

Answer to Referee #2

Dear Referee,

Thanks a lot for using your precious time in reading our paper and making valuable comments that, we believe, are relevantly improving the paper.

Generally speaking, we appreciated all your comments and changed the manuscript accordingly. Also, note that the professional proofreading of the article has been made.

In the sequel, we answer point by point to your report. All answers below and new text parts in the manuscript, specific to your comments only, are in green, while blue is for changes solicited by the 1st referee, or by you both, or (minor) changes independent on you.

Best regards,
Alessandro Fassò

General comments:

This manuscript contributes to highlight the importance of properly assessing uncertainty when imputation of missing data in atmospheric profiles is realized by interpolation, with novel ideas and tools within the scope of AMT. The statistical approach adopted is innovative, very valid and multifaceted, and it leads to reach substantial conclusions, even illustrated in its practical aspects. The overall presentation is well structured and clear, although some points need further clarifications and improvements, as specified in the following. Thus I would recommend the publication of this manuscript after a minor revision, believing it will be very interesting and useful for AMT readers.

ANSWER: ... thanks ...

overall changes ...

captions and titles of figures and tables have been reworked ...

Specific comments:

Figures: titles are redundant since information is already written in the caption, if necessary please add information in the caption but remove titles; moreover check axis names (e.g. missing in Figures 2-4) and limits.

We eliminated all figure titles but the panel titles in multi-panel figures.

Line 65: please clarify the sentence “thanks to the availability of appropriate data”.

Replaced with: thanks to the availability of "good" profiles without missing data.

Lines 118-120: the error term $\epsilon(t)$ should be introduced after its appearance in Equation (1).

Done in blue.

Lines 118- 124: although the assumption of GP is relaxed in a second phase, it would be suitable to justify or at least comment upon the choice of the two considered autocovariance functions.

Added some comments and a reference.

Line 127: here the assumption of zero expected value for the error term in Equation (1) is implied, while it could be written before.

The zero mean property has been stated earlier, close to Equation (1).

Line 174: the assumption of a GP as a good description of the problem comes with a specified autocovariance function, it would be useful to clarify this.

Added a reference to the role of the covariance.

Line 200: to facilitate reading, it would be useful to specify that m_1 refers to Equation (5) and m_2 to Eq. (9), and only for m_2 we need an estimation method (and so a hat).

Done.

Section 8: this section needs to be revised because a 3×3 simulation design is described but after there are comments about the 60 seconds case (e.g. line 237) and even results (e.g. in Table 2). Please correct consistently to have a 4×3 or 3×3 simulation design in all section.

Done.

Line 228: I would suggest to avoid the technical term "2-fold" since it is not introduced before and not necessary.

Done.

Line 270: it could be useful to clarify this sentence, consistently with the abstract where you state that both interpolation methods provide an underestimation.

We reminded that the underestimation is for both the linear- and the GP-interpolation.

Lines 299-300: it could be useful a line summarizing reasons to integrate the two approaches and then use the proposed bootstrap-corrected formula.

We tried to make this point more clear by focusing mainly on the two uncertainties and blurring the difference among the linear- and the GP-interpolation, which resulted to be equivalent.

Technical corrections:

Line 44: a reference for RAOB would be useful. RAOB is discussed in Finazzi et al. (2019), already cited there.

Line 61: "note" should be "noted". Done.

Line 76: "soruces" should be "sources". Done.

Line 89: a reference for the statistical analysis conducted by GRUAN would be

useful. It is an internal preliminary study, we rephrased to make it more clear: "A preliminary statistical analysis for the occurrence of data gaps in RS41 radiosonde soundings performed at 15 GRUAN stations in the period 2014-2019 shows that gaps occur in more than 20% of the soundings, virtually independent of the height ranges, with the majority ($> 95\%$) having less than 15 gaps per 1000 s."

Table 1: please specify what "Imported" and "Selected" mean. **Done.**

Line 186: add parenthesis for the two references. **Done.**

Line 225: maybe campaign? **Done.**

Line 237: "vey" should be "very". **Done.**

Figure 7: in addition to the general comments about figures, this one appears with a different look, I would suggest to use the same software for all plots. **Done.**

Line 261: in the same line seconds are written differently, please check throughout all manuscript. **Done.**

Figures 10 and 11: captions should be revised since altitude is not represented as axis, and the box with altitude intervals needs a title. **Done.**

Line 282 and Figure 13: please write in both points 22 or 23 km. **Done.**

Figure 11: "Lint" needs to be defined. **Done.**

Figure 12: axis names are missing. **Done.**

Figure 13: please change one blue color. **Done.**

Line 312: please delete "TEXT". **Done.**

Line 325: please explain QTF. **Done.**

Line 335: Finazzi et al. should have year 2019 (that is correct at page 2). **Done.**