Reply to RC2

Thank you very much for your reviewing our manuscript and providing us with valuable comments and suggestions. We reprocessed the IMS-100-GDP because we found some minor issues in the processing data used in the previous manuscript. Furthermore, we applied the updated screening of data following review comments. The number of samples has increased from 57 to 59. However, the main message of the manuscript has unchanged.

Hereafter, Cx represents the referee's comments and Rx represents the reply to Cx.

Major comments

C1) The methods described in sections 2 through 4 are possibly in several cases a bit too brief and too heavily reliant upon the reader going back to and reading a number of previously published papers. Perhaps some key additional aspects need to be documented to ensure broad-scale methodological reproducibility. In particular section 4.5 is too brief given its overall importance in a GRUAN product comparison and should be expanded.
R1) Detailed descriptions are added to the revised manuscript.

C2) For figure 3 and subsequent similar figures it would be useful to describe in the figure 3 caption what the different shapes of the boxes denote? What is a circle, a parallelogram, a diamond, and an oblong?
R2) Explanation of the shapes is added to the captions in the revised manuscript.

C3) It would probably be worth spending some time discussing the very marked seasonality of the rejection rates shown in Figure 10 which is cited but not really discussed. Rejection rates are low in winter but very high in spring and summer. Why is this?
R3) A brief discussion is added to the revised manuscript.

C4) I always hate making this comment because it is immensely impressive for non-native speakers to produce papers in English. However, the paper would be much more readable if you could get a native speaker to edit for clarity. There are numerous places where minor edits would improve the readability and make the messaging stronger.
R4) Thank you very much for your suggestions for edits.

Minor comments

C1) A reader would reasonably ask in the abstract why you mention 99 dual soundings but proceed to
analyze only 57 of these. Can a few words be added to clarify why this 57 subset of the sample were analysed? Something like “Following data quality checks 57 flights were considered of sufficient data quality to produce GDP profiles and this subset is analysed”

R1) We think that this comment is related to the major comment C3. The discussion is added to the revised manuscript.

C2) Line 8 I would delete “with RS92-GDP” as this is already clear from earlier in the sentence

R2) Deleted.

C3) Line 19-20 I would write. “While the RS92-SGP radiosonde has a GDP it was required to seek alternative radiosonde models to use for operational reasons as the payloads often fall within the greater Tokyo metropolitan region and for health and safety reasons use of lighter instrumentation is necessary.”

R3) Revised in response to your suggestion.

C4) In lines 21-22 I would suggest making explicitly clear that the RS-11G has already been developed and certified as a GDP

R4) The following sentence is added: “The GDP for RS-11G was developed and certificated in 2019.”

C5) Line 56 and not ant (typo)

R5) Corrected.

C6) Should line 105 not be 10’ x 10’? A geoid model at 10 degrees by 10 degrees feels implausibly coarse and the change to 5’ by 5’ later in paragraph is then huge.

R6) Co-authors at Meisei confirmed that the original geoid model used for iMS-100 is 10 degrees by 10 degrees.

C7) Perhaps in 2.2.4 make clear why no pressure sensor is fitted. Presumably this is to save weight and because the errors in the hydrostatic equation based approximation are considered sufficiently small to justify the omission of such a sensor?

R7) Yes, and the cost reduction is also one of the reasons. The following sentences are added: “Not installing a pressure sensor is to reduce weight and manufacturing costs. However, in recent years, radiosondes without a pressure sensor have tended to be used because the accuracy of the atmospheric pressure derived using hydrostatic equation-based approximation is sufficient (e.g., Nash et al., 2011).”

C8) Lines 205-206 are not possible to follow logically. How did a single payload last 7 months? What
is a logistic regression? Work required here to clarify this sentence please.

R8) Logistic regression analysis is used for probability prediction. Training data is the routine data (twice per day) at Tateno. The beginning of the paragraph is revised as follows: “For screening of ice-contaminated profiles, the probability of icing, $P_{\text{ice}}$, is derived using logistic regression analysis after variable selection from the length of ISSR, RH and the volume mixing ratio at several levels. The 452 routine observations (twice per day) with a single payload taken from April to November 2018 at Tateno are used as training data.” And the reference of the logistic regression is added to the revised manuscript.

C9) Lines 218-219 are again unclear. I think you mean something like: The criteria applied here apply solely to the present study and are not applied to the GDP. However, you also probably need to explain why this is the case. Surely the same processing should be applied to the GDP screening or is there some reason why this would not be possible?

R9) There are no authorized screening methods for the GDPs. This screening method is unique to this study. Users of the data may use their own methods of screening for their interests. The concept of screening with the uncertainties is added to the revised manuscript as follows: “The second screening is based on the uncertainty amounts (Sommer, 2013). This screening is based on the idea that data with uncertainties exceeding the criteria are of questionable reliability and need to be verified individually.”

C10) The paragraph starting line 302 is presumably applicable to all meteorological elements measured and not just temperature. As such its placement here rather than elsewhere – perhaps most logically the discussion section of the paper – feels strange to me. There is already similar text on lines 398-405 so maybe you can just delete this?

R10) This paragraph was deleted in the revised manuscript.

C11) Line 388-389 it is unclear for what parameter this finding applies. I assume pressure but it needs to be stated explicitly.

R11) The sentence will revised as follows: “While there are some cases where with significant differences for pressure are observed in the lower troposphere ($\geq$ 700 hPa) in the consistency check, the mean pressure difference is less than 0.4 hPa.”