Both reviewers have recommended publication of the manuscript subject to minor changes. However, the responses to these comments are in a few instances not sufficient. Please address the comments identified below more thoroughly prior to publication. The editor's comments to which the authors should respond are indicated in blue.

Page 13, line 10: The data in figure 5 has already been binned and then averaged over 9 days. Does the linear regression on the figure 5 data provide a reduced uncertainty relative to the data presented in figure 6? Errors on the fit should be included. I may have misunderstood, but don't both linear regressions use the same data (just one if further averaged into a diurnal)? Does the change in the regression slope as the data is averaged further suggest that the binning approach is biasing the correlation?

The binning is useful because there are occasional gaps in the time series (e.g., the morning of 14 July). Without the binning, the morning data is slightly "underrepresented" because of that gap. We have changed the caption as follows (including the fit errors):

"Figure 6. Correlation of ambient [XO_2] measured by ECHAMP with [HO_2*] measured by IU-LIF-FAGE. The linear fit is for data between 09:00 and 22:00, indicated by the points with green circles. The equation of the fit is [XO_2] = (1.08 ± 0.05) [HO_2^*] – (1.4 ± 0.3) ppt."

The reviewer asked several questions in this comment, but the response does not address the last, highlighted one above. The changes to the manuscript should include a response to this question.

Page 13, line 23: Although I appreciate that the authors do not know the reason why the measurements diverge on the 22nd, the possible explanation 'a transient interference in the HO2 measurement when sampling ambient air..' is rather vague. Could the authors elaborate on what they think this transient interference may be or what it may be related to?*

We agree that the explanation of a "transient interference" is vague, but feel that any possible reason offered at this point would be too speculative. We note that since HO_2* is measured as OH after conversion by reaction with NO, any interference in the OH measurement would affect the HO_2* measurements as well.

It is reasonable to state that the authors do not know the reason for the divergence of the measurements over a short period. The response does not make sense, however, since it first speculates that there is a "transient interference" in one measurement (without assigning a mechanism), but that any other explanation would be speculative (i.e, first speculates, then declines changes on the grounds that it would be speculative). If the authors are not able to follow the reviewer's suggestion as to the nature of the interference, then the speculation regarding the transient interference should simply be removed. The divergence can be noted, and the authors can also state that the reason for the divergence is not known.

Section 3.3: in general, there is a lot to consider when comparing HO2 and XO2 measured and modelled. The ratio varies with RO2 type present and calibration differences also need to be considered. A table detailing the measured HO2*, XO2 and XO2:HO2* and the 4 modelled HO2*, XO2 and XO2:HO2* on the individual days and campaign average would help to clarify the text.*

We hope that the majorly revised paragraph quoted earlier (starting with "A bi-variate linear regression...") has clarified these issues. Furthermore, the results from the 4 models are shown in the SI.

Can the authors provide the requested table, highlighted above? The comment has not been addressed.

The paper describes the measurement of peroxy radicals (HO2, RO2) with two different techniques. The LIF-FAGE technique by Indiana University was originally designed to measure solely HO2 radicals by chemical conversion with NO to OH, which is then detected by LIF. However, different experimental studies (including a study by authors of this paper) have shown that the technique is also sensitive to specific RO2 radicals with different sensitivities when the instrument is tuned for maximum HO2-to-OH conversion efficiency. The measured quantity is called HO2. The new ECHAMP technique, a chemical amplifier using ethane instead of CO, is designed to measure the sum of HO2 and all RO2 species. Due to different amplifier chain lengths for different radical species, the resulting quantity XO2 is a proxy for the total peroxy radical concentration. Comparing the measurements by the two techniques sounds like comparing apples with oranges. The present paper demonstrates that such a comparison can be done in a meaningful way, if the instruments are carefully characterized and additional information about the peroxy radical speciation is available (here from box model calculations constrained by measured trace gases). The direct comparison of the conceptually different calibration methods (photolysis of water vapor vs. photolysis of acetone) and the field comparison show that the measurement techniques yield consistent data within the specified experimental uncertainties. These findings suggest that the two described methods can also be used for meaningful tests of atmospheric chemistry models, if the measured peroxy radicals (HO2*, XO2) are appropriately simulated by the model by taking RO2-specific weighting factors of the instruments into account. This requirement should be explicitly stated in the conclusions.*

Furthermore, recent progress in the measurement of HO2 by LIF-FAGE instruments should be mentioned. It has been shown that the interference by RO2 can be avoided by reducing the concentration of NO that is used for conversion to OH (e.g. Fuchs et al., 2011; Whalley et al., 2013; Feiner et al., 2016; Tan et al., 2017).

This section has been re-written as follows:

"The sensitivity of the LIF-FAGE technique to each type of organic peroxy radical varies with the amount of NO added for the conversion and is instrument-dependent but in general is highest (up to ~90%) for β–hydroxy peroxy radicals derived from alkenes and lowest for those derived from small alkanes (Fuchs et al., 2011;Lew et al., 2018;Whalley et al., 2013). This RO_2 interference can be greatly reduced by use of lower NO concentrations or reaction times, yielding conversion efficiencies for isoprene-RO_2 lower than 20% (Feiner et al., 2016;Fuchs et al., 2011;Tan et al., 2017;Whalley et al., 2013)."

It is not clear that the change addresses the comment highlighted by the reviewer above. An explicit statement that the comparison between the two techniques is meaningful if an appropriate model simulation demonstrates them to be comparable should be included. Alternatively, if the authors disagree with the comment, they should state their reasoning.

(2) In the experimental section, the authors point out that the use of ethane instead of CO offers advantages. Safer operation is obviously a plus. However, I don't understand

why the choice of ethane reduces the sensitivity on relative humidity. Is this due to the reduced chain length? Is there evidence for water influence on the OH+CO reaction? To my knowledge, the water effect has been attributed to the reaction HO2+NO (e.g., Mihele et al. 1999, Butkovskaya et al., 2007). Why is the amplification factor lower,if ethane is used? Another advantage of ethane could be mentioned, although it is probably not relevant in a forest environment. Ethane avoids possible interferences from ClOx, which can lead to amplification in CO/NO systems (Perner et al., J. Atmos. Chem. 34, 9, 1999).

We have added the following text to briefly clarify the important issue of RH-dependence:

The cause of the RH-dependence of the CO-based amplification chemistry is the RH-dependence of the main radical termination step: the reaction of HO_2 with NO to form HNO_3 (Butkovskaya et al., 2007;Butkovskaya et al., 2005;Butkovskaya et al., 2009;Mihele et al., 1999;Reichert et al., 2003), with a smaller contribution from the RH-dependent wall losses of HO_2. These two RH-dependent radical termination steps affect the ethane-based amplification chemistry as well, but the most important terminations steps are from the formation of ethyl nitrite and ethyl nitrate – neither of which depends on relative humidity.

A water vapor dependence in the reaction of peroxy radicals, particularly HO_2 + NO, has been invoked to explain observed water vapor dependences in chemical amplifiers. Nevertheless, this is not the "main radical termination step" in such amplifiers, unless I misunderstand, but rather the one that leads to a water vapor dependence. The authors may want to consider rephrasing.

The box model was constrained with 30 minute average mixing ratios. As peroxy radicals show a strong non-linear dependence on NO, using 30 minute average values as constraint can lead to systematic bias in the model results. I would like to see the model results that are averaged to 30 minutes after the model has been run at the much higher time resolution of the NOx measurements.

The time resolution of the model is limited by the 90-minute frequency of the VOC measurements which we have interpolated to values every 30-minutes. Thus we are unable to run the model at higher time resolution.

This response does not make sense. If it is possible to interpolate from 90 to 30 minutes, then it is also possible to interpolate to a faster time scale. It cannot be the case that the authors are thus “unable” to run the model at higher time resolution. If the authors feel there is nothing to be gained in doing so, that would be an acceptable response, and the authors should make this case instead.

(6) Figure 4 - 6: Is it meaningful to adjust the result of the linear regression for the calibration difference (section 3.1)? This would only make sense, if the calibration would be done for the same peroxy radical speciation as encountered during the measurement days in the field.

We have intentionally included in the text both the “raw” regression/ratio results and those corrected for the calibration difference. Since both ECHAMP and LIF-FAGE are both sensitive (high α) to HO_2 and isoprene RO_2, we do think that “correcting” the comparisons for the 20% calibration difference helps to frame the discussion of the differences between the two measurements.

It is not clear why the word “correcting” is in quotes. Again, the response does not appear to make sense. Either the correction is justified, or it isn’t, but the justification should not include “framing the discussion.” A simpler response that simply states the justification for the correction is all that appears to be required.