Response to Referee #2

We thank the reviewer for the useful suggestions to improve the paper. The comments of the referee are repeated in bold letters while our response is given in normal text.

According to the comments of both referees, we changed to title of the paper, replaced X by [X], and added error bars to the TIMED/SABER observations in Fig. 1-5.
We further carried out sensitivity runs with different sets of Einstein coefficients and included a new Figure 2. We also increased the uncertainty of the Einstein coefficients and added uncertainties of SABER temperature, SABER OH(9-7)+OH(8-6) VER, and SABER O3, resulting in larger total uncertainties of [O(3P)] and [H]. The discussion of potential error sources of [O(3P)] was also extended.
The rate of OH(v=8)+O(3P) was reduced in order to obtain physically allowed [O(3P)] values, which are slightly lower than in the previous paper version.
Finally, a detailed comparison between the [O(3P)] derived here and [O(3P)] from other studies is also included in the section "Conclusions" and we explicitly state that out [O(3P)] should be regarded as an upper limit.

General comments:

This study proposes a new OH airglow model to retrieve O and H densities in the mesosphere. The OH model is empirically developed to simultaneously fit four OH emissions observed by the SABER/TIMED instrument at 2.0- and 1.6-microns as well as the OH(6-2) and OH(3-1) bands measured by SCIAMACHY/ENVISAT. The authors show that using adjusted rate coefficients and specific state-to-state relaxation mechanisms, the OH model reproduces the four emissions. However, they retrieve very high O and H concentrations.
The concept of fitting four emission bands simultaneously is promising as it may constrain unknown parameters involved in modeling OH emissions. The conclusions regarding their new OH model, however, are speculative as model inputs used in the study to simulate emissions have larger uncertainties than the authors claim (i.e. Einstein coefficients, ozone concentration). Accounting for these uncertainties will significantly alter the results of this paper.

Further, the authors show that the applied OH model retrieves unrealistically high atomic oxygen and hydrogen in the MLT. Recent publications (s.f. Kaufmann et al. 2014, Mlynczak et al. 2018, Panka et al. 2018, and Zhu and Kaufmann 2018) have shown that [O] densities retrieved using SABER and SCIAMACHY measurements are much lower (by up to a factor 2 and more) than those retrieved in this study, specifically from 85-100 km. The model development (and the rate coefficient adjustments) must have a goal to reliably retrieve atmospheric properties from the observations. The very high [O] and [H] retrieved with the help of the new model indicates that there are still major flaws (it does not matter that it fits all selected emissions, this system has a very large number of unconstrained variables). The paper needs major revisions before making physical sense and being suitable for publication.

We lowered the total rate of OH(v=8)+O(3P) and checked the validity of our [O(3P)]. Taking into account the uncertainty of O(3P) derived here, our [O(3P)] results are now physically justified between 80 km and 95 km based on radiative and energetic constraints reported in Mlynczak et al. (2013b). The estimation of total [O(3P)] uncertainty was extended by additionally including uncertainty of SABER

O3 (~10 %; Smith et al., 2013), SABER temperature (3 %, Garcia-Comas et al., 2008) and SABER OH airglow emissions (6 %).

We also carried out sensitivity runs with different data sets of Einstein coefficients to test their impact on our results. As a consequence, the uncertainty of Einstein coefficients was increased from 10 % to 30 %.

We added an extended comparison of our model results with recent results presented by Mlynczak et al. (2018), Panka et al. (2018), and Zhu and Kaufmann (2018). We further explicitly stated that our [O(3P)] should be viewed as an upper limit in the altitude region above 87 km. However, despite that our [O(3P)] results are larger than the [O(3P)] in these three studies, our results are physically allowed, while conditions of chemical equilibrium of O3 are also valid. Additionally, we did not claim that our results are undeniable truth. But we think that our study makes justified assumptions which is still of value for publication.

A detailed response is presented below, when answering your specific comments.

Instead of fitting four emission bands while simultaneously retrieving [O] and [H] densities, I recommend for the revised study, to concentrate on retrieving [O] and [H] densities but only fit three emission bands (as will be discussed below, the OH(6-2) emission band is unreliable and taking into account its large uncertainty will alter the results of the current study).

We agree that SCIMACHY OH(6-2) VER is relatively noisy. But this issue was taken into account by the relatively large error bars of these measurements and even these relatively large uncertainties cannot explain all the differences of OH(6-2) VER between the Base model and SCIAMACHY measurements. Additionally, the major impact of OH(6-2) VER on our model results is the suggestion that OH(ν≥7)+O2 primarily contribute to OH(ν≤5)+O2.

However, the corresponding changes in the OH model do account for an increase of [O(3P)] and [H] of about 10 %. And the impact on the derived [O(3P)] profile also decreases with increasing altitude. The large [O(3P)] values above 90 km are primarily caused by different OH(v)+O3P rates and the assumed deactivation paths. Thus, we kept SCIAMACCHY OH(6-2) VER observations in our study.

Further, the authors must demonstrate how the rates derived from zonal mean profiles fit real single scans in three emission bands.

This was done in Sect. 3.4 (l. 593-618) where we analyzed the relation between [O(^{3}P)] and OH(9-7)+OH(8-6) VER presented in Fig. 7.

Specific comments:

Line 1.
The title states "New insights in OH airglow modelling…". The proposed new "insights" are highly speculative and are inconsistent with existing theory and experiments. The authors need to first show that reliable [O] and [H] can be derived when their OH model is applied before claiming any new insights.

The title was changed to: "Model results of OH airglow considering four different wavelength regions to derive night-time atomic oxygen and atomic hydrogen in the mesopause region"

In the introduction section, discussion regarding the current progress of [O] and [H] retrievals using SABER and SCIAMACHY instruments is missing. Retrieving these two parameters are a key point of this study and no background is given. Please cite recent [O] and [H] retrieval studies and their proper discussion.

We added (l. 100-117):

"The newly suggested rates of OH(ν)+O(^{3}P) were applied in different models to derive [O(^{3}P)] in the mesopause region. Mlynczak et al. (2018) used SABER OH airglow emissions observed at 2.0 µm to derive [O(^{3}P)] and assumed rates of 3.0×10^{-10} cm^3 s^{-1} and 1.5×10^{-10} cm^3 s^{-1} for OH(ν=9)+O(^{3}P) and OH(ν=8)+O(^{3}P), respectively. They further stated that deactivation of OH(ν=9)+O(^{3}P) has to occur via single-quantum quenching and that the OH(ν=8)+O$_2$ rate has to be smaller than known from laboratory measurements to get global annual energy budget into near balance. Panka et al. (2018) simultaneously investigated SABER OH airglow emissions measured at 2.0 µm and 1.6 µm, while applying faster rates for OH(ν=8)+O(^{3}P) and OH(ν)+O$_2$. Their [O(^{3}P)] values agree within the corresponding errors with the results reported by Mlynczak et al. (2018) above ~87 km but are larger in the altitude region below. The authors also demonstrated the high sensitivity of the derived [O(^{3}P)] from O(^{3}P) quenching rates applied in their model. Zhu and Kaufmann (2018) analyzed SCIAMACHY OH(9-6) transition. They used a value of 2.3×10^{-10} cm^3 s^{-1} for OH(ν=9)+O(^{3}P) which is lower than the one applied in the two previous studies, resulting in generally lower [O(^{3}P)] values in the altitude region above 87 km. Their rate for OH(ν=9)+O$_2$ lies between the corresponding rates of the two other studies, and consequently their [O(^{3}P)] is also between the [O(^{3}P)] values of these two studies below 87 km. Thus, recent publications indicate that the rate of OH(ν=9,8)+O(^{3}P) might be slower than previously suggested in Sharma et al. (2015). But this problem needs further attention because all three papers derive different [O(^{3}P)], depending on the data sets investigated."

Lines 205-216.
The retrieval of [O] and [H] are both dependent on the [O3] volume mixing ratio. The authors used nighttime [O3] taken from SABER.
The nighttime SABER [O3] has never been rigorously validated, nor is there a paper discussing its retrieval approach. Differences between WACCM and SABER [O3] are roughly a factor of 2 (Smith et al. 2014) and, therefore, one cannot rely on SABER [O3] as an input parameter.
Additionally, in Mlynczak et al. 2018, the conclusion is made that current SABER daytime [O3] and, supposedly, the nighttime one is too high based on a significantly lower [O] retrieved in that study. It is clear from equation 4a and 4b that any variation in [O3] will have a significant effect on [O] and [H]. For the revised study, I recommend using inputs taken from a self-consistent photochemical model like WACCM instead of ones taken from retrievals, which are not supported by any other studies. Additionally, uncertainties in the retrieved parameters due to large uncertainties in the [O3] must be estimated and discussed.

It is correct that chemistry-climate models like WACCM optimally contain descriptions of the state-of-the art of all known processes and usually provide a quite realistic representation of reality. However, these model results do not describe the "true" state of the atmosphere at any given point in time and space. Consequently, a comparison of models results and observations might be used to validate a model, but certainly not the observations.

In particular, WACCM has a well-known deficit of odd-oxygen, so it is not surprising that the WACCM O3 is less than SABER O3 or O3 obtained from any other measurement. This is a long-standing issue with WACCM and its predecessors. The daytime SABER O3 excess is certainly clear between 60 and 80 km. It may also be above that, but both Mlynczak et al. (2018) and Smith et al. (2014) could not conclude that above 80 km the daytime SABER O3 is too large.

The same is true for SABER night-time O3. Mlynczak et al. (2018) state that since "the cause of the larger daytime SABER ozone is not known, it is also possible that the SABER night O3 is also too large." This means that they cannot exclude an overestimation of their SABER night-time O3 because they do not know the reason behind SABER O3 daytime enhancement. But SABER night-time O3 might not have the same problem as SABER daytime O3. There are two candidates for the daytime O3 enhancement: One is the interfering bands of daytime CO2 that are not properly accounted for. Secondly, there might be an out-of-band light leak in the spectral filter. Even if it is an out of band leak, it may not be an issue at night-time.

Furthermore, recent comparison between MIPAS O3 and SABER O3 (Lopez-Puertas et al., 2018) did not show conclusive evidence that SABER night-time O3 is generally too large between 80 km and 100 km.

Thus, we think SABER night-time O3 inside the time-space interval of interest is not less reliable than any other data set and we did not replace SABER O3. The corresponding sentences in the paper were rephrased and an uncertainty of 10 % of SABER O3 (Smith et al., 2013) was assumed and included in the calculation of the total [O(3P)] uncertainty.

We added in the paper (l. 549-556):
"Recent comparisons between MIPAS O_3 and SABER O_3 derived at 9.6 µm were performed by Lopez-Puertas et al. (2018). The authors showed that night-time O_3 from SABER is slightly larger than night-time O_3 obtained from MIPAS in the altitude region 80-100 km over the equator (their Fig. 8 and 10) but these differences are within the corresponding errors. Thus, at least to our knowledge there is no conclusive evidence stating that SABER night-time O_3 is generally too large. Nevertheless, we considered an uncertainty of O_3 of about 10 % (Smith et al., 2013). The uncertainty of SABER temperature was estimated to be lower than 3 % (Garcia-Comas et al., 2008) while the total uncertainty of SABER OH(9-7)+OH(8-6) VER was assumed to be about 6 % (see Sect. 2.1.2)."

Lines 296-327.
"…we exclude the Einstein coefficients as a potential fundamental error source."
I do not agree with this statement. The new constraint imposed using the OH(6-2) emission band is unreliable. This band has a very small Einstein coefficient. The authors do not go into detail regarding the numerical differences among the literature of the OH(6-2) emission rate, but state that they are consistent.
The authors use the OH(6-2) coefficient taken from Xu et al. 2012 which is 1.767 sec-1. A more recent publication by Brooke et al. (2015) recalculated OH Einstein coefficients and found a rate of 1.16 sec-1 for the same transition. The rate of Xu et al. (2012) is approximately 50% larger than that of Brooke et al. (2015) and would significantly change the OH(6-2) emission profiles in Figures 1-5 as well as the results in Tables 2 and 3. The ab initio calculations of van der Loo and Groenenboom (2007, 2008) give values that are even smaller than Brooke et al. (2015) - the OH(6-2) emission rate of Xu et al. (2012) is 75% larger than that of van der Loo and Groenenboom (2007, 2008). Evidently, the issue of the OH Einstein coefficients is not yet settled.
The problem of all these data sets of Einstein coefficients is that the results strongly depend on how good the representation of the Hamiltonian for the OH molecule is which is used to solve the Schrödinger equation. It is expected that the calculations improve with time, but not necessarily at these large quanta changes. Multi quanta transitions of more than 2 quanta have small Einstein coefficients and are generally hard to model and calculate.
Also, it is inappropriate solely focusing on the Einstein coefficient of OH(6-2) because errors of this single transition might be partly compensated by errors of other OH transitions. However, we agree

with the referee that different Einstein coefficient data sets have to be taken into account before excluding them as a potential error source.

Thus, we carried out sensitivity runs and the results are displayed in Fig. 2 (see next page). We also rephrased and extended the corresponding section in the text as follows:
"Since the overestimation of the Base model is especially large for OH(6-2) VER, an impact of the Einstein coefficient of the corresponding transition must be considered. Regarding this aspect, we have to point out that studies based on HITRAN 2004 data set should be viewed more critically, because of erroneous OH transition probabilities. The Einstein coefficients used in this study were recently recalculated (Xu et al., 2012, their Table A1) and correspond to a temperature of 200 K, which is very close to mesopause temperature. Furthermore, these Einstein coefficients are consistent with the values of the HITRAN 2008 data set (Rothman et al., 2009). However, there are several other data sets of Einstein coefficients found in literature that might lead to different results. We therefore carried out sensitivity runs, using the Einstein coefficients reported by Turnbull and Lowe (1989), Nelson et al. (1990), van der Loo and Groenenboom (2007), Xu et al. (2012; =Base model), and Brooke et al. (2016). The corresponding results are presented in Figure 2 and show considerably large differences in case of OH(6-2) VER which are about a factor of 4 between the highest and lowest model output. In contrast, the individual simulations of OH(5-3)+OH(4-2) VER and OH(3-1) VER are rather consistent and vary only by ~10 %. These results emphasize that the choice of the Einstein coefficients is a potential error source for higher quanta transitions.
Regarding the credibility of the Einstein coefficients, it is generally assumed that the calculation improve with time. However, this is not necessarily true at quanta changes >2 because it all depends on how good the representation of the Hamiltonian for the OH molecule is, that is used to solve the Schrödinger equation. Multi quanta transitions >2 quanta have small Einstein coefficients and are generally hard to model and calculate. The assessment of the Einstein coefficients requires a detailed analysis of the corresponding calculations, which is beyond the scope of this study. We therefore cannot exclude the values used in the Base model as a potential error source, but we also think that our choice of the Einstein coefficients from Xu et al. (2012) is reasonable. Additionally, these values represent approximately the average model output of all five data sets considered here, while the model results based on Nelson et al. (1990) and van der Loo and Groenenboom (2007) represent the variability. Thus, we will not replace the Einstein coefficients by Xu et al. (2012) in our model but keep in mind that they might be too large.
Furthermore, the best agreement between the observations and the model was obtained by applying the Einstein coefficients reported by van der Loo and Groenenboom (2007). But even in this case, the model still overestimates the observations of all OH transitions in the altitude region between ~80 km and ~86 km. This pattern strongly supports the suggestion stated above that the rates and schemes associated with OH(v)+O_2 are incorrect."

We further added in the Conclusions (l. 635-637):
"Also note that the Einstein coefficients used here might be in error (see Sect.3.1; Fig. 2). This does not affect the two general conclusions drawn above but would impact the empirically derived rates."
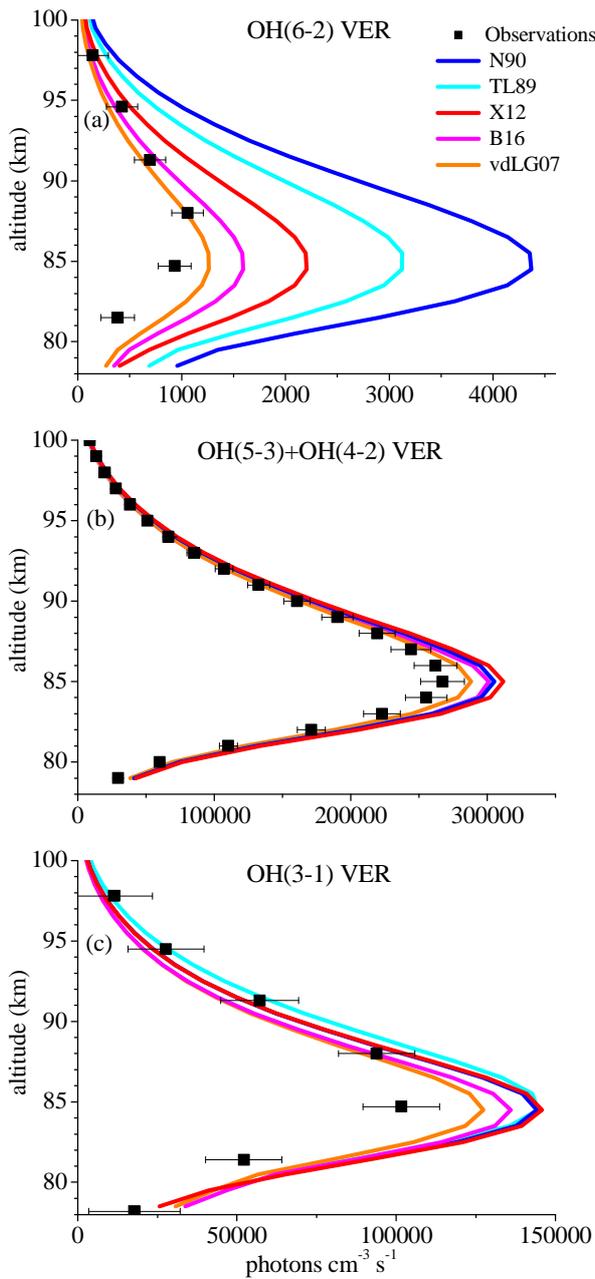
Figure 2 : Same as Figure 1 but for different sets of Einstein coefficients from literature, namely N90 (Nelson et al., 1990), TL89 (Turnbull and Lowe, 1989), X12 (=Base model; Xu et al., 2012), B16 (Brooke et al., 2016), and vdLG07 (van der Loo and Groenenboom, 2007).

To make matters worse, the SCIAMACHY OH(6-2) band displays a signal count two orders of magnitude smaller than that of the OH(3-1) band. The uncertainty of the VER signals for these two bands will be vastly different. Finally, because these two bands have a different Delta v and values of Einstein coefficients that are different by more than one order of magnitude, the uncertainty of the OH(6-2) band will be much larger.

Therefore, both the observed low VER signals and the large uncertainty in the value of the Einstein coefficient indicate knowledge of the OH(6-2) band is highly uncertain. As a result, the OH(6-2) emission band cannot be confidently used to constrain OH modeling parameters. For

manuscript revisions, I recommend to redo this study using only three OH emission bands. An alternative would be to validate the intensity of the OH(6-2) band by comparison with the OH(6-3) profile, which should be within the capabilities of SCIAMACHY.

We agree with the referee that SCIAMACHY OH(6-2) is relatively noisy but this was taken into account by the larger error bars presented in Fig. 1-5. These error estimates are based on the random noise at altitudes above the OH emission (l. 147-151) and we do not see any reason, why the OH(6-2) band should be omitted here because of low signal-to-noise ratio.

Additionally, including OH(6-2) does not considerably affect the [O(3P)] derived here (~10 %) and the impact also decreases with increasing altitude. However, based on the OH(6-2) emissions, we suggest that OH(ν≥7)+O2 primarily contribute to OH(ν≤5)+O2.

Thus, we did not exclude SCIAMACHY OH(6-2) from our study. Whether future studies will obtain similar results or strongly disagree with our suggestion cannot be known. But suggesting a new idea is not wrong even when based on relatively noisy data because these uncertainties were considered.

Lines 574-592.
"Applying their suggested limit, we found that in our case chemical equilibrium of O3 is probably true only above 80 km."
Recent studies have shown that the [O] and [H] retrieval approach used in this study may be flawed (Belikovich et al., 2018; Kulikov et al., 2017, 2018) and can introduce additional uncertainties. The authors addressed these issues very briefly here, but this needed to be more rigorously discussed. To say just simply "probably true" is insufficient. Additionally, uncertainties of the final results related to a probable chemical equilibrium breakdown need to be estimated and discussed.

This section was rearranged and extended as follows:

"The second aspect influencing the quality of the derived profiles is the assumption of chemical equilibrium of O_3, represented by Eq. (3). This issue was recently investigated by Kulikov et al. (2018), which carried out simulations with a 3-D chemical transport model and demonstrated that a wrongly assumed chemical equilibrium of O_3 may lead to considerable errors of derived [O(^{3}P)] and [H]. In order to test the validity of chemical equilibrium of O_3 locally, the authors suggested that OH(9-7)+OH(8-6) VER has to exceed $10 \times G \times B$, with B including several chemical reaction rates involving O_x and HO_x species. Note that this criterion requires simultaneously performed temperature and OH airglow measurements. Furthermore, this criterion is based on the assumption that the impact of atmospheric transport on chemical equilibrium of O_3 is negligible. Since our experiments fit these conditions, we applied their suggested limit and found that in our case chemical equilibrium of O_3 is valid above 80 km. We have to point out that the term "chemical equilibrium of O_3" refers to O_3 that does not deviate more than 10 % from O_3 in chemical equilibrium (Kulikov et al., 2018, their Eq. 2). Assuming that O_3 is always 10 % above or below O_3 in chemical equilibrium introduces an uncertainty of about 10 % at 80 km and 20 % at 95 km, additionally to the total uncertainty of [O(^{3}P)] and [H] estimated above. However, such a worst case scenario is rather unlikely while it is more realistic that O_3 actually varies around its chemical equilibrium concentration. Thus, an over- and underestimation of derived [O(^{3}P)] and [H] are assumed to compensate each other. Consequently, we conclude that the impact on the total uncertainty of [O(^{3}P)] and [H] due to deviations from chemical equilibrium of O_3 is negligible, but only because the previously used criterion is valid."

Lines 615-617.
"… we think that the O(3P) and H derived by the Best-fit model provides reasonable results between 80 and 95 km."

The [O] derived looks somewhat reasonable only below 87 km, but not above this altitude. At 95 km, the retrieved [O] is at least two times larger than Mlynczak et al. [2018] and more than a factor of 5 at 100 km. It has also been discussed in detail that high [O] will disrupt the energy balance in the MLT (Mlynczak et al. 2013, 2018) and influence temperature retrievals. If, in the revised study, the retrieved [O] and [H] remain high, then please demonstrate how it impacts the heating and cooling of MLT and discuss in detail possible ways to overcome the corresponding energy budget imbalance.

We compared our [O(^{3}P)] derived here with the maximum [O(^{3}P)] physically allowed by radiative constrains (Mlynczak et al., 2013b) and had to adjust our model to derive lower [O(^{3}P)] values. This was done by reducing $k_{11}(v=8)$ from 2.3×10^{-10} cm^3 s^{-1} to 1.8×10^{-10} cm^3 s^{-1}. Now, our [O(^{3}P)] matches the upper limit suggested by Mlynczak et al. (2013b) within the corresponding errors.

Thus, we added in the text (l. 474-486):
"At least to our knowledge, the total rate of OH($v=8$)+O(^{3}P)→OH(v')+O(^{1}D) was not measured. Nevertheless, results reported by Mlynczak et al. (2018) and Panka et al. (2017, 2018) indicate that this rate might be slower than the value of 2.3×10^{-10} cm^3 s^{-1} suggested by Sharma et al. (2015). This is also in agreement with our findings here, because applying 2.3×10^{-10} cm^3 s^{-1} for $k_{11}(v=9,8)$ results in non-physical [O(^{3}P)] values above 90 km. The corresponding value of [O(^{3}P)] e.g. at 95 km is about 1.25 times larger than SABER [O(^{3}P)] 2013 (Mlynczak et al., 2013a) which in turn is about 1.15 times larger than the upper limit of [O(^{3}P)] (Mlynczak et al., 2013b, their Fig. 4). This results in a factor of $1.15 \times 1.25 = 1.44$ (=44 %) above the upper limit and cannot be explained by the uncertainty of the [O(^{3}P)] profile derived here (40 %, see Sect. 3.4). In order to obtain reasonable [O(^{3}P)] values, it was necessary to lower the rate of $k_{11}(v=8)$ to 1.8×10^{-10} cm^3 s^{-1}, and we therefore recommend $k_{11}(v=8) \leq 1.8 \times 10^{-10}$ cm^3 s^{-1} as an upper limit to derive physically allowed [O(^{3}P)] values."

Line 638-639.
"Furthermore, it cannot distinguish between OH(5) and OH(4) as a well as OH(9) and OH(8), and consequentially errors in OH(5) and OH(9) might be compensated by errors in OH(4) and OH(8) or vice versa".
This is a troubling statement as your main results in Table 3 (R11a, R11b, R11c, R11d, and R11g) involve these levels and describe rate coefficients for specific state-to-state reactions. This statement needs to be clarified. It sounds as if you treat OH(9)+OH(8) as a combined, single level as well as OH(5)+OH(4). Is this true? If you cannot distinguish between certain vibrational levels, then how can you determine rate coefficients for specific vibrational levels?

Yes, the SABER OH airglow emissions are a sum of OH(9-7) VER and OH(8-6) VER as well as OH(5-3) VER and OH(4-2) VER. We therefore cannot distinguish between OH(9) and OH(8) as well as OH(5) and OH(4).

However, when analyzing OH($v=9$)+O(^{3}P)→OH($0 \leq v' \leq v-5$)+O(^{1}D), we can still draw some conclusion because this reaction can deactivate OH($v=9$) only to OH($v \leq 4$) but not OH($v=5$). Thus, even if we cannot distinguish between OH($v=5$) and OH($v=4$) we can estimate a branching ratio of OH($v=9$)+O(^{3}P) → OH($v=4$)+O(^{1}D) if the total rate of OH($v=9$)+O(^{3}P) → OH($v \leq 4$)+O(^{1}D) is known. Furthermore, the rates of the individual paths presented in Table 3 are only a suggestion and not the main result. The main result of Table 3 is that the total loss rates of R11 are indicated to be slower compared to the values suggested by Sharma et al. (2015). These total rates presented here are a simplified solution in accordance with the rare laboratory experiments available and the OH transition considered. But as stated in the conclusion (l. 632-634): "Including additional OH transitions, …, might result in other values and deactivation schemes.

We rearranged this section as follows (l. 638-646):

"Furthermore, our OH airglow model is based on the transitions OH(9-7)+OH(8-6), OH(6-2), OH(5-3)+OH(4-2), and OH(3-1) only. Therefore, our model does not provide any information of OH(v≤2). It further cannot distinguish between OH(v=5) and OH(v=4) as well as OH(v=9) and OH(v=8), respectively, and errors in OH(v=5) and OH(v=9) might be compensated by errors in OH(v=4) and OH(v=8) or vice versa. Consequently, the rates of the individual deactivation channels presented in Table 2 and Table 3 should be viewed as a suggestion only. But these issues will only be solved eventually when future laboratory experiments provide the corresponding OH(v)+O$_2$ and OH(v)+O(^{3}P) relaxation rates and deactivation channels. Nevertheless, we have to emphasize that the shortcomings of our model do not affect the two main conclusions drawn in this study."

Tables 2 3. It is not clear if the results in Table 2 and 3 describe the Best-Fit model discussed in the conclusion. Table 2 shows empirically determined branching ratios of the OH(v) + O2 reaction for only VER observations "below 85 km" while Table 3 shows the branching ratios of the OH(v) + O(3P) reaction for only VER observations "above 85 km". The lack of consistency adds confusion to the findings of this study. Please clarify this. Is there not a best-fit model for altitudes 80-100 km?

The individual model steps were always fit to the entire altitude interval 80-100 km. But OH(v)+O2 quenching is more important below 85 km while OH(v)+O(3P) becomes dominant above 85 km. Therefore, these altitudes were added in the caption of the tables.

But since this caused confusion we deleted "below 85 km" and "above 85 km" in the caption of Table 2 and 3.

Table 3.
The two most important processes (largest rate coefficients) estimated from the best fits are not energetically allowed! Processes R11a and R11c are highly endothermic processes by ~3000 cm-1 and 2000 cm-1, respectively.

The processes OH(v=9)+O(3P)→OH(v=4)+O(1D) and OH(v=8)+O(3P)→OH(v=3)+O(1D) are not findings of this study. They were adapted from OH(v)+O(^{3}P)→OH(0≤v'≤v-5)+O(^{1}D) suggested by Sharma et al. (2015) and were also included in other OH airglow models (e.g. Panka et al., 2017, 2018). Therefore, for details about their credibility, we refer to Sharma et al. (2015, and references within).

Additionally, the state-to-state rate coefficients in Table 3 for the OH+O(3P) reaction appear to be in contradiction with the findings of Kalogerakis et al. (2016), who measured a large rate coefficient attributed to the resonant reaction OH(9)+O(3P)→OH(3)+O(1D). These results are non-physical and must be revised.

Kalogerakis et al. (2016) only reported that OH(v=9)+O(3P)→OH(v=3)+O(1D) is an important deactivation channel of OH(v=9)+O(3P)→products. They did not provide any rate or branching ratio of the channel. To our understanding, "important" means "not negligible" but it does not mean "dominating". Thus, we assumed that OH(v=9)+O(3P)→OH(v=3)+O(1D) has to occur but this channel is not necessarily the fastest deactivation path of OH(v=9)+O(3P).

This was stated in the text (l. 490-493):

"However, not much is known about the individual branching ratios of R11 except that $OH(v=9)+O(^3P)\rightarrow OH(v=3)+O(^1D)$ is an important deactivation channel but not necessarily the dominating one (Kalogerakis et al., 2016)."

As stated above, it seems most likely that fitting the highly uncertain OH(6-2) signal that has large systematic errors have skewed the results of this paper. Removing this constraint may bring the revised OH model into better agreement with recent laboratory and modeling studies as well as retrieve reasonable [O] and [H].

As we wrote above, we took into account that OH(6-2) VER is a relatively noisy signal by considering the relatively large error bars. Also, OH(6-2) does not considerably affect derived [O(3P)] and [H] values (about 10 %).

Furthermore, we have to emphasize that, at least to our knowledge, the rates and deactivation schemes applied in the OH model are not in conflict with ANY laboratory measurements but partly disagree with other model studies. However, the intention of this study was not to match other model studies but to review recent rates and deactivation schemes, and consequently provide [O(3P)] and [H] values based on justified assumptions.

Figures 1-5.
Why are there no error bars displayed for the measured OH(5-3)+OH(4-2) VER emissions despite error bars displayed for the OH(6-2) and OH(3-1) emissions?
We added error bars of OH(5-3)+OH(4-2) VER in Fig.1-5 and a short description
as follows (l. 176-181):

"The total uncertainty of SABER OH airglow data used here comprises three different error sources. Since we used climatology of the measurements (see Sect. 2.2), there are sufficient samples that the random noise component of the total uncertainty is essentially zero. The remaining two major terms are the absolute calibration error (<5 %) and the "unfilter" factor error (<3 %). Assuming a root-sum-square propagation of the individual uncertainties, this results in a total uncertainty of about 6 % for all data points presented in this study."

In general, the concept of fitting the zonal mean profiles for three OH bands is questionable. Operating with zonal mean profiles only, the authors are essentially fitting a single scenario (four individual signal scans). They must demonstrate how the rates derived from zonal mean profiles fit real single scans in measured emission bands. This will show whether the derived rates have any value for practical analysis of measurements of both instruments.
We used climatology instead of single scans because the individual scans are too noisy to derive any reliable rate coefficients. We are aware that deriving [O(3P)] and [H] based on zonal mean climatology instead of a scan-to-scan basis does introduce additional uncertainties. In particular, this approach fails when linearity between [O(3P)] and OH(9-7)+OH(8-6) VER breaks down.
But this issue was considered in Sect. 3.4 (l. 593-618; Fig. 7):
"The last problem lies in the fact that the approach used here (see Sect. 2.2) has to be applied to individual OH airglow profiles to derive [O(^{3}P)] and [H] correctly. However, the individual scans of OH(6-2) were too noisy to analyze single profiles and we therefore used climatology for all input parameters. By investigating individual OH airglow profiles, we would derive individual [O(^{3}P)] profiles and eventually average them to the mean [O(^{3}P)] profile. While in our case, we directly derive the mean [O(^{3}P)] profile. This makes no difference as long as the relation between OH airglow and

[O(^{3}P)] is a linear one. But Eq. (4b) shows that the relation between [O(^{3}P)] and OH(9-7)+OH(8-6) VER is only approximately linear because G also depends on [O(^{3}P)], as represented by the terms C_ν and $C_{\nu\nu}$. The linearity between OH(9-7)+OH(8-6) VER and [O(^{3}P)] of an air parcel with a certain temperature and pressure is solely controlled by [O(^{3}P)]×G. Note that [H] too is affected by this non-linearity issue since [H] depends on G (Eq. (4a)). Thus, derived [H] values are only reliable as long as the derived [O(^{3}P)], and as a consequence G, is not seriously in error.

In order to test the linearity, [O(^{3}P)]×G was plotted as a function of [O(^{3}P)] and the corresponding results for Best fit model at five different heights are presented in Fig. 7. It is seen that the relation between [O(^{3}P)] and [O(^{3}P)]×G or OH(9-7)+OH(8-6) VER, respectively, is linear for small values of [O(^{3}P)], while a non-linear behaviour becomes more pronounced for larger values of [O(^{3}P)]. Furthermore, the starting point of the behaviour is shifted to lower [O(^{3}P)] values at higher altitudes. In order to estimate this threshold, we performed a visual analysis and determined an upper limit of [O(^{3}P)] before non-linearity of [O(^{3}P)]×G takes over. The approximated upper limits are added as dashed lines in Fig. 7. Finally, an [O(^{3}P)] value at a certain altitude is assumed to be true if this value is below the corresponding upper limit of [O(^{3}P)]. Otherwise, it should be viewed more critically. This was done for each altitude and we found that the [O(^{3}P)] and [H] profiles presented in Fig. 6 are plausible in the altitude region <95 km. In combination with the estimation of chemical equilibrium of O$_3$ and the maximum of physically allowed [O(^{3}P)], we think that the [O(^{3}P)] and [H] derived by the Best fit model are reasonable results between 80 km and 95 km. Note that these altitude limits do not affect the results with respect to OH(ν)+O$_2$ and OH(ν)+O(^{3}P) presented in the Sect. 3.2 and 3.3."

Technical corrections:

Line 29-31.
This sentence needs a citation at the end.
Done, we added Andrews et al. (1987) and Mlynczak and Solomon (1993) as follows:
"Atomic oxygen in its ground state (O(3P)) and atomic hydrogen (H) strongly influence the energy budget in the mesopause region (~75-100 km) during day and night (Mlynczak and Solomon, 1993), and consequently affect atmospheric air temperature, wind, and wave propagation (Andrews et al., 1987)."

Lines 631-646.
These sentences should be moved to section 2.3: The OH airglow Base model.
We rearranged this section as follows (l. 629-642):
"We have to stress that we performed an empirical model study and the total rates and deactivation channels suggested here heavily depend on the OH transitions considered. Including additional OH transitions, like OH(9-4), OH(8-3), and OH(5-1) from the Optical Spectrograph and InfraRed Imager System (OSIRIS) on board the Odin satellite, might result in other values and deactivation schemes. This could be a subject of a future study. Also note that the Einstein coefficients used here might be in error (see Sect.3.1; Fig. 2). This does not affect the two general conclusions drawn above but would impact the empirically derived rates.
Furthermore, our OH airglow model is based on the transitions OH(9-7)+OH(8-6), OH(6-2), OH(5-3)+OH(4-2), and OH(3-1) only. Therefore, our model does not provide any information of OH($\nu \leq 2$). It further cannot distinguish between OH(ν=5) and OH(ν=4) as well as OH(ν=9) and OH(ν=8), respectively, and errors in OH(ν=5) and OH(ν=9) might be compensated by errors in OH(ν=4) and OH(ν=8) or vice versa. Consequently, the rates of the individual deactivation channels presented in

Table 2 and Table 3 should be viewed as a suggestion only. But these issues will only be solved eventually when future laboratory experiments provide the corresponding $OH(v)+O_2$ and $OH(v)+O(^3P)$ relaxation rates and deactivation channels. Nevertheless, we have to emphasize that the shortcomings of our model do not affect the two main conclusions drawn in this study."

and added a shorter paragraph in Section 2.3 (l. 269-273):
"As described in the previous section, the OH airglow model is adjusted to fit OH(9-7)+OH(8-6) VER, OH(6-2) VER, OH(5-3)+OH(4-2) VER, and OH(3-1) VER. Thus, the model cannot provide information about $OH(v\leq2)$. It further treats $OH(v=9)$ and $OH(v=8)$ as well as $OH(v=5)$ and $OH(v=4)$ as a single level and the corresponding deactivation channels should be viewed more critically. "