Author response to Reviewer # 2, Paul Poli: Estimation of refractivity uncertainties and vertical error correlations in collocated radio occultations, radiosondes and model forecasts

Johannes K. Nielsen[1], Hans Gleisner[1], Stig Syndergaard[1], and Kent B. Lauritsen[1]

[1]Danish Meteorological Institute

[1]Lyngbyvej 100, DK 2100, Copenhagen, Denmark

Correspondence: Johannes K. Nielsen (jkn@dmi.dk)

1 Authors response:

We acknowledge the constructive suggestions from Reviewer #2, which have helped to clarify several aspects in the manuscript. We have considered all suggestions, and adopted most of them. All issues raised by Reviewer # 2 are addressed below, and we have also inserted some requested figures.

2 Detailed comments:

(1) Given that the authors make an explicit attempt to tie the terminology to established documents like the GUM, other prior relevant publications may deserve to be cited, namely those that already considered the GUM and its applicability to Earth Observation data, e.g., from the FIDUCEO project: Merchant, C. , G. Holl, J. P. D. Mittaz, and E. R. Woolliams. 2019: Radiance Uncertainty Characterisation to Facilitate Climate Data Record Creation. Remote Sensing 11, no. 5: 474. doi:10.3390/rs11050474

Answer 1: Thank you for suggesting this reference. We have inserted a citation.

(2) Section 1.2 mentions "RO reference coordinates" and "RO reference time": how are they defined in the present work?

Answer 2: We have inserted this in Sect. 1.2: "The RO reference coordinate is the point at which a straight line between the GNSS satellite and the receiving Low Earth Orbiter tangents the Earth ellipsoid."

(3) Is it possible to indicate (or cite the appropriate reference) for the step where radiosondes and ERA5 data are projected into refractivity space?

Answer 3: We have inserted this (see also Answer 7): "The refractivity calculation is done with the method described in the

ROPP user guide: https://www.romsaf.org/romsaf_ropp_ug_fm.pdf." The GRAUN data are mapped to the same 137 level grid as ERA5 before being forward modelled to refractivity space.

(4) The approach to calculate epsilon C and epsilon X needs to be detailed, and preferably with dedicated equations for clarity. This would also remove the need for forward references to sections 4.2 and 4.3 in section 1.2.

Answer 4: Instead of referring to Sections 4.2 and 4.3 we have added this:

“We are able to remove $\boldsymbol{\varepsilon}^C$ and the $\boldsymbol{\varepsilon}^X$ components of the three data sets, by adding the following additional analysis steps to the G3CH. The $\boldsymbol{\varepsilon}^C$ covariance matrix, $\mathbf{C}^C$, is eliminated by first calculating G3CH estimated covariance matrices $\mathbf{C}_i$ for a series of collocation subsets, sampled from areas of decreasing size around th RO reference coordinates. Next, the sequence of decreasing covariance estimates is extrapolated to the virtual zero-area case $\mathbf{C}_0$. $\mathbf{C}_i^C = \mathbf{C}_i - \mathbf{C}_0$. Subsequently the ε^X covariance matrix, $\mathbf{C}^X$, is eliminated by smooting all three data sets such that they have the same vertical footprint, and then calculate for each data set a covaraince matrix $\mathbf{C}_s$ with G3CH from the smoothed data sets. $\mathbf{C}^X = \mathbf{C}_0 - \mathbf{C}_s$. So the observation error covariance matrices that we estimate in the end includes only measurement error $\boldsymbol{\varepsilon}^I$ and representativeness error $\boldsymbol{\varepsilon}^R$.”

(5) A map showing the locations (or a density map) of the 15,997 selected collocations may be a useful information for the readers.

Answer 5: We will add this map as supplementary material.

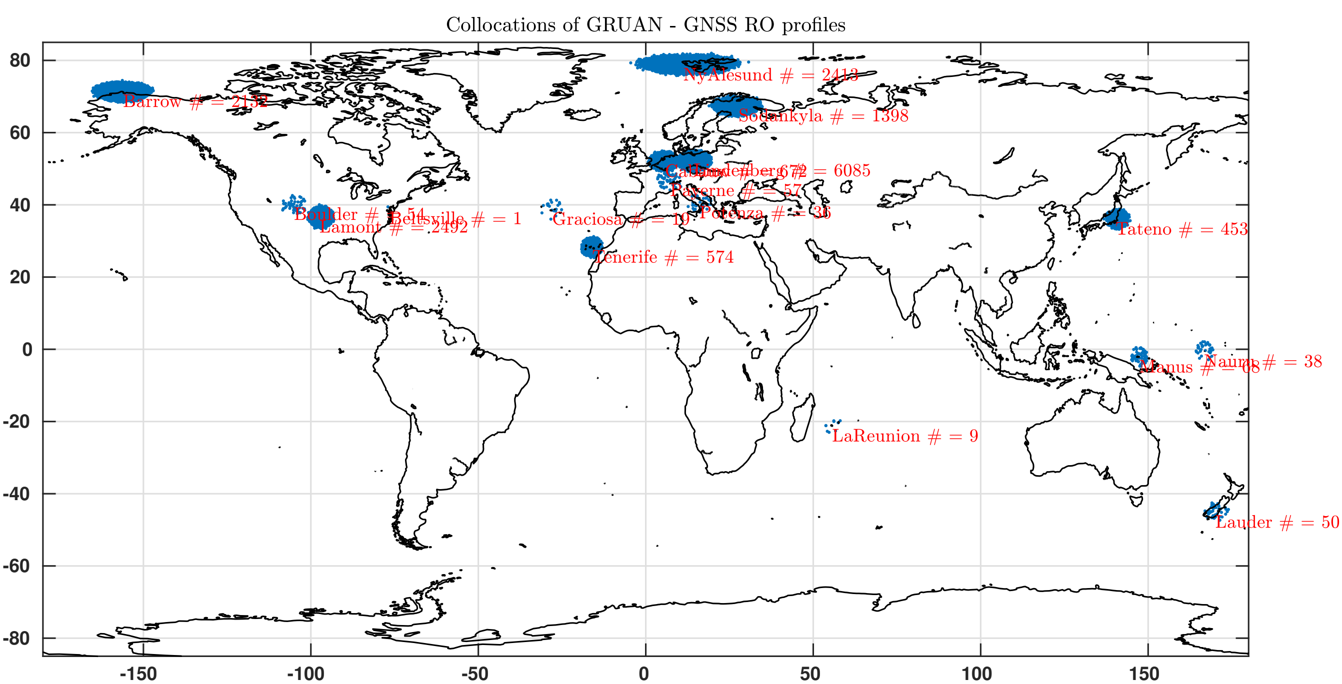

Figure 1. Reference locations of GRUAN RS92 sondes.

(6) Is it possible to comment on the possibility that the two steps of cubic spline interpolation, and of refractivity computations, both applied to ERA5 and radiosonde data, may (each) introduce correlations of uncertainties between these two datasets? Similarly, given that the assimilation of RO data improves the quality of a reanalysis, what are the prospects for some structural correlation between ERA5 forecast (even if, not analyses) and RO data?

Answer 6.1: The spline interpolation is done on a length scale which is smaller than the common footprint from ERA5 and therefore we do not expect it to create correlations on the length scales where we draw conclusions here (> 300 m)

Answer 6.2: The forward model does introduce some correlation through interpolation from the model grid to the observation grid. Again, these correlations will appear on length scales smaller than the common footprint, and as such they should not influence the estimated refractivity uncertainty.

Answer 6.3: Such contamination would manifest itself as a bias change in the analysis which subsequently spills over to the forecast. Given that biases are eliminated in the application of G3CH, it can be assumed that no such information spill over is taking place. (No changes made to the manuscript as response to (6))

(7) The numbers of vertical levels in each dataset may need to be introduced in the data section.

Answer 7: We have inserted this information:

Line 95: "Metop and COSMIC-1 missions (Gleisner et al., 2020), interpolated to 247 levels."

And on line 110: "The RS92 temperature, humidity and pressure variables have been interpolated with cubic splines to the 137 ERA5 model levels, hereafter the ERA5 and RS92 variables have been forward modeled to refractivity at the RO vertical grid of 247 levels(Lewis, 2009). The refractivity calculation is done with the method described in the ROPP user guide: https://www.romsaf.org/romsaf_ropp_ug_fm.pdf."

(8) As the work is using model forecast, whose quality decays as the integration time increases, may one expect a sensitivity of the results to the forecast integration time?

Answer 8: That is correct, we have chosen to ignore that, so the ERA5 uncertainty is representative for 3 to 15 hours forecasts. We have inserted this in line 107: "Effectively this implies that the used verification times runs from 3 to 15 hours, and the ERA5 uncertainty is assumed to be constant in this time range."

(9) Relying on correlations to pick-up a signal exposes one to be sensitive to any transient or structural correlation that may exist in the input data that are correlated, whatever the reason (true signal, artefact of pre-processing, matching bias, . . .). In the present case, the step of bias removal seems to be limited to a mean profile subtraction, carried out at the scale of each entire dataset (RO, RS, ERA5), is this correct? If so, this would leave, present in the data, all the bias(es) that may exist within each subset of analysis. Would it be possible to consider applying the bias removal in each subset (i.e. at the step when expectation values are computed, modifying slightly equation (7) to introduce the removal of the means), and then display (or report on) how much this changes the results (or not).

Answer 9: The biases are removed for each subset separately in the application of G3CH, such that for instance when looking at rising occultations at middle lattitude, the 3 bias vectors are calculated individually for this subset before applying 3CH. We can see how this is not clear from the paper, and we have changed the formulations:

Line 66: "...but for each subset of collocated triplets being analyzed, we remove systematic error differences between the three involved data subsets prior to the analysis."

Line 126:"In the present paper we only estimate the random uncertainties. In practice we remove biases in each subset of collocations where G3CH is to be applied by subtracting the subset mean of each of the three data sets prior to the analysis. So in the following derivation we can assume that all data are bias free."

(10) Do the brackets in the right-hand side of in Equation (7) reflect the actual implementation? (i.e. averages are computed after adding all cross-products?)

Answer 10: No, each cross-product is averaged individually before adding the 3 terms.

(11) Can you clarify how the data subsets (the sub-spaces in which expectation values and hence correlations are computed) are defined? This seems to be, at least initially, based solely by considering the vertical dimension, but then later in the paper other dimensions (for computing the correlations and presenting the results) are introduced. This may be done with various sets of subscripts (for the various dimensions: vertical, latitude band, . . .). In the ideal case where one would have many events, one could consider to compute these error estimates with subsets defined spatially (e.g. 5 deg x 5 deg). The resulting geographic patterns that may be obtained could be of interest.

Answer 11: The choice of the GRUAN RS92 profiles as demonstration data set is the limitation here. While being a limited quality checked data set of high accuracy, it only covers a very sporadic area. As it is also evident from the results, especially in the tropics, the data set is barely large enough to justify separation into 3 latitude bands. In future applications one might (probably should) choose a much larger set of sondes with better coverage.

(12) Would it be possible to define early on, i.e., in the methodology section, the 'raw uncertainty estimates' mentioned in the results section? Also, the sigma symbol may deserve to be introduced numerically with an equation.

Answer 12: Good question: "Raw uncertainty estimates" are uncertainty estimates calculated from some subset of triple collocations without correcting for collocation error and without performing any smoothing. Unfortunately we have used the term "raw" in Fig. 10, refering to estmates that have been collocation corrected but still unfiltered (this will be corrected). We insert this in the Figure 9 caption: "and e.g., $\frac{\sigma_{\mathbf{ERA5}}}{\sigma_{\mathbf{ERA5},0}}$ in subplot **(a)** means uncertainty estimate of ERA5, given filtering of the data set mentioned in the legend — in this case RS92, divided with the uncertainty estimate of ERA5, obtained without filtering."

(13) In figure 7, does the "0 m" line refer to no filter? If so, such a filter has an infinitely small width (Dirac), but is probably non-zero.

Answer 13: Yes "0 m" means zero width and is equivalent to no filtering; multiplication with unit matrix.

(14) The results in figure 7 indicate that as one filters out small-scale variability in both other datasets, the dataset that appears to be most affected in its 'error' estimate is ERA5 (this one presents the largest spread, nearly 1filter, in the lower troposphere).

This would be consistent with that dataset containing the least small-scale vertical information, given that equation (7) suggests that for an error estimate to increase, there are two pathways: the cross-products of differences with respect to the two other datasets increase (first two terms in the sum), and/or the cross-product of the differences between the two other datasets decreases (third term, negative sign). The latter may be the mechanism by which removing small-scale information in RS and RO data (thus reducing the differences RO-RS) leads to ERA5 to appear of worse quality (when in fact its quality should be independent of that, but here the method uses the other data as references). Such hand-waving comments (for lack of a better expression) may be tried with a simple toy model. Similarly, the dataset whose 'error' estimate is the least affected by filtering the small-scale variability in the two other datasets seems to be the radiosondes, which is also consistent with that data source possibly containing the most small-scale vertical information in the lower troposphere (or is the figure 7(c), truncated at a maximum of 2.0

Answer 14: We agree with the interpretation, and we only want to clarify that ERA5 estimated uncertainty should be independent on the quality of RO and RS: This is in principle true, but it fails to hold when the vertical footprints of the RO and RS data are smaller than that of ERA5. We apologize for the truncation of Fig. 7 (c), which was chosen to make the relative small differences higher up visible. We will exchange it with a version with expanded x axis.

(15) Figure 9 shows seemingly slightly different results because in this case one considers smoothing on only one (other) dataset a time. However, one finds consistency. When (only) RS or (only) RO data are filtered (in (a) and (c), respectively), then either one of the two may start to resemble more to ERA5 (but less to the other, i.e. RO or RS, respectively), so, in equation (7), the three terms that make up the total error estimate sees changes of different signs in its components (respectively, for the 3 terms in the right-hand-side of equation (7): decrease of differences ERA5-RS, no change ERA5-RO, and minus an increase of differences RS-RO – the net result is then a decrease of ERA5 estimated 'errors' when only RS is filtered). Similarly, this would explain that the 'error' estimate of RO increases in (b) when (only) RS data are filtered, making them resemble more ERA5 (respectively: increase of differences RO-RS, unchanged differences RO-ERA5, and minus a decrease of differences RS-ERA5). I note in passing that one missing piece of this puzzle would be to show what is happening to the error estimates of each dataset, when one filters that dataset only, and none of the other two datasets, as the results are not entirely predictable because they involve the sum of two terms moving in opposite directions, e.g., for RO error estimates, if filtering RO data: the differences RO-RS may increase, the differences RO-ERA5 may decrease, the differences ERA5-RS would be unchanged (so the net result is hard to predict – but such a thought experiment may help shed light on the optimal 'footprint' to characterize each dataset).

Answer 15: We show plots, where only the shown data set itself is filtered here:

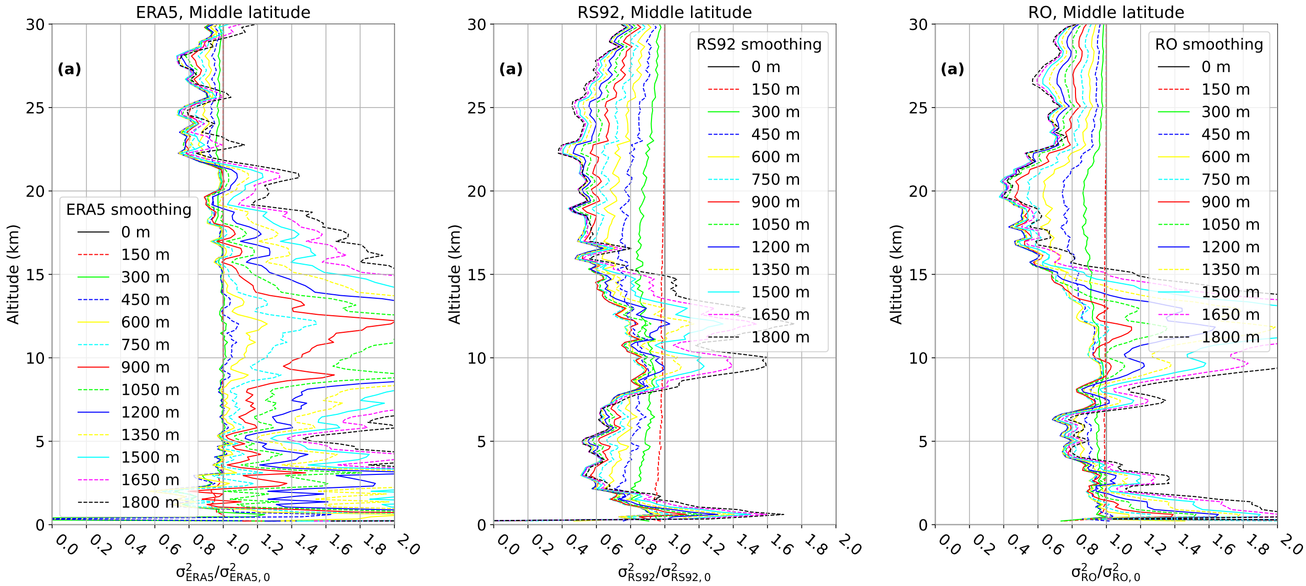

Figure 2. G3CH uncertainties when only the variable at question is filtered.

As expected ERA5 gets further away from the other data when filtering is applied.

(16) With respect to the expected long-range correlations expected in RS data (but not picked-up as well as expected by the method), this may also be due to the choice of sub-setting considered here, analyzing together day-time and night-time data. The effects of the radiative corrections (leading to consistently positive or consistently negative differences in each profile) may, if not cancel out, possibly be reduced, when considered together. However, redoing the exercise by separating clearly night- and day- ascents (and possibly leaving aside those profiles ‘in between’), may show slightly different results. Such a separation for the RS data would somehow echo the efforts made to separate between rising and setting events for the RO part.

Answer 16: Thank you for this suggestion. We think it is interesting, but since the main focus here is to understand RO uncertainty, we will leave this out, and possibly try it in future applications where GRUAN data are more in focus.

(17) Figures 11-12, I fail to see the labels (a) to (f) (either add these or amend the figure caption?).

Answer 17: Thank you for spotting this, we shall fix that.

References

Gleisner, H., Lauritsen, K. B., Nielsen, J. K., and Syndergaard, S.: Evaluation of the 15-Year ROM SAF Monthly Mean GPS Radio Occultation Climate Data Record, Atmospheric Measurement Techniques, 13, 3081–3098, https://doi.org/10.5194/amt-13-3081-2020, 2020.

Lewis, H.: GRAS SAF Report 08 ROPP Thinner Algorithm, Tech. Rep. Ref: SAF/GRAS/METO/REP/GSR/008, EUMETSAT, 2009.