General comments:

The paper compares radio occultation GPS/MET data retrievals from UCAR/CDAAC dual-frequency processing in 2007 with a newly processed GPS/MET dataset from JPL using single-frequency processing. The datasets partly overlap, and both contain periods with AS-on and AS-off. Together the datasets provide more data than before in the GPS/MET period of 1995-1997. Since the JPL dataset relies on L1 single-frequency processing, it is insensitive to the L2 encryption (AS-on). On the other hand, the single-frequency processing is affected by psudorange noise. All datasets are compared to MERRA-2 reanalyses in the 10-30 km range, concluding that errors at some altitudes and latitudes in MERRA-2 are likely larger than the errors from either approaches of retrieving the refractivity from the GPS/MET data. Thus, it is likely that future reanalysis projects could benefit from single-frequency processed GPS/MET data.

At times I found it difficult to understand exactly how the single-frequency processing is done, and I think a few additional equations may help (see specific comments below). Also, parts of the explanations seem to be contradictory, but I think it is just a question of being more precise in the language.

At the end, a possibly use of single-frequency processing is discussed, suggesting that it may provide some information on the magnitude of the residual ionospheric error due to raypath separation in dual-frequency processing. I have some doubt on that, partly because I don't think the single-frequency processing avoids raypath-induced residual ionospheric errors (see specific comment below), and partly because I think errors from pseudorange noise may be too large to be able to conclude much about residual ionospheric errors. However, I think it would be relatively easy to estimate the errors in bending angle due to the pseudorange noise, and I suggest a simple way to do that at the end of this review. This could then be compared to the expected size of residual ionospheric errors.

We thank the reviewer for a careful reading of the paper and the corresponding comments, which we address below, in an attempt to improve the manuscript.

Specific comments:

There are many acronyms in the abstract that are not defined. I assume they should be defined, but I'm not sure of the journal's policy on this.

Thank you. We agree that undefined acronyms are not desirable. We have inserted definitions.

Page 2, line 35-37:

The order and logic seems reversed here. Wouldn't it make more sense to say that "...

combining observations from multiple missions ... must be addressed". And then "Therefore, ... measurement accuracy needs to be characterized ..."?

We agree with this point. We have modified the text (line 39-40).

Page 4, line 125:

Would it be more correct to say "A benefit of the single-frequency data set is that we can assess the impact on the dual-frequency data set ..."? Was the CDAAC processing different (e.g., different filtering methods) in AS-on periods than in AS-off periods?

We appreciate this comment. We have modified the text on Line 133 in the updated manuscript.

Page 5, line 153:

"... are of equal but opposite sign". I suppose it should be 'size' after 'equal'.

We have clarified the meaning of this sentence (line 163).

Equation 1:

What is actually meant by 'delay' here? Is L (disregarding noise) the same as the optical path length minus the distance between transmitter and receiver?

Yes. We have clarified the meaning on Line 168 of the modified manuscript and cleared up some notational issues regarding what is referred to ionospheric delay.

Equation 3:

I think the signs on mu_1 and nu_1 should be opposite of what they are in this equation. Otherwise the math does not work out.

Thank you for catching this error. We have corrected it.

Page 6, line 170:

Why not just add (1) and (2) and divide by two? Wouldn't that give the same without the intermediate calculation of I? Or is there another reason for the intermediate calculation? I think it becomes clear later why, but that should perhaps be mentioned already here.

The purpose of equation (3) is to show how I is calculated from (1) and (2). Equation (3) shows how the ionospheric contribution is calculated from the range and phase measurements.

Adding (1) and (2) eliminates I, so it may be less clear as to how I is calculated. We have not modified the equations.

Page 6, line 175:

I do not understand this argument about the bandwidth. I am fine with that there in principle is no raypath separation at all between L1 and P1 measurements. These come from the same signal, and in the geometrical optics formulation at L-band frequencies there is only one ray path for that signal. Mathematically, L1 phase and P1 pseudorange relate to the integral of the refractive index and group refractive index, respectively, along this path. However, it is unclear to me if this avoids raypath-induced residual ionospheric errors in the further processing. When using eq.(3) to remove the ionospheric delay from L1 in eq.(1), you still have the integration along the original L1 path embedded in the eta term (as I understand the equations and the notion that this is a calibrated GPS signal, eta is the integral along the L1 path of the neutral atmospheric part of the refractive index minus the distance between transmitter and receiver). Thus, a part of eta depends on the ionospheric gradients because the path of the L1 signal depends on the ionospheric gradients. I don't think that is negligible in the context of residual ionospheric errors. How do you deal with this in the further processing? What are the equations that you use? I think these equations needs to be given, so that the results can be reproduced by others, and so that it can be understood to which degree raypath-induced residual ionospheric errors are present.

We appreciate that the reviewer has brought up these points. There are two main comments here: 1) bandwidth and 2) the eta term.

Regarding the bandwidth, our point is that the P-code signal can be decomposed into a signal of multiple frequencies, and is not characterized by a single frequency. The frequencies comprising the P-code signal span 10 MHz. Each frequency in this signal will travel a slightly different path, but this effect is negligible in comparison to raypath separation between L1 and L2. We have clarified the language near Line 188.

The comments regarding the eta term are well founded, but do not contribute to residual ionospheric error in the sense referred to in this paper. The ionosphere affects eta because bending in the ionosphere will slightly shift the tangent height of the raypath in the neutral atmosphere. Since ionospheric bending is quite small, this shift is similarly small and generally ignored in radio occultation processing, for either single or dual frequency processing. (See Hajj et al., 2002 for a processing description).

Page 7, line 191-195:

Here it becomes clear that the sampling rate is not the same for L and P. Is smoothing of I in eq.(3) and/or interpolation there not necessary? It is mentioned that smoothing is applied directly to P_1. Is that really correct? Don't you have atmospheric variation in P_1 that you don't want to smooth (eta in eq.(2))? Why not apply the smoothing to I in eq.(3) where the atmospheric variation has been removed? Well, I think it is when I read on. Revision in the text here is probably needed.

Thank you for this comment. We agree this section was not quite expressed correctly. We have made clarifications on Lines 220-224.

Page 8, line 225-226:

Is it correct to use the word bias here? If it is different from profile to profile (with different noise), it is not a systematic error, and thus not a bias. The approach seems good, although I am not an expert in filters and smoothing. There will of course be a residual error after the smoothing no matter how you do it, and there will be error correlations between adjacent points, but I don't think you can call that a bias.

We appreciate this comment. We agree that the term "bias" leads to confusion here. The use of the term bias here is not the same as we refer to elsewhere (e.g. climatological bias or systematic error as used in the reviewer's comment). We have introduced the additional modifier "statistical" (near line 256) to characterize the bias and included a reference where statistical biases due to weighted averages are discussed.

Page 8, line 234-235:

Here it is explained that the smoothing is applied to the ionospheric estimate in (3). I think that contradicts the information on page 7, where it was stated that smoothing is applied directly to P_1. It makes more sense if it is applied to I at 1-second intervals in (3). But I think the text on page 7 needs to be revised to be consistent with this.

We thank the reviewer for bringing a manuscript error to our attention. The smoothing is applied to the ionospheric estimate I (Eq. 3), not to P1. I is available at 1-sec cadence. We have corrected the manuscript (line 220).

Page 8, line 247:

I think I understand the approach, but I don't understand the sentence here. What is 'formal variance'?

We have clarified the sentence and added a reference (lines 290-294).

Page 9, line 262:

This is basically the same as just mentioned two lines before.

We have removed the repetition in the preceding lines (line 304).

Page 9, line 274-275:

Is it really smoothing of L2, or is it rather smoothing of L1-L2? In any case it is not the frequency that is smoothed, but the phase. And not the lower SNR that is mitigated, but rather the negative effects of it. The language in the paper could be more precise here.

We have modified the language to be more precise (line 319).

Page 9, line 315-317:

The sentence here does not really make sense to me. Could it be rephrased?

We believe sufficient explanation is made in the preceding sentences of the paragraph, so we have removed this sentence.

Page 11, line 332:

Could you provide one or more references to support "Other climate-related work" here?

References are now added at the end of the sentence (line 387).

Figure 4:

The figure shows median values. Would it be similar with mean values? Or are there a number of outliers that makes the median and mean significantly different?

Yes, due to outliers the mean and median can differ significantly, particularly for the small differences found in these comparisons. We believe that median is an acceptable metric for the comparison as it is a robust statistic.

Page 12, line 365:

Common profiles in three data sets? I suppose there are not common profiles between the CDAAC AS-on and CDAAC AS-off data sets. Either AS was on or it was off. It cannot be both at the same time for the same profiles. So should it rather be common profiles in two datasets here?

Thank you for this comment. We now refer to two pairs of profile sets and provide more clarity on why the differences can occur (near line 423).

Page 12, line 366-369:

Could the reason for the differences also be that these are median values? With mean values one would expect to be able to see consistency between differences in Figure 4 and Figure 6, since this is more or less linear algebra. But with median values it is a more complex and different story, and one cannot generally expect such consistency. It is difficult to see the differences in altitudes without a grid in the figures, but they seem quite similar (I do see a small offset at 10 km).

Thank you for this comment. This is a valid point. Even in the case of the mean, full consistency would not occur because differences are formed between the JPL and CDAAC data sets at altitudes that differ from the differences formed between the CDAAC and MERRA-2 data sets. We have changed the language to be more explicit and also suggest the point made about median differences (lines 425-434).

Page 13, line 390:

ECMWF-Interim? Do you mean ERA-Interim?

Corrected.

Page 14, line 449-450:

I was not able to find this estimate (1% near 30 km) in (Danzer et al., 2013). The number seems at least an order of magnitude too large. Danzer et al. (2013) show mostly errors/differences in bending angle, but at much higher altitudes. They show also the effect on temperature profiles (in their Fig. 8), but biases in temperature at 30 km can be very different from biases in refractivity, because of downward error propagation via the hydrostatic integration and large biases in the retrieved pressure. In the introduction of Danzer et al. (2013), they cite error estimates in previous works (Schreiner et al., 2011) of 0.045% at 30 km for refractivity, which sounds much more reasonable to me.

A couple of questions related to this: What would be the size of errors in bending angle (in micro-radians), that typical pseudorange noise could create in single-frequency processing? And how would this compare to expected residual ionospheric errors in dual-frequency processing? I think you would be able to answer these questions with the data that you have: You could take the derivatives with respect to impact altitude (in m) of the differences between the L1-L2 and CA-L1 fits in figure 3 (examples c, d, e, and f). That should give you four examples of bending angle errors (in radians) due to pseudorange noise between 15 and 60 km. I don't know the answer myself, but me feeling is that it will be difficult (even when averaging over many profiles) to say anything conclusive about the residual ionospheric errors using the single-frequency processing because of the pseudorange errors. In any case, it would be very interesting and very relevant to see example estimates (and perhaps also ensamble averages) of the bending angle errors from the single-frequency processing with this straight-forward approach. I strongly suggest such estimates to be included in the paper.

Citation: https://doi.org/10.5194/amt-2021-241-RC2

Thank you for this comment. We agree that an error was made in the paper. We have now corrected the error and added a figure and paragraph related to bending angle precision by comparison with JPL dual frequency processing for AS-off periods. See lines 538-558.