Authors' Reply

The authors would like to thank the referees for their constructive feedback, that helped in improving the manuscript. In the following, all revision points are addressed and the resulting text edits are included in the following way:
The comments are repeated and the responses are given below. Changes made in the manuscript are indicated in blue. Figure numbers with "R" correspond to figures in this reply not included in the manuscript.

Reply to Anonymous Referee #1

Comment: A slightly clearer distinction between total liquid water path, cloud liquid water path and rain liquid water path could be made throughout the paper. I would also like to see some more details of the distinction between cloud and rain liquid water path. This is hinted at on P7 line 17, which implies that it is taken from the ICON microphysics scheme, but it would be helpful to specify the difference in terms of the different size distributions etc.
Response: To clarify the discussion we added the following to the the second paragraph of Sec. 1: "Thus, we define LWP as the sum of CLWP and RWP." We checked throughout the manuscript the consistency and corrected especially Sec. 1 accordingly Furthermore, it is made clear that the size distributions from ICON microphysics. They are specified in Sec. 2.3, paragraph 2, now. "Cloud and rain particles are simulated with a 20 µm diameter mono-disperse and exponential distribution of water spheres, respectively. The exponential distribution has its intersect N_0 classically fixed to $0.08\,\mathrm{cm}^{-4}$ (Marshall and Palmer, 1948)."

Comment: Why are all the available microwave channels not used in the retrieval? I would expect that particularly the 183 GHz channels would contain additional information on the IWV, including its vertical distribution, and the quasi-window channels on the far wings of the 118 and 50-60GHz O2 bands will also respond strongly to liquid water. Since the data are already screened for cloud ice then scattering at 183GHz should not be a concern here.
Response: We thank the reviewer for proposing this retrieval extension. It certainly is an interesting objective for a follow-up study, but we chose the frequencies for two main reasons. First, we wanted to use only frequency bands that are currently also used by spaceborne microwave sensors to obtain a comparable product. Second, including additional channels to the LWP retrieval inhibits the calibration crosscheck with the IWV retrieval. This is important to avoid additional bias errors, as the calibration and the absorption characteristics in the O2 band are still uncertain (Maschwitz et al., 2013). Furthermore, the 183 GHz channels unfortunately suffered from hardware instabilities during both campaigns as pointed out by Konow et al. (2018a) and their data is not available all the time.

Comment: I would like to see some further discussion on the impact of surface wind speed (and the minor impact of surface temperature) on the retrieval of LWP and IWV. How does the frequency-variation of the brightness temperatures differ for surface wind speeds compared to that for IWV/LWP/RWP shown in figure 2? Is there any independent information content on the wind speed contained in the radiometer measurements, or does it effectively just add noise to the LWP/IWV retrievals?
Response: The error of the LWP retrieval shows no dependence on the 10 m wind speed as it can be seen in Fig. R1. A slight mean underestimation of the retrieval between 5 and $10\,\mathrm{g\,m}^{-2}$ for calm wind decreases toward wind speed of about $6\,\mathrm{m\,s}^{-1}$. There seems to be a retrieval overestimation for stronger wind speed above $12\,\mathrm{m\,s}^{-1}$, but this effect is of minor significance, as wind speed above $12\,\mathrm{m\,s}^{-1}$ occurs only rarely (as determined by ICON).
The IWV retrieval error as a function of wind speed is shown in Fig. R2. A slight systematic dependence of the error on wind speed is notable. However, the largest errors are again related to rare wind speeds.
To address this issue included a variation of the surface wind speed in Fig. 2. It shows a brightness temperature change of up to 2 to 3 K per $5\,\mathrm{m\,s}^{-1}$ wind speed change. The potential to retrieve surface wind speed from this rather weak signal, however, is low. The surface temperature influence of less than 1 K difference for sea surface temperature (sst) variation of 4 K is too weak for any reasonable sst retrieval with the HAMP channels. We added to the 3rd paragraph of Sec. in the manuscript: "The near surface wind speed slightly alters the BTs through modification of surface reflectivity and emissivity as also shown in Fig. 2. This influence will act as a random source of error to the LWP and IWV retrievals as no independent information to correct for

wind influence is available.”

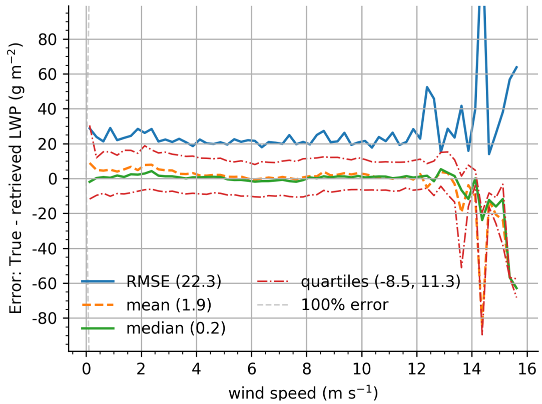

Figure R1. Neural network LWP retrieval error as function of 10 m wind speed.

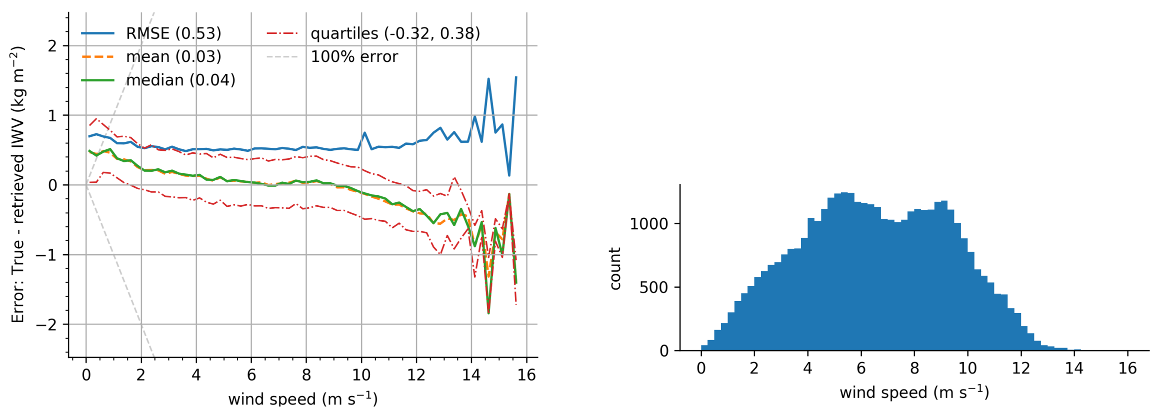

Figure R2. Left: Neural network IWV retrieval error as function of 10 m wind speed. Right: 10 m wind speed frequency distribution of the test database.

Comment: I find it slightly surprising that there is more liquid water during the dry season than the wet. I would like to see some more discussion about how the results in sec. 6 may be influenced by the choice of flight paths during the two campaigns. If specific conditions were either targeted or avoided then this could significantly bias the results.

Response: Indeed we find the higher LWP during the dry season to be one of our most interesting results. While you are right that the flight patterns could be responsible for part of the difference, also changes in the organization of clouds could cause the differences in cloud fraction and LWP. The fact that the medium LWP range from 100 to 400 gm-2 is less frequent in the wet season could be due to the higher degree of organization causing more clear sky areas and more intense clusters with higher amounts of precipitation. In that sense the latter would be missed by our flight patters as we avoided strongly convective scenes with formation of large ice particles. The aspect of organization is currently investigated by different LES modeling groups and will be a major objective of the EUREC^{4}A campaign.

The discussion was added to the paper as follows: “The higher LWP in the dry season might partly be explained by the choice of flight patterns. However, an analysis of ground-based LWP measurements at the Barbados Cloud Observatory (Stevens et al., 2016) over the years 2013-2018 confirms the generally higher LWP values during December than August (not shown). Thus, also changes in the organization of clouds could cause the differences in cloud fraction and LWP. The fact that the medium LWP range from $100\,\mathrm{g\,m^{-2}}$ to $400\,\mathrm{g\,m^{-2}}$ is less frequent in the wet season could be due to the higher degree of organization causing more clear sky areas and more intense clusters with higher amounts of precipitation. In that sense the latter would be missed by our flight patters as we avoided strongly convective scenes with formation of large ice particles."

The abstract was extended by "We hypothesize that higher degree of cloud organization on larger scales in the wet season reduces the overall cloud cover and observed LWP."

Within the conclusions we added: "An extension of the NARVAL observations is planned by the EUREC^{4}A field study in early 2020 ('elucidate the couplings between clouds, convection and circulation'; Bony et al. (2017)) which among other objectives will investigate convective aggregation. The algorithms presented here will be applied and together with additional measurements a better understanding of the governing processes that cause differences between dry and wet season will be analyzed. For that the campaign will provide additional observations of large scale dynamics, horizontally resolved remote sensing observations by a second and in situ observations by additional aircraft in the cumulus layer. Also, more locally targeted flights, distributed over the daytime are planed to study the diurnal cycle. Together with ship, drone and buoy measurements a unique dataset for a better understanding of precipitation onset will be generated."

Comment: The authors note that the WALES IWV measurements are only available in clear sky conditions so they only provide validation for the MWR retrievals when there are no confounding effects from liquid water. Is it possible to split the dropsondes into clear and cloudy scenes to demonstrate if there is any impact of liquid water on the quality of the IWV retrievals?

Response: The separation between clear sky and cloudy sondes is the reason, why we separate Tab. 2 in "observed pairwise" (all sky) and "observed by all" (WALES available, clear sky). We added the following paragraph to make this more explicit. "A small confounding effect from liquid water in cloudy scenes can be derived from the separation of the HAMP – dropsonde comparison into all ('observed pairwise') and clear sky ('observed by all', i.e., when also WALES is available) in Tab. 2. In the NARVAL1 dataset, the bias for cloudy sky sondes ($0.24\,\mathrm{kg\,m^{-2}}$) is somewhat smaller than that for clear sky ($0.32\,\mathrm{kg\,m^{-2}}$). However, RMSD and STD in cloudy scenes are about $0.3\,\mathrm{kg\,m^{-2}}$ larger than in clear sky. NARVAL2 also shows a larger bias in cloudy sky of about $0.53\,\mathrm{kg\,m^{-2}}$ in comparison to clear sky ($0.28\,\mathrm{kg\,m^{-2}}$). The cloudy sky RMSD and STD of $1.32\,\mathrm{kg\,m^{-2}}$ and $1.21\,\mathrm{kg\,m^{-2}}$, respectively, are only slightly larger than their clear sky counterparts. An increase of the random error for cloudy scenes is expected as also higher water vapor variations are expected in heterogeneous cloud fields."

Comment: P6 line 18 Why is there a need to convert from water vapour number density to volume mixing ratio? It is the former that is required to calculate integrated water vapour mass.

Response: Indeed, you are right. We note this conversion from water vapor density to water vapor mixing ratio because we want to show mixing ratio in Fig. 4.

Comment: P8 figure 2 I suggest using a logarithmic colour scale to show the relative frequency to highlight any detail away from the strong "clear sky" line

Response: We followed the suggestion and changed the color scale of Fig. 3, that was on page 8 of the previous manuscript version and shows BT frequency distribution.

Comment: P9 line 31 Biases with respect to what?

Response: The biases are related to the deviation of LWP from 0 during clear sky scenes. We adjusted the sentence accordingly. "When retrieval algorithms are applied to HALO measurements, slight biases of LWP from 0 with slow changes over time are observed during clear sky scenes."

Comment: It would be useful to have an indication of along-track distance on figure 4 rather than just time

Response: We followed the suggestion and updated Fig. 4 as well as Fig. 10 accordingly.

Comment: P13 final paragraph – it would be nice to refer to fig 6 early in this discussion.
Response: We added an earlier reference.

Comment: P14 discusses the impact of negative LWP retrieval values on the bias. These could be avoided by performing the retrieval in logarithmic space (I.e. retrieving log(LWP)). Would this have a significant impact on the results?
Response: Retrieving log(LWP) would avoid negative values, but it would inevitably result in a overestimation of the retrieved mean LWP as noise is always positive. Another reason for not retrieving the logarithm is, that the natural first order approximation relation between brightness temperatures and LWP is linear.

Comment: In figure 10 it might be clearer to plot the LWP and RWP on a logarithmic scale – in the current plot it is hard to see the cloud LWP retrieved by the MWR between 17:38:30 and 17:39:10 that is discussed in the text at the end of page 17
Response: We changed Fig. 10 by using a piecewise linear scale now, that magnifies the range from -20 to 20 $\mathrm{g\,m^{-2}}$. We decided against a logarithmic scale as this can not display negative retrieval artifacts.

Comment: The paper is clearly written and generally easy to follow, although I find a number of sentences do not read well and should be redrafted. There are also a few typos: [...]
Response: All remaining minor corrections by the referee were agreed on and are incorporated in the revised manuscript.

Comment: P1 line 1 "...identified especially marine low level clouds to play a critical role for the climate."
Response: We removed the first sentence on recommendation from referee #2.

Comment: P1 line 5 ". . .to better understand *the* LWP of warm clouds..."
Response: "the" is added.

Comment: P2 line 5 "Especially, shallow marine clouds are attributed to contribute largely to intermodel spread of climate models"
Response: We redrafted this sentence to: "Sherwood et al. (2014) attribute especially shallow marine clouds to contribute largely to intermodel spread of climate models."

Comment: P2 line 28 "Visible/near infrared techniques *such* as those applied to MODIS..."
Response: "such" is added.

Comment: P3 line 9 "...allow to study clouds with similar, however, more sensitive and higher spatially resolving instrumentation than available on satellites." Perhaps "...allow the study of clouds with similar, but more sensitive and higher spatially resolving, instruments to these available on satellites."
Response: We thankfully incorporated this comment.

Comment: P3 line 13 "Their study shows *the* sub-footprint variability of spaceborne Special Sensor Microwave Imager/Sounder..."
Response: We added "the".

Comment: P3 line 24 "The assessment of LWP (Sec. 4) reveals the importance *of using* ancillary measurements, e.g. lidar measurements for low LWP values and cloud radar measurements for lightly precipitating cases."
Response: We thankfully incorporated this comment.

Comment: P3 line 26 ". . . between dry and wet *seasons*"
Response: Changed "season" to "seasons".

Comment: P4 figure 1 I think the caption mis-labels the thin and thick lines based on the dates in the legend (I.e. NARVAL 1 looks like it should be the thick lines)
Response: Corrected.

Comment: P5 line 12 "more dominant in the higher *frequency* window channels"
Response: Corrected “window” to “frequency”.

Comment: P5 line 16 remove the comma after "both"
Response: Removed comma.

Comment: P6 line 9 "cloudy conditions as *well* as possible"
Response: Corrected “good” to “well”.

Comment: P6 line 25 "...prevent having data during some flights..."
Response: We redrafted that sentence to: “While all research flights during NARVAL1 can be used, no data is available for some NARVAL2 flight days due to hardware issues as summarized in Tab. 1.”

Comment: P7 line 10 is ambiguous. Are all profiles with ice excluded, or only ones with ice water path above 1000 g/m^2 ?
Response: We reordered that sentence for better understanding. “In general, the training and test data excludes cases with LWP greater than 1000 g m^{-2} and cases with ice, i.e., 86 % of all profiles over the ocean are used.”

Comment: P7 line 11 "...the ocean *are* used"
Response: Corrected “is” to “are”.

Comment: P8 line 3 "...are *visible* as a line..."
Response: Corrected “visibly” to “visible”.

Comment: P9 line 11 "...never deviates *more* than..."
Response: Added the missing “more”.

Comment: P14 line 7 "...decided *to use* a retrieval"
Response: Replaced “we decided for” by “we chose”. Further we slightly changed the beginning of the following sentence to: “Instead to include clear sky directly in the retrieval, we make use of lidar measurements, which are better suited than MWR for cloud masking.”

Comment: P21 line 11 "...consistent *with*"
Response: Replaced “to” by “with”.