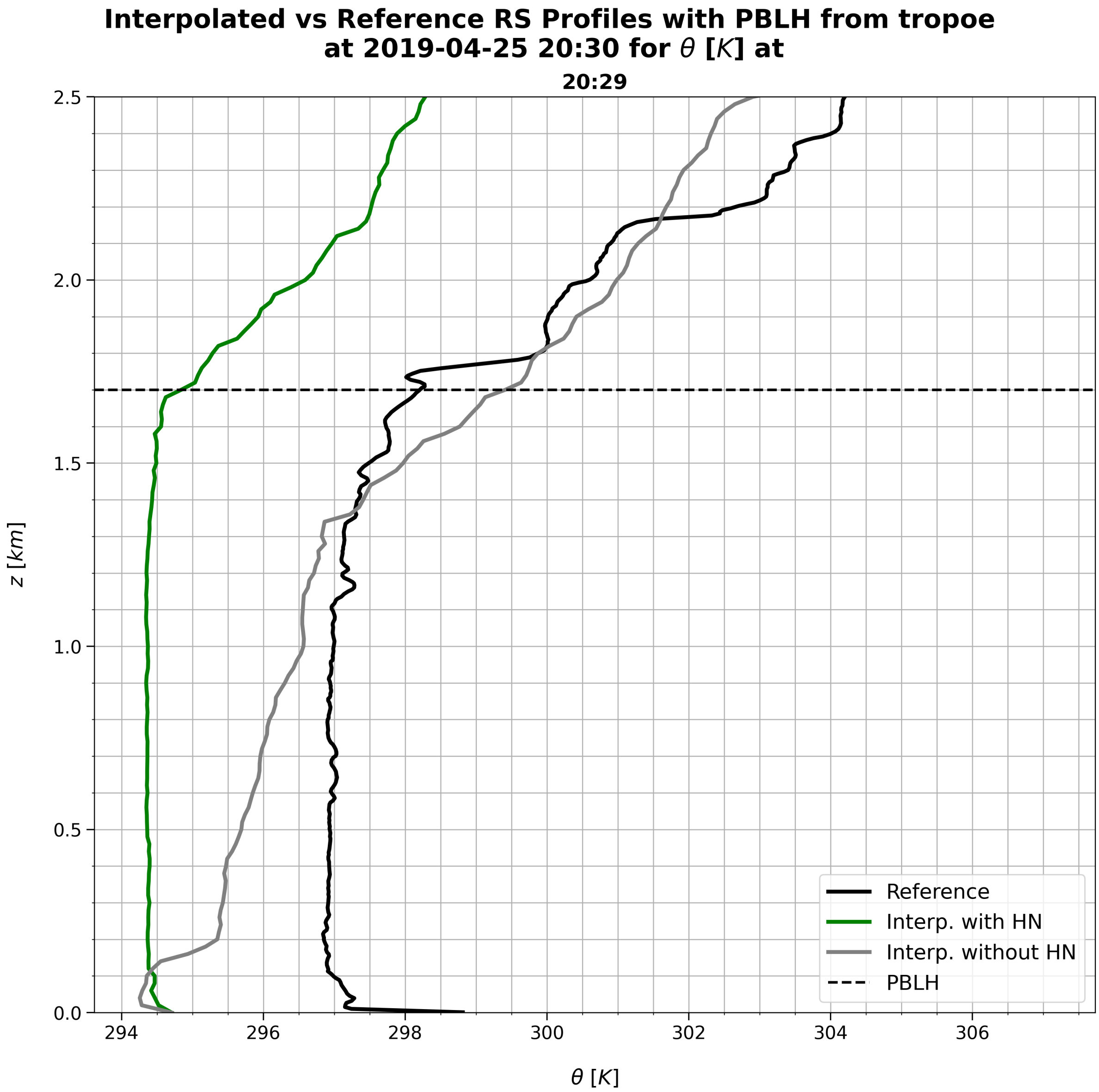

Figure 1: An example comparison between the interpolated profiles with (green) and without HN (grey) and the reference profile (black) of potential temperature on April 25, 2019, 20:30 UTC. This figure is not shown in the manuscript.

Comment: Lines 84-91: Thank you for the simple description of the TROPoe data product. It would be helpful to also include the time resolution of the product and a description of how the PBLH is determined, since it is a critical input to the new interpolation method. It would also be helpful to describe why an interpolated sounding product is needed if TROPoe already provides continuous profiles of temperature and water vapor mixing ratio. Also wonder how the TROPoe algorithm is impacted by the presence of boundary layer clouds and how this would impact the interpolation methodology.

Reply: *The TROPoe data product used here had a 5-minute resolution, and the PBLH was determined using a parcel method (i.e., by lifting a parcel at the surface along a*

constant virtual potential temperature profile until it encountered a higher value above, following the approach by Nielsen-Gammon et al., 2008).
The interpolated sounding product might be, without considering any specifics of the products, indeed redundant in the specific example that is given by our study. With available profiles from TROPoe or lidar systems, there might not be a need for this interpolation at the ARM SGP site. However, this study primarily aims to demonstrate and test this new interpolation technique, and the real value of this improved interpolation technique is only revealed in a more limited setting, considering the available instruments. We added a clarifying paragraph to the manuscript.
Regarding the influence of boundary layer clouds, or more broadly speaking, different atmospheric conditions, on TROPoe and the interpolation methodology, we first want to stress that it is not the intention of this study to get into the details of this specific PBLH product, as it is just one out of a broad variety of options (see also our reply regarding the second reviewer's comment about the PBLH estimates). Any other continuous product could have been used for this study. So we want to answer a little bit more generally here in acknowledging that any bias in the external PBLH time series will have an impact on this interpolation technique. And this impact will depend on the size and variability of the bias with different atmospheric conditions. Our attempt to counteract this bias is the employment of a bias-correction.

Action: To complement the technical description of TROPoe, we added the description above to the corresponding paragraph: "[...], while the temporal resolution of the used data product is 5 min" (lines 103-104 of the revised manuscript) and "The PBLH was determined using a parcel method (i.e., by lifting a parcel at the surface along a constant virtual potential temperature profile until it encountered a higher value above, following the approach by Nielsen-Gammon et al., 2008)." (lines 109-111 of the revised manuscript).
At the end of Section 2.2 (AERI and TROPoe), we added the paragraph "At this point, it should be made clear that this study is about the demonstration of an improved interpolation technique utilizing additional PBLH information. In *this* specific demonstration, the value of the interpolated product is questionable, as it might be more straightforward to use the temporal high-resolution (compared to single soundings) thermodynamic profiles of TROPoe or the Raman lidar directly. However, the additional value of the proposed technique lies in the fact that it *only* requires the PBLH estimate as additional information. It is therefore conceivable that in a very limited setting, the *only* chance of retrieving profiles of potential temperature and humidity is the use of the presented technique, at least for a CBL. A fictional example could be a measurement site that only employs radiosondes and a Doppler lidar from which the necessary PBLH information is extracted (e.g., Sivaraman and Zhang at the ARM SGP site)." (lines 120-127 of the revised manuscript).
Also, we added to Section 2.2, "Regardless of what specific PBLH time series is used, it will contain a bias that will vary with the atmospheric conditions. This bias will impact the presented interpolation more or less, depending on its size. An attempt to reduce this impact is the employment of a bias correction, which is explained later." (lines 116-119 of the revised manuscript).

Comment: I find Figure 2 very difficult to read and interpret as a scatter plot. Both discerning the different markers used to represent the different interpolation scheme and radiosonde separation, but also the fact that most of the

points are clustered in one are with lots of white space among the plot. Have you considered some type of contouring or 2D histogram (or something else) to make this easier to read and interpret? I do find the table communicates the differences much more clearly.

Reply: *This is a fair point and we agree that the presentation of our results would benefit from a clearer visual presentation. We therefore modified Figure 2 in two ways. First, we zoomed into the center area of the figure to reduce the wasted space at the borders (the r-axis begins now at 0.4 instead of 0 and the axis representing* $\log_2(\sigma_{interp}/\sigma_{ref})$ *covers now the interval between -1.6 and 1.6 instead of -3.2 to 3.2). This leads to a few more data points that are not shown; however, the overall message of this plot is not compromised by this in our view. Second, we added the shadows of the planes (for HN and non-HN respectively) to the plot that are span by those 85% of the corresponding data points that are closest to the ideal values of* $r = 1$ *and* $\log_2(\sigma_{interp}/\sigma_{ref}) = 0$. *This immediately demonstrates that the HN interpolation technique produces more data points closer to the ideal values.*

Action: We updated Figure 2 with the changes described above, as demonstrated also in Figure 2 of this response. Additionally, we updated the figure caption from "~~For better visualization, seven data points (five without and two with an HN) between $r = 0$ and $r = -1$ are not shown in (b).~~" to "For better visualization, twelve data points (seven with and five without an HN) are not shown in (a), as well as 16 data points (six with and ten without an HN) in (b). This results in 72 and 74 data points shown in (a) for HN and non-HN interpolations, respectively. In (b), 73 and 69 data points are shown, respectively. The shown planes are spanned by those 85% of data points that are closest to the ideal values of $r = 1$ and $\log_2(\sigma_{\mathrm{interp}}/\sigma_{\mathrm{ref}}) = 0$." and added a description for the added planes (lines 247-249 of the revised manuscript): "The shown planes (semi-transparent) visualize how close most of the data points for the HN and non-HN interpolation are to the ideal values of $r = 1$ and $\log_2(\sigma_{\mathrm{interp}}/\sigma_{\mathrm{ref}}) = 0$ respectively by being spanned by those 85% that are closest." Additionally, we strengthened our discussion of this updated figure by adding (lines 255-257 of the revised manuscript) "This is visualized especially well by the shown planes. The plane that is spanned by those 85% of data that is closest to the ideal values is obviously smaller for the HN interpolation than for the non-HN one, representing a smaller spread to less ideal values."

Comment: I think an important point to make, perhaps in the conclusions, is that this height-normalized interpolated sounding has an explicitly different purpose/target compared to the ARM interpolated sounding product. While this new technique focuses on the convective boundary layer structure only, the interpolated sounding product is designed to provide thermodynamic profiles through the depth of the troposphere under all atmospheric conditions (in near real-time). A statement about how these might be combined would be useful.

Reply: *This is a fair point, so we added a paragraph to the conclusion that explicitly states the different scopes and therefore purposes of the two interpolation techniques. A combination of both might not be trivial, as the HN interpolation is only tested for the CBL, limiting it temporally to the CBL situation and spatially to the boundary layer itself,*

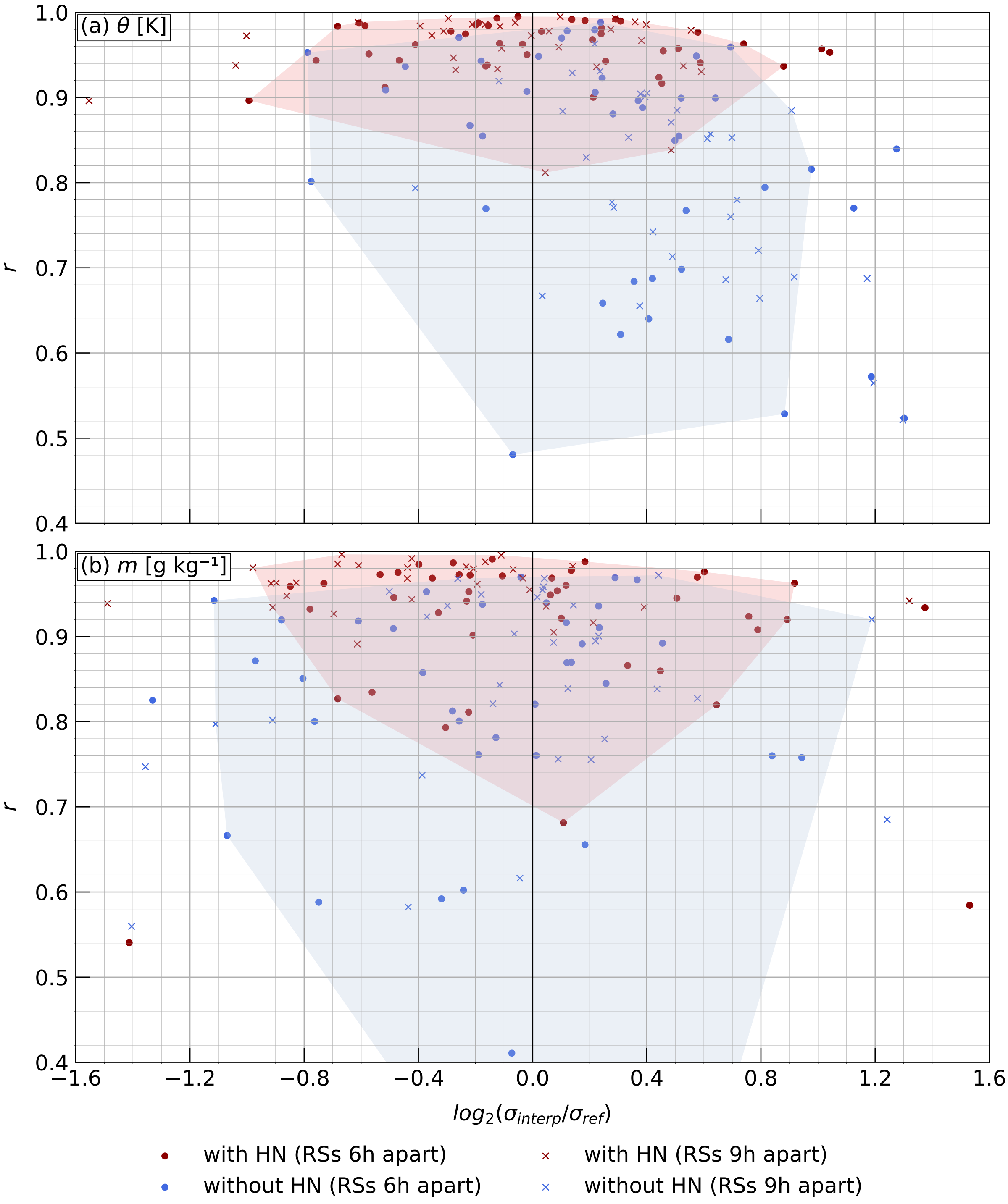

Figure 2: Improved version of Figure 2 from the Preprint. This updated version covers only a zoomed-in area and includes the planes that represent those 85% of data that are closest to the ideal values of $r = 1$ and $\log_2(\sigma_{\text{interp}}/\sigma_{\text{ref}}) = 0$.

as the HN technique was not tested for the free atmosphere.

Action: In the conclusion, we added (lines 376-382 of the revised manuscript) "Because the HN interpolation additionally requires the PBLH and is presently applicable only to CBL conditions as well as the boundary layer region, it is important to clarify its intended purpose. The routine, non-HN interpolation delivers thermodynamic profiles for the entire troposphere under all atmospheric regimes and therefore remains the best choice when continuous temporal and spatial coverage is required. By contrast, when the goal is to obtain the most accurate representation of thermodynamic conditions within the boundary layer during a CBL situation, the HN interpolation can provide a substantial improvement. It is not meant to replace the standard interpolation product, but rather to complement it for the specific use cases of boundary layer analysis."

2.2 Specific Comments

Comment: Line 1-2: Suggest adding that this improved technique is for the planetary boundary layer.

Reply: *Agreed.*

Action: We added "[...]in the planetary boundary layer[...]" to the first sentence of our abstract (line 2 of the revised manuscript).

Comment: Line 2: Not sure there is a need to define or even use acronyms in the abstract, especially if they are not used again within the abstract.

Reply: *Agreed. This disturbs the flow of reading through the abstract, and they are all defined later anyway.*

Action: Any acronyms defined in the abstract are removed.

Comment: Line 4 (and 17, 56): Minor detail but ARM is a "Facility" rather than a "Program."

Reply: *True, we corrected this.*

Action: We replaced "~~Program~~" with "Facility" (lines 4, 17, 75, and 359 of the revised manuscript).

Comment: Line 16: The frequency of launches can vary much more than between 2 and 4 times per day depending on many different parameters. You might say that at many operational sites radiosondes are routinely launched two to four times per day or something similar.

Reply: *We changed the corresponding introductory paragraph.*

Action: "~~This dependency on RSs has drawbacks, as the sounding locations are relatively scarce (Wulfmeyer et al., 2015), the soundings are not vertical, and the launches occur routinely only two to four times per day, even at larger facilities such as the Atmospheric Radiation Measurement Program (ARM) Southern Great Plains (SGP) site considered in~~

~~this work (Jensen et al., 2016).~~" is replaced by "This dependency on RSs has drawbacks, as the sounding locations are relatively scarce (Wulfmeyer et al., 2015), the soundings are not vertical, and the launches occur routinely only two to four times per day at many operational sites, such as the ones operated by the Atmospheric Radiation Measurement (ARM) Facility (Jensen et al., 2016), which also incorporates the Southern Great Plains (SGP) site considered in this study" (lines 15-18 of the revised manuscript).

Comment: Line 30: Normalizing by the height of the PBL to understand PBL structure has a long history and it should probably be acknowledged here. One of the earliest studies I know of that used this technique was: Augstein,E., H. Schmidt,and F. Ostapoff, The vertical structure of the atmospheric planetary boundary layer in undisturbed trade winds over the Atlantic Ocean, Boundary Layer Meteorol., 6, 129-150,1974.

Reply: *Fair point, we added this as a side remark.*

Action: We added "Normalizing the boundary layer with the PBLH to investigate the lower atmosphere is generally not a novel approach; an early example of this technique can be found in Augstein et al., 1974." to the corresponding introductory paragraph (lines 36-38 of the revised manuscript).

Comment: Line 32: Replace "level" with "value." Level is ambiguous because it could refer to the level (i.e. height) in the atmosphere.

Reply: *Agreed.*

Action: "~~level~~" was replaced by "value" (line 40 of the revised manuscript).

Comment: Line 41: Suggest adding an "e.g." ahead of the list of references since there are many, many studies in this subject area.

Reply: *Agreed.*

Action: We added "e.g., ahead of the two listed references (line 59 of the revised manuscript).

Comment: Line 56: ARM and SGP were already defined.

Reply: *True.*

Action: We replaced the definitions with the acronyms (line 74 of the revised manuscript).

Comment: Line 108: "in-between' seems too colloquial. Maybe "intermediate" would be a more appropriate word?

Reply: *Agreed.*

Action: "~~in-between~~" was replaced by "intermediate" (line 143 of the revised manuscript).

Comment: Line 109: "an" should be "any"

Reply: *True.*

Action: "~~an~~" was replaced by "any" (line 144 of the revised manuscript).

Comment: Line 110: Rather than referring to this as the "standard procedure of the ARM program," I would suggest using "standard procedure used in the interpolated sonde product."

Reply: *Agreed, this sounds more specific.*

Action: "~~standard procedure of the ARM program~~" was replaced by "standard procedure used in the interpolated sonde product by ARM" (line 145 of the revised manuscript).

Comment: Line 110-111: May be better stated as ". . . .and second, using the normalization of the height coordinate using the smooth PBLH estimate."

Reply: *True, this sounds indeed better.*

Action: "~~[...], and second, the one with such a normalization using the smooth PBLH estimate.~~" was replaced by "[...]. And second, the newly proposed interpolation technique that utilizes a smooth PBLH estimate to normalize the height coordinate." (line 145-146 of the revised manuscript).

Comment: Line 117: Can you describe "the characteristics indicative of a convective boundary layer?" Does this include thresholds for the stability? Does it include decoupled cases?

Reply: *With these characteristics, we do not mean anything more than the ones mentioned in the introduction (lines 31-34 of the Preprint), i.e., the overall layering of the CBL and the shapes of the potential temperature and humidity profiles. To be more concrete, to "accept" a sounding for our analysis, we wanted to see a relatively constant potential temperature in the mixed layer and a rather pronounced gradient, signifying the interfacial layer. We did this "by eye" and did not employ any thresholds. Regarding the inclusion of any decoupled cases, please refer to our response to the comment regarding lines 34-36 of the Preprint.*

Action: We added the half sentence "[...], i.e., relatively constant values in the mixed layer and a pronounced gradient on top, signifying the interfacial layer." (lines 152-153 of the revised manuscript).

Comment: Line 122: I am not sure “normal” or “old” are the appropriate here. “Current” seems like it would better here and elsewhere (but “old is better than “normal”). Also, see comment regarding Line 110.

Reply: *Agreed.*

Action: We replaced statements such as ”old” or ”normal” with ”current” or ”non-HN” throughout the manuscript. Especially, we replaced ”~~[...]’normal’ or ’old’[...]~~” with ”[...]’current’ or ’non-HN’[...]” (line 158 of the revised manuscript).

Comment: Line 138: See general comment #3. Are you requiring the CBL be well-mixed throughout its depth?

Reply: *Yes, we require this in the sense that we only selected cases that showed a relatively constant (and therefore well-mixed) potential temperature value between the surface and the inversion layer. See also our replies to the above comments regarding lines 34-36 and line 117 of the Preprint.*

Action: In our opinion, this question was addressed with the replies to the comments mentioned above.

Comment: Line 148-149: Should we expect the TROPoe-derived and the Raman Lidar derived PBLH to agree? Different measurement and retrieval methods, and different resolutions will make the agreement difficult.

Reply: *We should not expect an agreement due to the reasons mentioned by you. But as the Raman lidar has a much higher vertical resolution compared to the AERI-TROPoe combination (active versus passive remote sensing), we can expect the Raman lidar to describe reality better, meaning in this study, the actual PBLH. This non-agreement is the motivation for us to employ the mentioned bias-correction.*

Action: For better clarity, we added the half-sentence ”[...], which is therefore expected to describe the real PBLH the best, [...]” (lines 185-186 of the revised manuscript).

Comment: Line 195: “constant mixed layer and a sharp gradient at the top” is not clear. I think you are referring to potential temperature or water vapor mixing ratio being constant with height within the mixed layer with a sharp gradient at the top.”

Reply: *True.*

Action: For improved clarity, we replaced ”~~[...] with a constant mixed layer and a sharp gradient at the top, [...]~~” with ”[...] with a mixed layer that is signified by nearly constant values of potential temperature and WVMR as well as a sharp gradient over the interfacial layer for those quantities, [...]" (lines 238-239 of the revised manuscript).

Comment: Line 203: I would not refer to these as distributions, rather just scatterplots.

Reply: *Fair point.*

Action: "~~distributions~~" was replaced by "scatter plots" (line 245 and 250 of the revised manuscript).

Comment: Line 215: What is meant by "introduction of nonphysical atmospheric layers as artifacts of insufficient interpolation?" This is unclear to me.

Reply: *Fair point. We hope to make this clearer by giving a concrete example. The averaging of non-HN profiles may lead to the situation where, at a certain height, the warmer (in terms of potential temperature) free atmosphere of the earlier profile (the CBL is not that far developed earlier in the day) is averaged with the cooler mixed layer of the later profile (as the CBL developed during the past time and now has a mixed layer at the height of interest). The average profile will be warmer than the earlier and cooler than the latter input sounding at this height, so it will appear as if there is an additional layer (when examining this interpolated profile solely). However, this is a non-physical artifact resulting from the insufficient interpolation. We present a visualization of this in Figure 3 (not shown in the manuscript).*

Action: To make our analysis clearer, we incorporated a brief additional description (lines 262-265 of the revised manuscript): "For example, averaging an early-day potential temperature sounding with a lower PBLH together with a later-day sounding that has a higher PBLH (using the non-HN technique) will produce an apparent extra 'layer' at heights where the warmer free atmosphere from the first sounding is averaged with the cooler mixed layer from the second."

Comment: Line 216-217: Does this mean you do sometimes have decoupled layers that could cause problems for the height normalization method?

Reply: *This point should be clearer now with our explanations regarding the comment to line 215 explaining what we meant by "additional layers" as well as the comment to lines 34-36 where we discuss the decoupled case.*

Action: See actions regarding the above mentioned comments.

Comment: Line 257: "level" is ambiguous here, suggest "value instead." Level could be height in the atmosphere.

Reply: *Agreed.*

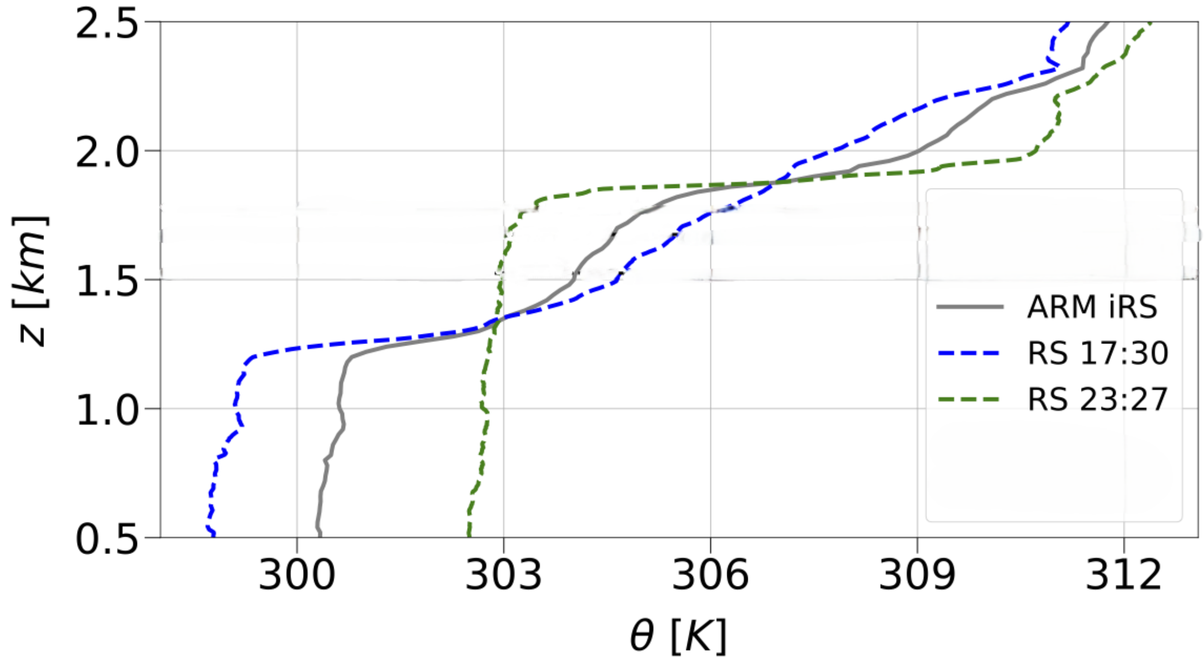

Figure 3: Radiosonde (RS) profiles of potential temperature from 30 August 2017 at 17:30 (blue) and 23:27 UTC (green). The current non-HN interpolation of these profiles to 20:00 UTC (grey, ARM iRS) clearly shows two "inversion layers" at the inversion heights of the input soundings. These are the above-mentioned nonphysical artifacts that can arise for a non-HN interpolation. This figure and the shown soundings are not part of the manuscript.

Action: "~~level~~" is replaced by "value" (line 306 of the revised manuscript).

Comment: Line 280: Can you explain what is meant by "non-physical artifact layers?" Perhaps provide an example?

Reply: *Please look at our reply regarding the comment to line 215, which also gives an example.*

Action: See actions regarding the comment above to line 215.

Comment: Line 304: Can you define what is meant by "suitable source?" For example, since you are interpolating thermodynamic variables, should the PBLH be defined from similar thermodynamic variables? i.e. will a lidar-based retrieval that depends on aerosol scattering, or a a retrieval based on turbulence profiles raise some issues with the height normalization?

Reply: *With "suitable source", we really just meant any PBLH product that provides a temporally continuous estimate of the PBLH. However, this is a good question that we cannot answer satisfactorily in the scope of this study. Different methods will definitely yield different PBLH estimates and contain different method-specific pitfalls, e.g., in the* *case of the mentioned aerosol-based PBLH estimates, there is the risk of confusing the PBL by detecting aerosol layers above it. See, e.g., Kotthaus et al., 2023, for a good summary of different methods and their limitations. And although it sounds reasonable that the PBLH should be defined from the same thermodynamic variables as the variable of interest for consistency in the PBLH definitions, this might be impossible depending on the setting, as those variables might be explicitly not available for a PBLH estimate; otherwise, there might be no reason to use the interpolation in the first place.*

Action: For improved clarity, we replaced "~~[...]from any suitable source.~~" with "[...]from any source providing a temporally continuous PBLH estimate." (line 362 of the revised manuscript).

Comment: Line 325: Are the height normalized interpolated soundings available somewhere?

Reply: *No, unfortunately not, as our processing routines only save the resulting metrics (correlation coefficients, ratios of the standard deviations, and biases). A change of the routines to also save the interpolated profiles would be possible, but be an intensive effort.*

Action: None.

3 Comments by Referee #2

3.1 General, Scientific, and Technical Comments

Comment: This manuscript presents a new method for interpolating temperature and water vapor mixing ratio (WVMR) profiles between radiosondes. The method uses a height normalized grid interpolation approach to more accurately represent the structure of the boundary layer throughout the day and requires continuous planetary boundary layer height (PBLH) measurements as input to the algorithm.
The normalized height (NH) approach was applied to convective boundary layer cases and compared qualitatively with Raman lidar based measurements of potential temperature and WVMR. The approach shows improvement to the precipitation free boundary layer though with a bias. The mean bias and standard deviation near the surface are very similar for the NH and non-NH methods.
Overall, the paper is well organized, and the writing is clear, though some clarification of the analysis could be improved and are noted below. The NH method demonstrated does provide value in this well constrained example, but the broader applicability in a continuous product is a much farther reach. Despite this I am inclined to accept this manuscript for publication with minor revisions because it is the first demonstrated improvement to interpolated radiosondes that could be valuable for specific applications and case studies. I encourage the authors to pursue further evaluations under more varied conditions, as discussed in the conclusions, so that a more universal application

can be realized.

Reply: *Again, we agree that some improvements to our descriptions are necessary for a better understanding of our results. We also agree that this study only treats a well-constrained example. In the context of our broader work and motivation (see our reply to the first reviewer's comment regarding lines 21-27), a more general study was out of our scope.*

Action: Any actions are taken in response to individual comments. In case a specific point was already addressed in the context of the first review, we refer to our corresponding reply.

Comment: Introduction and Motivation: The motivation for this study is to explore ways to improve the interpolation of thermodynamic profiles between radiosonde measurements, which has many positive benefits. This study is focused on thermodynamic structure in the hydrometeor free boundary layer (no fog, precipitation etc.) but much of the discussion in the introduction is around the ARM routine interpolated sonde data and its use in cloud microphysical retrieval products, where the interpolated radiosonde measurements are only used in cloud above the boundary layer. They also demonstrate in their analysis that the NH and non-NH approaches both converge above the boundary layer. This discussion is essentially irrelevant to the paper. I suggest removing this discussion (lines 19-27) or revising it to be more focused on the improvements in the boundary layer and describing applications where an improved product would be beneficial to the community.
I would also note that it is very likely that many more research groups besides ARM use simple interpolations between radiosondes for a variety of applications, and you don't use the ARM product in the analysis. You may want to make a note of this in the introduction and focus on the different methodologies instead.

Reply: *It is true that our attempt to motivate the presented interpolation technique via the MICROBASE product is not expedient. We therefore replaced the corresponding paragraph with a more concrete example of how this new product could be used by describing our own motivation behind this study.*
Regarding the second point, we agree that we should clarify that we did not use the ARM interpolation explicitly in our analysis, but rather a "simulation" of it.

Action: For the improved motivation, we refer to our reply to the first reviewer's comment regarding lines 21-27.
To clarify our non-use of the ARM interpolation in our analysis, we added "The mentioned ARM product, as the currently available RS interpolation product, is the benchmark for this proposed improved technique. However, the actual ARM product is not used explicitly in this work, but rather a 'simulation' of it for yet-to-be-explained reasons." to the corresponding introductory paragraph (lines 31-33 of the revised manuscript).

Comment: PBLH estimates: As the authors state, the PBLH is an input to

the NH method, and any measurement could be used. But the accuracy of the PBLH measurement is critical to this approach. Deriving a continuous PBLH product over all conditions is not trivial, requires multiple measurement methods, and errors would cause discontinuities in the NH interpolated radiosonde data. The authors should acknowledge this in Sec. 2.2 or the conclusions.
Line 43-45: Can you elaborate on what simple instruments could be used to determine continuous PBLH? The authors use AERI measurements to derive continuous PBLH. The AERI instrument and associated TROPoe algorithm are both complex and expensive.

Reply: *Thanks for the remark, we agree and added a side remark, acknowledging this non-triviality of deriving a continuous PBLH estimate.*
Regarding our use of the term "simple instruments", we must admit that this formulation was perhaps too sloppy and oversimplistic. Initially, we meant with this statement that it is not necessary to employ active remote sensing instruments, although this does not correlate with complexity or cost. Therefore, we think that the quality of the corresponding description benefits from replacing this formulation with a more neutral one. This corresponds then also well to the admittance of the non-triviality of retrieving PBLH estimates.

Action: We added "To avoid any misunderstanding, deriving a continuous estimate of the PBLH is far from trivial and constitutes its own field of research. Numerous methods exist, often differing significantly in their results (for RSs, e.g., Seidel et al., 2010), and their performance depends on the atmospheric conditions and underlying instrumentation (e.g., Zhang et al., 2022; Kotthaus et al., 2023)." to the end of Section 2.2 (lines 128-131 of the revised manuscript).
The misleading sentence "~~Fortunately, this can be derived from comparably simple instruments, making it possible to acquire a relatively precise picture of the boundary layer using only basic equipment. It may not be feasible for smaller measurement sites, nor necessary, depending on the requirements, to employ complex, active remote sensing systems, which also depend on the weather situation (Wulfmeyer et al., 2015).~~" is replaced by the more neutral statement "Fortunately, and as briefly discussed later, there are different available techniques utilizing different types of instruments, increasing the possibility that even smaller measurement sites can produce such an estimate." (lines 60-62 of the revised manuscript).

Comment: Figure 2: For the discussion regarding Figure 2 you are using cases and profiles interchangeably. I interpret a "case" as being a convective boundary layer event and a profile as representing a 10 min interpolated vertical profile corresponding to the 10 min average Raman lidar profile (based on earlier discussion related to Fig. 1). Please clarify what the 79 profiles represent, how many profiles are compared per event (would it be 2?) and state how many profiles are represented in Fig. 2a and 2b for the HN and non-HN profiles.
The scatter plot in Figure 2 as presented is hard to interpret, though Table 1 provides quantitative values to support the analysis. A joint PDF might be easier to visualize the differences, though there may not be enough points.

Reply: *First, we indeed used the terms "cases" and "profiles" interchangeably in different contexts, which is neither precise nor consistent. To make our descriptions exactly this, we adjusted our expression and defined a "case" as the individual comparison between an interpolated profile and the corresponding reference profile. This means that there are 79 cases per interpolation technique. Before, we also used "case" to describe an individual sounding selected for our analysis, whether for input or reference, but now any other context is described properly. In the context of Figure 2, "cases" is therefore a synonym for "data point".*
Second, we also updated the description of Figure 2 to state the number of shown cases (now used after the new definition).
Finally, regarding the readability and interpretability of Figure 2, we agree that the previous presentation was convoluted in some parts and wasted space in others. As already stated in our reply to the comparable critique of the first review, we improved Figure 2 by zooming more into the relevant part and adding shadows of the planes that are span by those 85% of cases/data points that are closest to the ideal values of $r = 1$ *and* $\log_2(\sigma_{interp}/\sigma_{ref}) = 0$. *For more details, please refer to the reply above.*

Action: See our actions regarding Figure 2 in the corresponding section for the first review.
"Each comparison between two profiles is called a 'case' from now on, meaning that there are 79 cases of either interpolation technique that can be analyzed in terms of the mentioned metrics." was added to the introductory paragraph of the Results Section (Section 4) to define the term "case" in the mentioned way (lines 199-201 of the revised manuscript). Additionally, we changed "~~case~~" or "~~cases~~" to "soundings during CBL conditions", "soundings", "soundings", "interpolated soundings in comparison to their reference profile", "pairs of interpolated and reference profiles", "favorable example", "procedure", "situations", "Potential temperature profiles" (lines 70, 144, 151, 181, 192, 215, 220, 224, and 322 of the revised manuscript) to fit our more consistent definitions.
The caption of Figure 2 now includes "This results in 72 and 74 data points shown in (a) for HN and non-HN interpolations, respectively. In (b), 73 and 69 data points are shown, respectively."

3.2 Minor Revisions

Comment: Line 26: The Microbase algorithm uses the interpolated radiosonde data and radar reflectivity to determine the cloud phase. The interpolated radiosonde data is not used to determine the radar reflectivity. Please correct this inconsistency in the text if this sentence is not removed (per the discussion above under regarding the introduction).

Reply: *Thanks for the clarification. However, we decided to completely replace the paragraph concerning MICROBASE.*

Action: For the improved motivation, we refer to our reply to the first reviewer's comment regarding lines 21-27.

Comment: Lines 122, 313 throughout: Suggest referring to the method without height normalization as non-HN rather than 'old' or 'normal' as it is more descriptive.

Reply: *Agreed.*

Action: We changed this to "current" or "non-HN" throughout the manuscript. See also our reply to the first reviewer's comment regarding line 122.

Comment: Sec. 4.2.1 You may want to remind readers that the analysis in this section compares the interpolated sondes (both methods) with the radiosonde launches that were not used in the interpolation. This was described in Sec. 3.2 but did not point to which comparison uses the sondes. Anyhow, breadcrumbs are always nice.

Reply: *We agree that the reader should be reminded of the specifics of the conducted comparisons. However, we think that this was already taken care of in the first introductory paragraph of Section 4.2 as a reminder ahead of the whole analysis. Anyhow, to avoid any misunderstandings, we mention this now more explicitly.*

Action: We added "To be clear, the interpolated sondes of both methods are compared to soundings that are not used in the corresponding interpolations." to the introductory paragraph of Section 4.2 (lines 222-223 of the revised manuscript).