Author response to Reviewer # 3, Anonymous: 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 #3, which has lead to substantial changes of the manuscript. Reviewer # 3 calls for clarification of assumptions and certain aspects of the argumentation, and indeed the questions raised has pointed out parts of the text that has lead to misunderstandings. We apologize for the lack of clarity, especially regarding the definition of "vertical footprint", which has understandably frustrated the reviewer. By addressing every question raised by Reviewer # 3 below, and changing many formulations in the manuscript we hope that we have sufficiently closed all outstanding issues. Reviewer #3 calls for figures to illustrate some concepts of the paper. We cannot accomplish that, but we think that the explanations below and the revised manuscript eliminates the need for drawings.

1.1 Detailed Comments:

Line 36: this sentence is imprecise. One of the purposes of the paper is to take into account error correlations. Errors contain a random component. Therefore stating that the random error components are independent appears mis-leading. Random errors can be dependent and correlated. This should be re-phrased.

Answer, line 36: Right, that is unclear. The intention is to state that we anticipate error cross correlations, as defined in the paper, to be negligible. We have changed this formulation: "We apply the generalized 3CH (G3CH) to three data sets where the random error components can be assumed to be truly independent." into: "We apply the generalized 3CH (G3CH) to three data sets where the random errors components can be assumed not to be interdependent, meaning that their error cross correlations are assumed to be negligible."

Line 39: while it is true that the authors focus on vertical error correlation, they have not made a convincing case that error correlation might not arise for other reasons. It seems that the current analysis could proceed at a particular vertical level, in which case it would seem incorrect to assume that all error correlation is from the vertical dimension. More on this point later.
Answer, line 39: It is correct that the analysis can be performed on a particular vertical level. That would be the classical three cornered hat analysis, which has been studied several times in literature. One would obtain the diagonal elements of the covariance matrices, a perfectly valid analysis. We would like to note that the focus, in this paper, on vertical correlations does not imply that we assume that all errors are driven by vertical coupling. Likewise, in classical 3CH, it is not assumed that that there are no vertical correlations, they are just not analyzed.
Reviewer # 3 is right in that there can be error cross correlations between the data sets that we are not aware of. We have chosen 3 data sources, obtained with completely different techniques, as different as they can possibly be within available data sets. On this background assume that the cross correlations are negligible. However, we should add that the mechanisms that produce vertical correlations in refractivity profiles might include processes that depend on horizontal coupling or horizontal representation. For instance, ERA5 and RO are both sampled over a horizontal range of the order of 100 km, which may imply that some errors arising from horizontal coupling could be cross correlated in the two data sets.
We have included this point in the paper on line 76, where we have substituted: "Hence the ε^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 ε^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."

Line 41: in light of the earlier sentences in this paragraph, are we to assume that the ERA5 and RS92 error covariance matrices contain off-diagonal terms only because of vertical error correlation? Please clarify.
Answer, line 41: No, as it is also stated above, we make no assumptions about the origin of the vertical error correlations. The G3CH approach presented does in fact allow correlations in any spatial or temporal dimension to be examined, but we only examine vertical error correlations. It is a perfectly valid and mathematical sound approach which is being used widely in NWP and satellite retrieval: For an atmospheric profile one can define the covariance of errors e_i, e_j of some property at two altitude levels z_i and z_j, $\mathrm{cov}_{ij} = < e_i e_j >$. Horizontal error correlations may exist for instance between two ERA5 profiles at different locations, but that does not prevent one from studying the vertical error correlations separately.

Line 58: this sentence is not understood. Vertical footprint for RO profiles will be of order 10 km, which is much less than the distance between RO measurements for the data sets considered here.
Answer, line 58: Reviewer # 3 refers to this sentence: "The vertical footprint will typically be larger than the distance between data points." We see how the sentence can be misunderstood, and we have rephrased it to: "The vertical footprint will typically

be larger than the distance between the vertical height levels which the data values refer to." We remark that the vertical footprint of RO profiles is found in the paper to be of the order of a few hundred meters.

Line 63: we suggest the authors add a figure to the paper that defines precisely what is meant by “footprint” for the data sets and for truth. This same figure should clarify the term “observation grid” (Line 69), since the observations are available at random places and times, and not on a grid.

Answer, line 63: We realize that the term "vertical footprint" may stir some confusion here, and this issue was also raised by Reviewer # 1. We have tried, on line 56, to be more concise in the definition of this term with this formulation (and 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.*"

Answer, line 69: The word "grid" is used here in the meaning of vertical levels on which the refractivity is expressed. It has nothing to do with horizontal representation. We have changed line 69: " ... as it is being mapped to the vertical observation grid."

Line 66: assuming that systematic errors are removed, i.e. errors have no bias, is a confusing aspect of this paper. While it is true that the authors remove global means from the data sets, this does not imply the errors as analyzed contain no bias. The reason is that the authors use data subsets in the analysis (e.g. latitude subsets, collocation-distance subsets, etc.) and these subsets may contain bias. An example would be lack of global bias arising because there are equal and opposite biases in the northern and southern hemispheres. The authors need to consider the possibility of biases in subsets of the data. If they take this into account, the analysis can proceed apace.

Answer, line 66: We have not been clear about this, so it is valid critique. As we explain in the response to Reviewer # 2, 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."

And in line 126: "Systematic differences between data subsets are in practice removed by subtracting the mean of each data subset prior to the analysis.”

But still we pool the Northern and Southern Hemispheres, and also for instance land and ocean, so within each subset there may be groups of data that contain different observation biases. So indeed, if one data set contains separate bias regimes, this will lead to an overestimate of random uncertainty for that variable. We cannot do anything about that with the data at hand, except for requiring larger data sets.

Line 69: we again recommend a figure be used to carefully define how the truth data set is “distorted” when mapped to the observation grid, and to carefully define what is meant by “representativeness error”.

Answer, line 69: We cannot provide a drawing, but we can explain: The word "distortion" is chosen to include multiple sources of representativeness error, all contributing to the inability of the observation to represent the truth. Here the definition of representativeness error differs from the NWP definition, which — as also pointed out by Reviewer # 1 — attributes the model representativeness errors and the forward model errors to the observation representativeness error. The representativeness errors are composed of several components, including limited resolution, spatial and temporal differences in sampling, processing errors, interpolation errors and errors caused by smoothing. They come about in very different ways in the three data sets. For instance, by far most of the RO observation error comes from representing the radio occultation as a 1 dimensional profile. In the processing a crucial assumption of spherical symmetry neglects horizontal gradients in the atmosphere. This component is regarded as a representation error.

Line 83: based on the writing so far, "error cross correlations" are error correlations between data sets. It is not immediately obvious to this reviewer how finite footprints of the data sets would lead to such correlations. (Again, the suggested figure might help here). For example, if the data sets are not overlapping in space, why would finite footprints lead to error correlation? Also, please clarify whether the footprints alluded to are horizontal, vertical or both. This question is posed because in several locations of the paper it is implied that vertical correlation is what leads to non-zero off-diagonal covariance, so vertical footprint would be relevant.

Answer, line 83: It is instructive to think of the following thought experiment: Consider two identical vertical profiles $\{x_i, y_i\}$, with good resolution, that is with small vertical footprint, and with no errors. For simplicity, one may think of the truth as a profile with the same resolution. That is, we have two profiles identical with the truth t_i. Suppose that we now smooth these two profiles with the same filter. The smoothing introduces an error, ϵ_i, which is identical for the two profiles $\epsilon_i = x_{\mathbf{smoothed},i} - t_i = y_{\mathbf{smoothed},i} - t_i$. Now consider a set of such paired profiles $\{x_i, y_i\}_{i=1..N}$ sampled across the globe, that is we have two extremely well collocated data sets $\{x_i\}_{i=1..N}$ and $\{y_i\}_{i=1..N}$. Let these two data sets be smoothed with the same filter, such that they have similar vertical footprints. The smoothing will cause both systematic and random error. Let us assume for simplicity that the systematic error is zero such that $1/N \sum_i \epsilon_i = 0$. When we calculate the error cross correlation by averaging over the set of pairs we still get a non-vanishing error cross covariance $\mathrm{cov}(x, y) = 1/N \sum_i (x_{\mathbf{smoothed},i} - t_i)(y_{\mathbf{smoothed},i} - t_i) = 1/N \sum_i \epsilon_i \epsilon_i$. We hope that this extreme case can help to illustrate how similar footprints can cause error cross correlations between two data sets.
We can see how the distinction between vertical and horizontal footprint is assumed to be implicit from the context some in the paper, and as such not stated clear enough. We have made sure that "*vertical footprint*" is stated explicitly throughout the paper.

Line 120: Please clarify the notation. Do the different epsilon terms (x,y,z) each decompose into components I, R, C, X in equation (1)? We assume that the vector here represents different values along the vertical dimension. That could be stated explicitly.

Answer, line 120: Yes, Sec. 1.2, describes the error components of all three data sets. The formulation "For a given refractivity data profile, x,..." is not clear. We will change this to "For a given refractivity data profile from either of the three data sets RS92, RO, or EAR5 ..."

Lines 127-128: we have remarked earlier how the bias free assumption may not apply. Can the authors verify they have removed bias from all data subsets they have worked on? Subtraction of one global bias will not guarantee there are no biases in subsets of the data. Also, the statement that randomness implies “bias free” is mis-leading. Please modify this statement.

Answer, lines 127-128: True, it has not been stated clear enough that we remove biases for each subset individually. So this sentence:

"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 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."

Line 156: we raise again the concern that all data subsets might not be bias free. Has this been confirmed in the analysis?

Answer, line 156: As also mentioned above, each subset has been centered such the its mean is zero prior to G3CH analysis.

Line 167: we ask the authors to define the “vertical footprint of truth”. The figure asked for earlier would again help here. ....

Answer, line 167: For heuristic reasons we think it is convenient to think of the truth as referring to a certain vertical scale. However, to avoid further discussion we have removed this phrase: ".... differing from the vertical footprint of $\mathbf{t}$"

Line 167: why wouldn’t similarity of horizontal footprints also lead to cross-correlated errors?

Answer, line 167: That is a good point. As also mentioned in "Answer, line 39", there can be a cross-correlation arising from similarity between e.g. ERA5 and RO in horizontal footprint. We have a method to assess the vertical footprint and eliminate cross-correlations arising from similarity in vertical footprint, but we have no way to access the error cross correlation arising from similarity in horizontal footprint. This is a limitation of the method, and we have extended the formulation:

"... if two data sets have similar vertical footprints, or if they are sampled at similar horizontal scales, these two data sets may have cross-correlated errors, and possibly biases. All biases are removed prior to application of G3CH, but the error cross correlations introduced by finite vertical footprints or similar horizontal scale may influence the result of G3CH.

We also added a this statement at line 78:

"Hence the ε^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."

Line 171: what is meant by “common grid”? What are the spacings of this grid?

Answer, line 171: The common vertical grid is 247 levels defined in impact height, written explicity in the appendix of (Lewis 2009).

Line 180: it would be useful to clarify the mathematical relationship between the uncertainty estimate and the footprint. What assumptions are made to derive this relationship?

Answer, line 180: The relationship would depend on the vertical correlation function. An example: Suppose we have a profile with Gaussian uncorrelated random noise ε, such that $\langle\varepsilon\rangle = 0$ and $\langle\varepsilon(z)\varepsilon(z')\rangle = \sigma^2\delta(z-z')$. If we smooth this profile with a Gaussian filter of half width L, we get for a given altitude level, z_i, a vertically correlated error $\varepsilon_i = \frac{1}{L\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(-(z-z_i)/2L^2)\varepsilon(z)dz$. If we now calculate the variance of ε_i we get $\sigma_i^2 = \frac{\sigma^2}{L\sqrt{4\pi}}$, which scales as $1/L$.

Line 199: I believe what is being stated here is that G3CH is incorrectly assigning collocation error to RS92 error. We expect then, that if ERA5 is collocated to RS92 rather than RO, the RO would show the large uncertainty. Is that the case?

Answer, line 199: Yes that is true, for the raw uncertainty estimates.

Line 218: is there a way to justify this interpretation using a mathematical model and showing it mathematically? Otherwise, it’s difficult to assess the validity of this interpretation.

Answer, line 218: We can define the procedure of letting the area go to zero and define the collocation corrected error covariance matrix as the error covariance matrix in that limit in that limit. We have added this to section 1.2: "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 $\boldsymbol{\varepsilon}^R$."

Line 239: please refer back to the equations where this error covariance is defined. See the earlier comment about how the errors break down into the different components. One could, for example, insert those components into the covariance equations (7) and identify specific outcomes depending on the properties of these error components.

Answer, line 239: We can unfortunately not refer directly to eq.7, since the presented matrices have been collocation-corrected after the G3CH was applied.

Line 270: see question raised earlier of how “footprint of the truth” is defined.

Answer, line 270: We have removed this sentence: "In the derivation of G3CH representativeness is defined with reference to given scales in space and time of the truth."

Line 275: the concept of “physical variability” is introduced here for the first time. How does it relate to the error components I, R, C, X defined earlier? Or is it a new component of error? In general, the statistical properties of the error distributions are not explicitly described (are they gaussian?) except that they are mean zero. If statistical error distribution is not relevant, and any mean-zero error distribution is acceptable, it should be stated.

Answer, line 275: Physical variability refers to the underlying truth, which has "physical" correlations in it; temperature at altitude $\mathbf{z}_i$ is correlated with temperature at altitude z_j. This is not to be confused with error correlations.

Answer, line 275: The theory applies for any zero mean distribution. There is probably some obscure counterexample which from a mathematical point of view breaks this clause, so please read it as "any reasonable zero mean distribution". We have added this sentence on line 127 "Besides this no assumptions are made about the particular shape of error distribution functions."

Line 295: The Rieckh paper uses RO, radiosondes and analyses and forecasts. Please be more explicit why these data sets are not suitable for 3CH analysis, since they appear to be similar to the data sets used in this paper.

Answer, line 295: None of the many refractivity uncertainty experiments they do are really free from error cross correlations between the data sets because they either use two models (which have assimilated the same observations) or radiosondes and model analysis in the same triplets. We have changed the formulation (a bit simplified, since we cannot go into a deep discussion of the Rieckh paper): "... are less well suited for 3CH analysis than the data sets used here, because the errors of the used ERA5 analysis fields must be expected to be correlated with errors of other data sets"

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.