Atmospheric Chemistry and Physics Discussions

Atmos. Chem. Phys. Discuss., referee comment RC1
https://doi.org/10.5194/acp-2021-668-RC1, 2021

Comment on acp-2021-668

Anonymous Referee #1

Referee comment on "Atmospheric stratification over Namibia and the southeast Atlantic Ocean" by Danitza Klopper et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2021-668-RC1, 2021

This paper presents an analysis of atmospheric profiles measured over the Southeast Atlantic Ocean derived from GPS-RO satellite retrievals and radiosonde profiles, and ends with a discussion of the broader atmospheric context over the study region. However, the paper is confusingly structured and unclear in many parts. Specifically, 1) the manuscript's structure jumps back and forth between the different components (satellite, radiosonde, different BLH definitions) in a way that makes it very difficult to follow the methodology; 2) it is not clear what is the connection between the satellite and the radiosonde portions of the analysis, or indeed between the different BLH definitions just using the radiosonde data (difficult to follow the analysis); and ultimately 3) the intended scientific focus of this work is thus quite difficult to determine. I hope that the authors will address the below comments before the paper is considered for final publication.

1) The paper begins with a discussion of the COSMIC satellite data and its processing algorithm. For those not familiar with the COSMIC satellite data, Section 3 is rather difficult to follow regarding which processing (Abel inversion algorithm, "atmPrf" dataset, ECMWF "1-D var" moisture correction?) is provided or what is additional processing/analysis done by the authors e.g. following the Shyam reference?

Sections 3 and 4 seem disjointed, as, for example, Section 3.1 talks about the data processing for COSMIC and Section 4.1 also talks about COSMIC data processing and definitions, and Sections 3.2 and 4.2 both discuss radiosondes. As I said above, it's not really clear in Section 3.1 which processing was done by the authors in the present work, and which was in an external dataset (as the references seem incomplete). Would it be better to combine the two COSMIC GPS-RO sections, and the two radiosonde processing sections into one Data/Methods section, followed by analysis of the results?

- Also, the paragraph starting on Line 134 seems to belong more in the background/introduction sections.

- Section 5.1.1: The conclusions section states that the GPS-RO method underestimated temperatures in the temperature profiles, but that doesn't seem to be strongly supported by Section 5.1.1 (an absolute error of -0.3+/- 1.3C seems fairly evenly distributed between positive vs negative differences) or Figure 2, which shows pretty good agreement between the two methods at least as it is presented there (see below "other comment" about Fig 2 as well). I'd recommend finding a different visualization if the point of this figure is to say that one method systematically underestimates temperature.
- The authors mention it multiple times, but it's not clear how the issue of superrefractivity might be affecting the analysis. Line 100: "even with the applied corrections, no reliable information about atmospheric structure can be collected below where the signal is super-refracted": where is this point, typically, and how frequently do these conditions occur in the region of interest? Then Line 280 states the inversion heights were on average 190+/-480m lower than MG heights, "beyond these layers, the profile will likely be superrefracted" but doesn't 190+/- 480m indicate a sizable fraction of cases where the MG of refractivity is lower than the inversion height? Are these cases included in Table 1?

- Then in Section 4 the authors mention three or four separate methods to calculate BLH only from the radiosondes, but the refractivity definition compared with the other three is never really explored. It's mentioned briefly on Line 494 that it isn't consistent with the RN definition, but by then we're already in the conclusions. To my mind this needs to be addressed far before that because the analysis of Section 5 uses both refractivity definitions and inversion definitions, so those need to be reconciled. What are the considerations of each calculation? What are we supposed to take away from these different definitions beyond "the BLH can vary rather widely based on what definition of BLH you use" (this issue of definition was also discussed somewhat in this same special issue by Ryoo et al, https://doi.org/10.5194/acp-2021-274).

2) The value of the refractivity definition is understandable in that it allows a direct comparison to the satellite-based retrieval (see above), but Figure 3 shows it doesn't do a particularly good job in that respect, and Figure 4 shows that the other three definitions aren't consistent with one another either. So what's the use of any of this? And the somewhat arbitrary throwing out of 6 points in Fig 3 doesn't lend any confidence to the time series of these parameters in Fig 6 either.

- I'd first suggest a clear delineation of each BLH calculation description, either as a bulleted list or maybe even a table. Also, come up with clear names/abbreviations for each of the BLH definitions (e.g., one of them is described as "the point where the virtual potential temperature (VPT) aloft is the same as at the surface," multiple times, when you could just call it BLH_VPT or something similar after Section 4). It's difficult to keep all of those straight. And Line 141 says "the point of MG of refractivity (hereafter MG height)" but later in the text uses "MG height" and "the height of the MG of refractivity" (Fig 7) and "the height of MG N-refractivity profiles" (Line 269) etc… are these all the same thing? And "low-level inversions" and "surface-based inversions" are the same? It's quite hard to follow.
- Section 5.1.2: I see the value in comparing the refractivity BLH calculations for GPS-RO

and radiosondes, but Figure 3 doesn't seem to support that these are comparable. In this section the authors eliminate several potential explanations for the poor agreement, but then exclude the worst-comparing points based on nothing other than they are the worst-comparing points. What's to say that the majority of points in Fig 6 don't show that same discrepancy, then? How can these really be compared?

- Then throughout Section 5.2 and 5.3, the analysis jumps back and forth between GPS-RO and radiosonde analysis, under the headings of "Spatial and temporal variability" although really only the GPS-RO data can give spatial variability here, right? Given the results of the earlier sections, it seems to me the takeaway is that they aren't really interchangeable, although the structure of Tables 1/2 and Tables 3/4 make it rather difficult to compare the results from the two methods.
- Confusingly, Sections 5.2.3 and 5.3.1 are both titled "temperature inversions" but refer to either low-level or mid-level temperature inversions. It's not clear how spatial plots of low-level temperature inversions are derived from only the GPS-RO data, given the superrefractivity questions above and the clear altitudinal limitations of these data as shown in Fig 2, especially relative to the radiosonde-based inversion height in Fig 4.
- I'd also move Fig 6 up to significantly earlier in the paper, e.g. just after Fig 3, as they are showing similar things and the context for how GPS-RO and radiosondes compare with one another is a necessary prerequisite before talking about the spatial and temporal patterns in their results.

3) Finally, it's difficult to see how these observations (which I think are worth describing if the above issues can be addressed) fit into the broader meteorological picture, which I think is what the authors are trying to do in Section 6. These connections are tenuous at best. Most of the spatial maps presented (Figs 7-16) primarily show seasonally-averaged values and then standard deviations (or try to; the stdev figures are extremely hard to interpret, see additional comments below), so it's not clear to me how this relates to transient meteorological events e.g. as discussed in Sections 6.1 and 6.2. Section 6.3 discusses cloud fraction but the spatial analysis will rely on the GPS-RO profiles which are less reliable in cloudy conditions, is that right? How is this addressed? The authors mention MERRA-2 (and also mention MODIS in the "data availability" section but apparently nowhere else in the paper?), I can't help but think that a comparison of the profiles here with a large-scale reanalysis or model that gives atmospheric motion (MERRA-2 or perhaps ERA5 which performs better in the region; see Ryoo et al., 2021 https://doi.org/10.5194/acp-2021-274 or Pistone et al., 2021 https://doi.org/10.5194/acp-21-9643-2021) is necessary if the goal of this work is to contextualize these boundary layer height variations within the larger context of the regional atmospheric circulation.

Perhaps some of the above comments can be easily addressed simply by restructuring the manuscript, but I spent more time than I'd like to admit right now trying to understand it and these are the major questions I still had. I'd suggest the authors decide what they hope to convey with this manuscript as they revise their work, and I have hope this will lead to a greatly improved paper.

Other comments:

- It’s not very clear to me what you’re trying to convey with Figure 2. If the focus is on the BLH difference between the two datasets, then why show the full altitude scale up to 10.5km? It’s very difficult to see what the differences are between ~2-5km. On the other hand, if the point is to show that the lapse rate is generally in agreement, I don’t think you need 36 panels to do that (also, I’d suggest making the lower line thicker, it’s really difficult to see the underlying blue line there). These aren’t every single coincident profile, correct? How many are in the circle shown in Figure 1? Lines 112-113 indicated 4007 within the coastal region, were only 32 comparable? Also, why so many more valid retrievals the ocean? Based on just the surface area shown in in Fig 1, I’d expect maybe 2-3 times as many profiles over the ocean versus the continent… is there a further difference in what makes valid retrievals for land vs ocean beyond just the atmospheric moisture? I’d mention that.
- How many retrievals are going into Fig 4? Are there systematic differences in May vs June re: the 10am vs 9am launch time? Or could the sharp increase in the RN BLH range between those two be due to diurnal BLH development, or fewer radiosondes being launched in May 2015 vs April or June?
- Figure 5: I’d recommend a different color scheme especially for panel A; wind direction of 0 = 360 degrees, so having one be red and the other being blue is difficult to interpret e.g. northerly from easterly (also I’d recommend adding to the caption 0=east for ease of reading, assuming that’s the convention being used here).
- The 2x8 figures are overall very difficult to interpret; it’s not at all clear what is the main message of each of these very similar figures. Beyond that, the color scale on Figs 7b, 8b, 9b, 11b, 14b, and 16b makes it extremely difficult to interpret, beyond “they’re all small”. If that’s the message, you can lose all these panels altogether. If it's not, then a different scale should be used to show the variations between different panels. And are the black contours the same parameters? At what interval are those lines? And how much data is included in these figures (how many overpasses; is this also limited to mid-morning or is this all times of day; are the retrievals regularly distributed in time and space or are there particularly retrieval-rich times or overpasses which could bias the results preferentially towards a certain time or condition)?
- Relatedly, it’s not clear what the authors are intending to convey with the standard deviations throughout the paper, especially when the ranges are much larger than the mean values themselves. For example, how can you have an inversion depth of 200m +/- 300m or a temperature inversion strength of 0.55 +/- 0.56 C/depth [depth = 60 +/- 40 meters?]? Isn’t that saying a nonnegligible fraction of the data would have zero-to-imaginary temperature inversions? I think another metric of variability might be more instructive, either in terms of percentiles or just showing the frequency distributions of select parameters.
- Also, Figures 6-11 (except 10) show heights above mean sea level in a region where ground level is ~1-2km (e.g. Fig 1)? I’m not sure amsl is the most instructive height metric here. If “mean MG heights were consistently higher over land” (Line 265), were they relatively higher compared to magl?
- Figure 8: Why show the seasonal variability of the refractivity BLH in Fig 6, but show the diurnal variability this way? For all the discussion about diurnal variability (e.g. also Section 6.3), I think this would be better served as a time series. As with other figures,

I'd like to know how many retrievals go into each of these.

- Section 6.4: if it is decided to keep in the larger discussion, is there any indication regarding whether the presence of aerosols would affect the validity of the GPS-RO profiles, as humidity/clouds do?
- Is "subcontinent" (Line 38) a common term to refer to this part of Africa? It seems to be more just the southern continent-proper.
- Lines 120-125: this is a bit confusing. Was the primary set of radiosondes always at 10am local, with an additional set between 10 and 11? Also, why say you're converting time to UTC, and then describe the dataset in local time?
- Line 199: isn't very low vapor pressure = very dry conditions, not a moist atmosphere?