

## ***Interactive comment on “Interpolation uncertainty of atmospheric temperature radiosoundings” by Alessandro Fassò et al.***

### **Anonymous Referee #1**

Received and published: 28 July 2020

#### General comments:

This paper addresses an important question that researchers come across when aiming to estimate the uncertainty on interpolated profile data. The paper is a valuable contribution to Atmospheric Measurements Techniques and I would recommend the publication of this manuscript after a minor revision, which should include some clarifications and improvements to the figures. It is clear that much effort has been put into this work and I'm sure it will be a valuable contribution to the atmospheric research community.

In the following, I will answer the questions posed on the AMT website to be answered in the review and then provide specific comments and technical corrections. 1. Does the paper address relevant scientific questions within the scope of AMT? Yes.

C1

2. Does the paper present novel concepts, ideas, tools, or data? Yes.
3. Are substantial conclusions reached? Yes.
4. Are the scientific methods and assumptions valid and clearly outlined? I am not a statistician and may therefore not be qualified if all assumptions and equations are valid. However, to the best of my knowledge, the methods seem to be valid. I would describe the explanation suitable for an expert user and could mostly follow the steps.
5. Are the results sufficient to support the interpretations and conclusions? Yes.
6. Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)? Without trying this myself, I don't feel in the position to confidently answer this question.
7. Do the authors give proper credit to related work and clearly indicate their own new/original contribution? Yes.
8. Does the title clearly reflect the contents of the paper? Yes, however, I will suggest an improvement below.
9. Does the abstract provide a concise and complete summary? Yes.
10. Is the overall presentation well structured and clear? Yes.
11. Is the language fluent and precise? Yes, mainly. I have suggested some improvements below.
12. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used? I suggested some changes to the units and abbreviations that have not been introduced, but overall the standard is good.
13. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated? Tables and Figures: Your table and figure captions are rather scarce and need to be improved to ensure the reader does not need to guess what has

C2

been shown. I made some specific comments for individual plots, but please also have another careful look over them. I also suggested in several instances to use common axis limits to aid the reader in understanding the plots.

14. Are the number and quality of references appropriate? Yes.

15. Is the amount and quality of supplementary material appropriate? Yes.

Specific comments:

Title: I wonder if the title could be improved, e.g. 'Interpolation uncertainty of atmospheric temperature profiles' or 'Interpolation uncertainty of atmospheric temperature profiles measured with radiosondes'. The reason for this suggestion is, that I would not search for the word 'radiosoundings' but rather for radiosonde or for profile.

Line 18: I would argue it is not just the ULTS, but the entire troposphere and stratosphere

Line 22: Also, here I would argue that GRUAN provides reference observations from the surface, through the troposphere, and into the stratosphere (rather than in the UTLS)

Line 30: Please rewrite sentence starting on this line e.g. "If interpolation is applied to fill in the missing values, the uncertainty introduced through interpolation should be taken in the uncertainty budget."

Line 44: RAOB is not a network, but a software/program, see <https://www.raob.com/>. I would take out any mentioning of RAOB as I do not think this is useful. Some people use RAOB as an abbreviation for radiosonde observations, however, RS may be better suitable for your paper as you are using the RS41 radiosonde. What you probably mean is the operational radiosonde network. However, nowadays most radiosondes measure with a high vertical resolution, so it is more for historical reasons as the old TEMP format is providing low resolution profiles. An increasing number of sites now report data in high resolution BUFR format.

C3

Line 57: The sentence starting on this line is not clear to me. Either take it out or try to improve please.

Line 71: 'the latter being used as missing data for interpolation uncertainty assessment' if I understand it right, you are not using the testing set to assess the interpolation uncertainty, but rather the lack of this data is used to assess the uncertainty. Thus, please reword.

Line 73: Could you please give a reference for 'block-bootstrap cross-validation scheme'?

Line 90: I think a better word for 'station' in this context is 'site' which is also used on the GRUAN website. I would suggest correcting that throughout the paper.

Line 117: Please explain what  $e(t)$  is.

Line 118: I don't understand what you mean with a local dynamics. Could you please clarify?

Line 141: Is  $E$  the expectation value? Please clarify.

Line around 203: Atmospheric output layering. Could you please clarify what you mean?

Line 216: I do not really understand this sentence. Please clarify what do you mean with the layering problem? Have you then selected a 400 second layer? If so, please write second out. However, please rewrite the entire sentence as I also am unsure what you mean with the atmospheric adaption. You may need to explain more what you mean so that the reader can follow you.

Line 231: Please clarify what you mean with 'overall linear interpolation uncertainty'? Is that an average of the two approaches or which approach is shown here?

Line 236: I would suggest to always use 's' as unit for second and not " in some instances. s is the SI unit for second and is therefore preferable and clearer. Most

C4

importantly though, be consistent throughout the entire paper including the Figures.

Figure 10: As mentioned for other Figures as well, your figure caption should be improved. It should be clarified that the different lines show ranges of altitudes. Please use [s] as the unit of second based on SI standard as done earlier in the paper. Either in the caption, or in the text, can you please explain how interpolation distances of 70s are possible. Does that happen by merging the 30 and 10 second gap sizes? Delete the title of the Figure, as the description should be given in the caption.

Figure 11: As you are comparing this Figure to Figure 10, it would be helpful to have the same y-axis limits. You use "Lint" in the Figure labels, but have never introduces this abbreviation, please make sure you always introduce them. Also, please include the suggestions made at Figure 10. Delete the title of the Figure, as the description should be given in the caption.

Line 270: You state that the SE systematically underestimates the interpolation uncertainty, however, looking at the plots, they initially look very similar. Then I noticed that you did not use the same scale, which does make it harder to understand. Please correct this for the readers ease.

Line 270: Please also explain why you expect the standard error SE and the RMSE to be about equal? This is probably obvious to you but may not be to the reader and therefore a sentence to refresh that in peoples thinking would help.

Line 282: What I am missing here is a discussion of the results of the Figure 13. Depending on where in the profile values are missing, the estimated uncertainties are able to enclose the pseudo-missing data or not. At the fourth gap from the bottom, even the corrected, bootstrap does not agree with the actual measurement, as the measurements were missing in a local minimum of the profile. This should at least be discussed as this is a fundamental issue with interpolation which cannot be resolved easily.

## C5

Line 299 and 300: It again confuses me that you are first stating the two approaches are interchangeable and then, in the next line, you explain that you suggest a bootstrap corrected formula to integrate both approaches. There is something missing to clarify this.

Line 311: Is there any way the proposed method can be used by individual researchers? I.e. when using radiosonde data in the old, low-resolution TEMP format and interpolating to a given pressure level. Could lookup tables produced by GRUAN help to estimate this uncertainty as well? Or could lookup tables for averaged interpolation uncertainties be provided to give an estimate to other researchers?

Technical corrections:

Line 12: 'Since both ...' Please clarify that you mean both approaches here.

Line 27: high vertical resolution

Line 38: I think "contemplate" may not be the right word here as it is another word for "consider" indicating it is not applied in OSSSMOSE.

Line 48/49: "in this frame" maybe better say: in this publication

Line 53: use 'similar' instead of 'the same' and delete 'to a large extend'

Line 61: correct 'note' to 'noted'

Line 63: replace 'integrating' with 'using'

Line 76: Correct spelling of 'soruces' to 'sources'

Line 95: correct 'radiosonde's' to 'radiosonde'

Line 96: Write out GNSS

Line 96: Maybe better: The raw data are corrected for known or experimentally evidenced systematic effects such as adjustments from ...

## C6

Line 104: Please try to reword this sentence

Line 139: The formatting here is weird. Is the equation meant to be in-line? Or is it lacking a number? This seems to be a recurring problem in the following equations, i.e. some have an own line, but no number. Something also went wrong with the line numbering between 140 and 145 and at later instances where equations are present. In those instances, I provide approximate line numbers.

About line 144: I don't understand the sentence starting with: Since, using field... Please rewrite

Line 149: Try to reword 'with abuse of'

Line 162: write Equation (1) instead of model (1)

Line 182: replace 'little missing gaps' with 'few missing data'

Table 1: I would suggest changing the table caption to something along the lines of: GRUAN sites included into the Few\_nan dataset and the respective number of launches (imported) and the number of profiles that have less than. . . .

Table 1: I can only guess from the text and comparing the numbers what the column 'Imported' and 'selected' means, please clarify. Also, neither Ny-Ålesund nor Payerne are spelled as on the GRUAN website.

Line around 187: Please correct the sentence starting with: Mudelsee (2014) . . . as it misses a verb.

Line three below 187: take out 'able'

Line 216: Use were instead of where in: . . .the results were. . .

Line 218 . . . RMSE compared to. . .

Line 223: Maybe better: . . .set-up, we decided to use or settled on using the simplest . . .

C7

Line 225: champaign – I assume you picked the wrong word here.

Line 238: be more specific in: . . . have slightly larger values. . . about what you mean with values, i.e. it is the RMSE.

Table 2: This table should come before Figure 7.

Line 242: Could you please state for which interpolation method you will show the plots in the latter?

Line 245: Maybe better . . . Lauder typically have the largest values.

Figure 8: You use the abbreviation CV in the figures but have not introduced it. Also, in the label you say that the x-axis is based on the RMSE. If I understand correctly what you did, it us the RMSE in [K]. Please reword. To clarify the effect, I suggest fixing the x-axis limits, or at least mention in the caption that the limits change from one panel to the next.

Line 261: Please write 70 seconds (or s) rather than 70".

Line 267: Please rewrite the long sentence starting on this line.

Line 269: "above two graphs" please give the Figure numbers instead as they may have moved around during the process and I assume you mean Figure 10 and Figure 11 which are below.

Line 278: Maybe better ' . . .deleted (pseudo-missing)' as this sentence is currently not nice.

Line 280: Write (Equation 7) and (Equation 15)

Line 282: The SI unit for kilometre is km rather than Km.

Line 282: Here you state the plots shows values around 22km and later you state around 23km. While both is not wrong, it would be nice to be consistent.

Line 294: Write 60s or 60 seconds instead of 60"

C8

In Acknowledgements: Please say what the QTF group is as this abbreviation has not been introduced.

Thank you very much for this nice research article!

---

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2020-161, 2020.