Author response to Reviewer # 1, John Eyre: 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:

The authors wish to thank for the review. We acknowledge the comments about the "truth" being in principle unknown in 3CH analysis, and we shall reformulate sentences that explicitly assumes a specific "truth". However we will maintain the view that, for heuristic reasons, it may be convenient to imagine an underlying true profile for each collocation set during the derivation or application of G3CH. Certainly one must not be mislead to conclude anything about the nature of that truth from 3CH. We are also grateful for the clarifying and editorial comments raised by the reviewer.

All the raised issues are addressed below:

1.1 Detailed Comments:

5. p.1-2, l.23-25: "This is likely because of the requirements — that the errors of the three data sets must be uncorrelated, and that the data sets must truly represent the same property with the same footprint in time and space — that are seldom met." This is not the main reason; it is that NWP DA theory requires that the errors of representation (all of them) are considered as part of the observation error, not part of the NWP forecast error, even though they arise because of the NWP system's limited ability to represent the real world. With the 3CH method it is not clear how the errors of representation will be portioned between the 3 data sets. (The Desroziers method does not have this problem.)

Answer 5: We have replaced the sentence with:
"This is likely because in NWP data assimilation all the model representativeness errors, including forward modeling errors, are considered as a part of the observation error. The 3CH method is not targeted specifically at NWP applications. This means that all three data sets involved are treated equally as a start, thus they are all assumed to contain representativeness errors,

with respect to the underlying truth. In order to use results obtained from the 3CH analysis it is necessary to consider, for each particular application, how representativeness errors are distributed among the involved data sets, and this is not always possible to find out."

6. p.2, l.56-56: "The term vertical footprint of a data set is used here in the meaning width of an ideal physical refractivity feature, shaped as a delta function, mapped to the resolved representation of refractivity, for the given data set." This is not very clear, partly because a (Dirac) delta function has zero width. Please define more clearly what is meant to "vertical footprint", how it differs from "vertical resolution", and how it is related to the vertical smoothing discussed later.

Answer 6: We have changed the mentioned sentence to this formulation (with a reference) "The term *vertical footprint* of a data set is used here in the same way as in Semane et al. (2022): *The vertical scale that an observation value represents.*"

7. p.3, l.62-63: "t is the actual refractivity at a vertical line above the RO reference coordinates at the RO reference time". 3CH method, i.e. the solution of simultaneous equations described later in the paper, makes no assumption about exactly what the reference profile is. In fact, this is the paradox of the 3CH method, as discussed by O'Carroll et al. 2007 (appendix to their paper). The solution to the paradox is through the appreciation that non-zero correlations of error (assumed to be zero in the 3CH method) arise because each of the data sets represents different spatial scales. Alternatively, these can be considered as correlated errors of representation when the data are assumed to measure the same scales. There are also the error correlations caused by space/time collocation differences. In general, the paper discusses very well the issues of scales and error correlations, but it would be helpful to point out that they also are also related to the problem in defining the "true" profile.

Answer 7: We have made this substitution at line 62: "In the context of this paper t is the actual refractivity at a vertical line above the RO reference coordinates at the RO reference time, defined with respect to given finite footprints in space and time, which may differ from the footprints of all three data sets."

— has been changed to:

"The G3CH does not make any assumptions about exactly what the true profile $\mathbf{t}$ is. $\mathbf{t}$ may be thought of as defined with respect to a given but unknown finite footprints in space and time, which may differ from the footprints of all three data sets."

Error cross correlations caused by space/time collocation differences are addressed through the discussion of ε^C and ε^G components in line 75 to 78. We assume that RS92 and RO sampling paths are uncorrelated, and that could in theory be disputed, but we anticipate that this issue is of minor importance. But one thing that could play a role, as also mentioned by Reviewer #3, would be horizontal cross correlations between ERA5 and RO, which are at least of similar horizontal scale.

We have substituted:

"Hence the $\boldsymbol{\varepsilon}^G$ term contains no cross correlations, and consequently it will be correctly attributed to the RO and RS92 data by the G3CH procedure."

with:

"Hence the $\boldsymbol{\varepsilon}^G$ term can be assumed to contain no cross correlations. However, there are potentially error cross correlation

components arising from spatial correlations between the data sets, that we cannot assess. This could for example be the case for ERA5 and RO, because these are sampled on similar horizontal scales."

8: p.3, l.66: "we assume that". This is a little confusing, because you say later that you actually processed the data to ensure that this was the case. (You did not just assume that it was true.)

Answer 8: The sentence have been changed to "..., but for each subset of collocated triplets being analyzed, we remove systematic error differences between the three involved data subsets prior to the analysis."

9. p.3, l.68-69: "The ε^R component represents the distortion of the underlying truth in a data set, as it is being mapped to the observation grid." This is not clear. To which observations does it apply?

Answer 9: In the "Error components" section error components are discussed generically such that the discussion applies for all 3 datasets. The model forecast is virtually treated as just another observation in this context. We have inserted a clarifying sentence at line 41: "The three data sets are treated on equal terms such that none of them are considered more or less representative for the truth a priori, thus the analysis will also provide estimates of ERA5 ... "

10. p.3, l.71-72: "representing the departure of the RO and RS92 trajectories in time and space from the vertical profile at the RO reference time." Again, this is identifying the reference profile with the "true" profile, and the problems discussed in point 7 above apply.

Answer, 10: Correct. We have changed the formulation to: "Especially ε^R contains a geometric error component, ε^G, representing the departure of the ERA5, RO and RS92 vertical or skewed profiles in time and space from the unknown true profile at the RO reference coordinates."

And in line 74 we have added: "The used forward operator estimates refractivity along a 1 dimensional assumed vertical line, and this has an impact on the uncertainty estimates. Thus, the RO observation errors estimated by the 3CH method in this paper are applicable for variational assimilation with a 1D operator, but not for 2D/3D operators."

11. p.3, l.83: "intrinsic error". Is this the same as epsilon, as introduced on line 61?

Answer 11: Yes. "Intrinsic error " is maybe confusing. We have changed it to "observation error"

12. p.3, l.88: "ε^R". This is the representation error for this definition of the true profile, but it would not be the appropriate representation error for use in NWP DA. See point 5 above.

Answer, 12: We agree with this statement.

13. p.4, l.105-107: "model forecast ...". What is the range of this forecast, e.g. 3-, 6- or 9-hour forecasts interpolated to the RO observation time? Also, "is model forecast" → "contains forecasts"?

Answer, 13: The ERA5 forecasts have been downloaded in 3 hourly intervals. For each profile two fields of as little forecast time as possible are chosen, such that the RO and RS92 profiles are outside the analysis windows used for initialization of these 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."

85 We have changed "is model forecast" → "contains forecasts"?

14. p.4, l.125: "we assume no cross correlation components". Yes, this is what the algebra of the 3CH method assumes, but it is the weaknesses in this assumption that represent the problems with the method – see point 7 above.

Answer 14: That has been noted.

15. p.4, l.127: "it can without loss of generality be assumed that all three data sets are bias free". Again, this is ambiguous –
90 you did not just assume it – you actually ensured that it was true through the data processing (lines 151-152).

Answer 15: We have changed the formulation accordingly: "In the analysis in the present paper we only estimate the random uncertainties, so it can without loss of generality be assumed that all three data sets are bias free. This can in practice be ensured by subtracting the mean of each data set prior to the analysis." has been changed to: "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
95 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."

16. p.6, l.167-168: "However, if two data sets have similar vertical footprints, differing from the vertical footprint of t, these two data sets will have cross-correlated errors". Yes, they will have correlated errors, almost independently of how t is defined – see point 7 above.

100 **Answer 16:** We agree on that, no action needed.

17. p.7, l.180-191: "Uncertainty estimates ... . In the three cornered hat analysis, the data set with the largest footprint determines the common footprint to be used for all three data sets.". I don't think this is true – see point 7 above. It is certainly not consistent with the assumption that the "true" profile (relatively to which all errors are assessed) is a vertical profile at the nominal RO location. Also, the use of the ERA5 scale as a common scale to which all data are smoothed is certainly a good strategy for the
105 reasons stated but, again, it is not consistent with the definition of the true profile.

Answer 17: Right, the formulation is wrong. We have changed this: "In this analysis.." —> "In our approach..."

18. p.8, l.205: "between 50 km and 300 km". This must reduce the sample size by a factor of 36. It raises the question of question the sample size is still big enough. From the results it appears to be so, but this may be worth a comment. Also, please could you comment on the related problem of temporal collocation window.

110 **Answer 18:** The covariance estimates near 50 km collocation distance are indeed uncertain, as is also seen in Fig 5. We believe this is the main reason for the noise in the uncertainty profiles. We tried at an early stage to include both distance in time and space in the analysis, and found that the impact of time distance between GRUAN and RO profiles on the GRUAN - RO temperature standard deviation was negligible, at least as long as the time criterion was less than 3 hours, so we left that

out of the analysis.

115   The text has been changed to this: "We are calculating the G3CH estimates of covariance matrices for a sequence of collocation criteria (between 50 km and 300 km) and use these to extrapolate all covariance matrices to 0 km collocation distance, with a linear fit to the full covariance matrices as function of the squared collocation distance. The effect of varying the temporal collocation window is small, so we have excluded that from the analysis."

19. p.8, l.213-214: "For each data set filtering has been performed, not on the data set itself, but on the two other complementing
120   data sets (see figure legends)." This is not very clear, and the figure caption is not any clearer. I think I understand what has been done, but a few more words of explanation would be helpful, e.g. for the calculation for the curve shown for RO, only the RS92 and ERA5 data have been smoothed?

Answer 19: That is correct. It is as explained, but maybe it would help with more motivation or explanation of why this procedure is chosen. We have inserted:
125   "The idea is basically to probe the vertical footprint of one data set with two other data sets of varying footprint." On line 214.

20. p.9, l.253-254: "It is worth noticing that the estimated vertical correlations of RS92 are larger for setting than for rising RO at high latitudes, especially between 6 and 22 km. So the G3CH fails to give an independent estimate of the RS92 correlations." This is a helpful warning; it is another illustration of a weakness in the 3CH and G3CH methods, through their implicit assumption that correlated errors (between data sets) are zero.
130   **Answer 20:** We agree that the explanation can be error cross correlations.

21. p.10, l.269-273: "In the derivation of G3CH representativeness is defined with reference to given scales in space and time of the truth. The truth is assumed to have smaller footprint 270 than any of the involved data sets. We choose for all data sets to report the estimated uncertainty boundaries with reference to the estimated footprint of the ERA5 data set. ... This operation is equivalent to define the truth t with reference to the ERA5 footprint if one will." This is not consistent with 1.2 para 1. It again
135   illustrates that the choice of t is somewhat arbitrary with this method, and that the error correlations will change according to the scale of the "truth".

Answer 21: We agree that 1.2 para 1 is misleading because it suggests that the G3CH scheme works for a specific truth, and we have taken that part out of the paper, and we have removed this sentence at line 270: "The truth is assumed to have smaller footprint than any of the involved data sets." And even though it is not really wrong, we have removed this sentence at line
140   272: "This operation is equivalent to define the truth t with reference to the ERA5 footprint if one will."

22. p.10, l.287: "The increase of uncertainty in the troposphere is smaller at high latitude". Why is this? It may seem an obvious point, but you would expect larger errors in the tropical lower troposphere because the absolute values of humidity are highest there and hence the collocation errors in refractivity will also be highest.

Answer 22: We agree that tropospheric humidity is causing a lot of uncertainty in the tropical troposphere, but since the paper
145   is not really addressing physical interpretations, we think that a comment on this would appear unmotivated in this context.

23. p.11, l.331: "model" → "NWP"?

Answer 23: Changed.

Editorial comments 24. Throughout. "data sets", "datasets" or "data-sets". Consistency.

Answer 24: Changed to "data sets"

150   25. Throughout. "data" is usually used as plural, but in a few places as singular. Consistency.

Answer 25: Thanks.

26. p.4, l.98: "has" → "have". Also, l.103 and l.104.

Answer 26: We have chosen singular

27. p.4, l.109: "spans" → "span".

155   **Answer 27:** We have chosen singular

28. p.6, l.155: "is" → "are".

Answer 28: Not found

29. p.8, l.224: "happens" → "happen".

Answer 29: Changed

160   30. p.10, l.274: "includes" → ", include"?

Answer: 30: Changed

31. p.10, l.289: "does" → "do".

Answer: 31: Changed

32. p.11, l.321: "promises" → "promise".

165   **Answer: 32:** Changed

33. p.11, l.325: "forecast" → "forecasts".

Answer: 33: Changed

34. p.22, Figure 7 caption: "easiest" → "most easily"

Answer: 34: Changed

170 References

Semane, N., Anthes, R., Sjoberg, J., Healy, S., and Ruston, B.: Comparison of Desroziers and Three-Cornered Hat Methods for Estimating COSMIC-2 Bending Angle Uncertainties, Journal of Atmospheric and Oceanic Technology, 39, 929–939, https://doi.org/10.1175/JTECH-D-21-0175.1, 2022.