Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC4
https://doi.org/10.5194/amt-2022-121-RC4, 2022

Comment on amt-2022-121

Anonymous Referee #3

Referee comment on "Estimation of refractivity uncertainties and vertical error correlations in collocated radio occultations, radiosondes, and model forecasts" by Johannes K. Nielsen et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2022-121-RC4, 2022

Paper Summary:

The authors develop and apply a variation of the three-cornered hat (3CH) method of analysis for three independent collocated (to within some limits) refractivity data sets: 1) from Metop-A, B, and COSMIC-1 radio occultation (RO); 2) from RS92 sondes as processed by GRUAN; and 3) from ERA5 forecast fields. The 3CH analysis is generalized to include off-diagonal elements such that certain error correlations are accounted for. Various sources of uncertainty are defined and analyzed. Derived RO uncertainties are reasonably consistent with previous "first principles" error analyses for RO. Altitude and latitude dependent uncertainties are presented, to gain insight into the RO data set and improve assimilation of RO into numerical weather models.

Review Summary:

The paper has the potential to be an original and useful advance in analysis of RO error characteristics. The generalization of 3CH (called "G3CH") is potentially very valuable. However, the paper suffers from lack of clarity in 1) the definition of certain assumptions, and 2) how certain conclusions are reached. The paper requires major revision to improve and clarify the presentation. Clarifying certain assumptions might alter conclusions of the paper. It is difficult to be certain. In any case, the paper can be greatly improved with an altered presentation. Detailed comments follow.

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.

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.

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.

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.

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.

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.

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".

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.

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.

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.

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

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

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

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

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?

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?

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.

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.

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

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.

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.