Reviewer 2.

This paper analyses the variability of boundary layer moisture at the ARM ENA site using ground-based observations and ERA5 reanalysis data. The authors compute mixed-layer water budgets at monthly timescales and analyze the respective contributions of the different terms, and then assess mesoscale moisture variability on sub-daily timescales and their relationships to updrafts, liquid water path and precipitation. The analyses seem sound, and the paper is well written and quite easy to follow. My major concerns are that (i) many assumptions are not justified or discussed, (ii) I did not understand if the authors just focus on stratocumulus or also shallow cumulus conditions, (iii) the story and sequence of analyses is hard to anticipate, and the novelty and connection among different sections is not really clear, and (iv) uncertainties are not systematically quantified.

These comments, along with some more minor comments, are addressed in more detail below.

We thank the reviewer for the constructive comments that have led to changes and substantial improvements to the manuscript. Please find below our point-by-point responses to the comments. The comments are in black, and our responses to it are in green. Any changes to the article text are mentioned in blue.

Major comments:

1.**Justification & discussion of assumptions:** For many of the assumptions made here, I miss a justification or discussion. For example, are the boundary layers analyzed really well-mixed up to the boundary layer top? For stratocumulus conditions I'd assume this is the case, but for decoupled shallow cumulus conditions, only the sub-cloud layer is well mixed. (See also comment 2 regarding the cloud regimes below). Could the authors demonstrate the well-mixedness of their boundary layers, and discuss how different cloud regimes might affect the budgets?

We thank both reviewers for pointing this out. We now discuss this topic and infer the uncertainty due to the inclusion of cases that are not well mixed. We repeat here the discussion provided for the 1st reviewer.

The assumption of a well-mixed boundary layer is important to define the framework over which to study processes that affect clouds and moisture. The framework has been used for multiple decades now to understand the primary controls on the boundary layer vapor and (cloud) liquid field (e.g. Betts 1976 JAS for cumulus; and Brost et al., 1984 for stratocumulus). In our case, because we are using all data, there will be a mix of conditions, similar to the profiles shown by Lock et al. (2000 MWR). In theory the budget terms could be modified to include multiple mixed layers, or a mixed layer on top of a stable layer, or vice-versa. However, it is not possible to implement such a framework to the large data (~7 years) used in this study. It should be also noted that decoupling indices are usually calculated from radiosonde profiles, and at the ENA site we only have two radiosondes (00 and 12 UTC) per day, making it difficult to calculate budgets of a thermodynamically decoupled boundary layer. This is primarily because any stable or mixed layers above the surface layers are transient and do not persist for more than 6 hours. However, the boundary layer inversion is omnipresent as it is primarily controlled by large scale subsidence. Lastly, in a decoupled boundary layer the surface moistening and the entrainment drying are mediated by boundary fluxes through another layer. Hence assuming a decoupled boundary layer is well-mixed will essentially cause an imbalance in the equation 1, thereby yielding a residual term. However, the seasonal and annual average residual term (Table 2) are far smaller than any of the calculated terms. This suggests that although the boundary layer may not be well-mixed, it can be assumed to be well-mixed at seasonal to annual timescales.

To address the reviewers concern in more detail we have therefore examined all radiosondes and radiosondes during marine conditions and calculated a Decoupling Index (DI) defined as:

$DI= (Z_{CB}-Z_{LCL})/Z_{CB}$

Profiles with DI < 0.25 are classified as strongly coupled and profiles with 0.25 < DI < 0.4 as weakly coupled. With this definition more than half of marine profiles are strongly coupled and about 68% of the marine cases have DI < 0.4. Marine cases with low cloud base height (< 1.2 km) are statistically more coupled with 64% having DI < 0.25 and 80% having DI < 0.4.

	All	Marine	Marine with CB<1.2 km
Number of cloudy cases	1825	681	469
DI<0.25	739 (40%)	351 (52%)	301 (64%)
DI<0.4	1205 (66%)	461 (68%)	374 (80%)

Selecting marine cases increases the percentage of coupled and weakly coupled cases compared to the entire dataset. The annual cycle of DI for the marine cases, shown in Fig. R1, shows higher variability in the summer months.

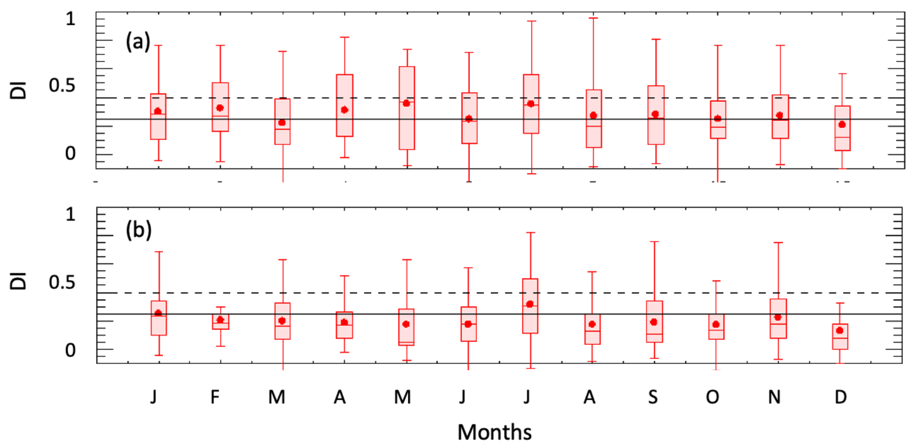

Fig. R1: Annual cycle of the decoupling index (DI) for (a) all marine cases and (b) marine cases with cloud base < 1.2 km. The solid line indicates the 0.25 threshold (strongly coupled) and the dashed line indicates the 0.4 threshold.

To evaluate whether the inclusion of decoupled cases affects the moisture budget, we repeated the calculations of eq. 1 including only cases with cloud base height < 1.2 km. This is shown in Fig. R2. When compared with the new Fig. 7b, the main features of the results are not affected. However, the inclusion of decoupled profiles in the analysis (Fig. 7b) enhances the contribution of the entrainment fluxes to the total moisture sink from ~18% to ~25%.

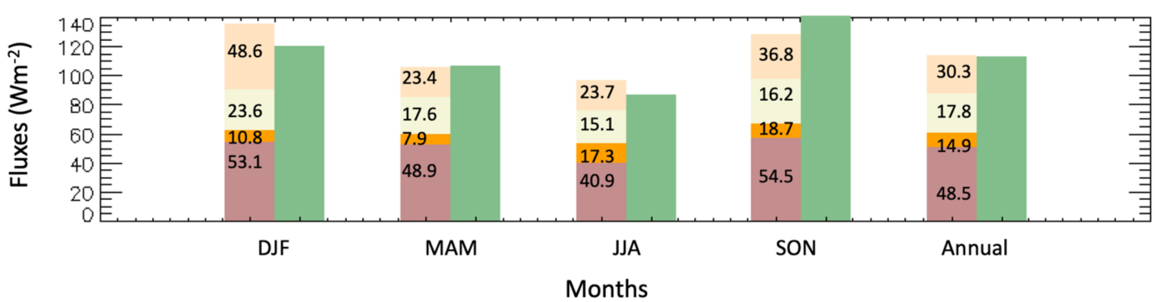

Fig. R2: Seasonal fluxes calculated only for cases with cloud base height < 1.2 km where coupled conditions prevail. Colors are the same as in Figure 7b (Advection: Dark brown; Local tendency: Orange; Beige: Entrainment; Precipitation: Pink; Evaporation: Green)

	DJF	MAM	JJA	SON	YEAR
$\frac{\mathcal{L}}{g}\hat{p}\frac{\partial\langle \boldsymbol{q_t}\rangle}{\partial t}$	-10.8± 68.5 (7.9%)	-7.9±62.1 (8.1%)	-17.3±59.6 (17.8%)	-18.7±68.8 (14.8%)	-14.9±64.6 (13.5%)
$\frac{\mathcal{L}}{g}\hat{p}\langle \boldsymbol{v} \cdot \boldsymbol{\nabla} q_t\rangle$	-53.1±53.8 (39.0%)	-48.9±49.7 (50.0%)	-40.9±43.7 (42.2%)	-54.5±50.4 (43.2%)	-48.5±48.9 (43.8%)
$\frac{1}{g}\omega_e \Delta q_t$	-23.6±12.6 (17.4%)	-17.6±8.9 (17.9%)	-15.1±9.0 (15.6%)	-16.2±9.1 (12.9%)	-17.8±9.8 (15.4%)
$\mathcal{L}$ P	-48.6±97.5 (35.7%)	-23.4±54.4 (23.9%)	-23.7±45.9 (24.4%)	-36.8±60.5 (29.2%)	-30.3±54.4 (27.3%)
E	105.3±50.2	87.9±54.4	84.9±65.4	122.4±79.8	99.1±69.9
Residual	-30.8	-10.1	-12.0	-3.8	-11.7

Table R1: **Seasonal average values and standard deviation of the budget components for cases with cloud base < 1.2 km. In parenthesis are the contributions of each negative term to the total boundary layer drying. Residuals are computed as the difference between source and sinks (Wm^{-2}).**

This point is now mentioned in the introduction Lines 65-68:
“The validity of the mixed layer framework has recently been shown to be sufficient to explain synoptic and monthly variability in the sub-cloud layer (Albright et al., 2022) however our dataset includes a mix of coupled and decoupled cases, and it is therefore important to understand how often the assumption of a well-mixed boundary layer is verified at the site, and how it affects the results.”

Section 4 Lines 203-207
“To this end we examined 1825 soundings of which 681 where marine conditions and calculated a decoupling index (DI) defined as $(Z_{CB}-Z_{LCL})/Z_{CB}$. We then classified as strongly coupled cases with DI<0.25 and as weakly coupled cases with 0.25 < DI < 0.4. According to this classification most marine cases (68%) are weakly or strongly coupled. The decoupling index is generally smaller when the cloud base is lower and cases with cloud base < 1.2 km have DI<0.4 in 80% of the cases.”

And the uncertainty associated with this assumption is discussed in Section 4.3 at lines 327-331:
“At this point we are in the condition to evaluate the impact of including decoupled conditions in the analysis. We repeated the budget computations including only a subset with cloud base < 1.2 km which present mostly coupled conditions. The results showed a diminished contribution of the entrainment fluxes that decreased annually from 26% to 18%. It is therefore

likely that the inclusion of decoupled conditions in the analysis leads to an overestimation of the moisture sink due to entrainment fluxes."

In the discussion at lines 508-510:
"Although the majority of marine cases at the site can be classified as coupled or weakly coupled, inclusion of decoupled cases in the analysis introduces uncertainties leading to an overestimation of the contribution of entrainment fluxes to the budget."

I understand that ERA5 data is necessary to complement the observations, but did the authors check whether the moisture profiles are consistent with the radiosonde and Raman lidar profiles? E.g., are the moisture profiles consistent enough that ERA5 can be used to computed the gradient? Even if this can be checked only for a limited data sample, it would greatly increase confidence in the approach. Similarly, for the local tendency and the PBL heights, could the Raman lidar profiles be used to check the hourly variability and a potential diurnal cycle in the terms that would be missed with the twice daily radiosondes?

ECMWF moisture profiles were used for the calculations of the large-scale advection of qv and for the advection of the PBL height used in the mass budget equation. For everything else (i.e., PBL height for the integration in eq. 1, qv for the local tendency, and Δq_t) we used radiosondes. We agree that it would be great to have Raman lidar profiles. Unfortunately, we don't have those yet, but are developing the retrievals to derive RL profiles at 10 s resolution. We provide in Fig. R5 a comparison between ECMWF and radiosondes expressed as mean and standard deviation of differences and the mean difference expressed as a percentage of the average mixing ratio at the site. Statistically the ECMWF seem to underestimate the mixing ratio of about 10-15%. The differences are more pronounced near the top of the boundary layer where the humidity gradient is often located.

We added this discussion in the paper in section 2.1, Line 115-119:
"ECMWF profiles were compared to the local soundings between 2015 and 2020 and found to underestimate the mixing ratio of about 10% with standard deviation of 1.-1.6 gkg^{-1} between 0 and 3 km with a maximum underestimation of 15% at the PBL height. The ECMWF profiles are therefore suitable for the estimation of the vapor advection component of the budget but not suitable for the estimation of the PBL height. For this purpose, we use radiosondes."

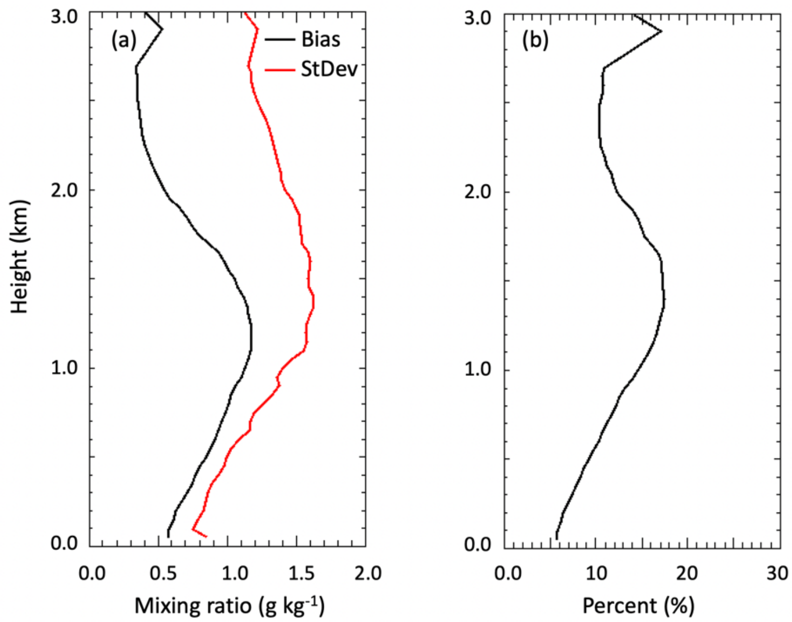

Fig. R5: (a) Mean difference between RS and ECMWF (black) and standard deviation around the mean difference (red) for N=4139 soundings between 2015 and 2020. (b) Mean difference between RS and ECMWF expressed as a percentage of the average mixing ratio.

2. Cloud regimes: As alluded to in Sec. 1 and 2, both stratocumulus and shallow cumulus conditions are frequent at ENA. But these cloud regimes are associated with coupled vs. uncoupled boundary layers, which is a relevant difference for this study. Except for Sec. 5, which clearly addresses stratocumulus conditions, only in L147 a cloud criterion is mention: "In the following discussion only boundary layer clouds with cloud fraction from the ceilometer greater than 0.99 were selected (total of 3580 hours)." So does this mean that the entire discussion of the budgets focuses on stratocumulus conditions with ~100% cloud cover? And does the 0.99 threshold refers to hourly values? Please clarify, and discuss more prominently.

We realize that this inclusion caused confusion and was mentioned by the other reviewer as well. The strict condition of overcast cloudiness was set only for the purpose of adiabatic computations for the reason that the adiabatic computations are very sensitive to the cloud boundaries. We understand that the digression made the paper hard to follow and have moved it to the appendix in the revised version. This way it is clear that the restrictive conditions only apply to the data in the Appendix.

3. Story, structure of the paper and novelty: I found it hard to anticipate the story and the structure of the paper from the abstract and the introduction. For example, Sec. 3.1 on the cloud adiabaticity seems rather peripheral and came quite unexpectedly. And at first, I expected just an analysis of the water vapor budget, and then I realize the paper focuses on the total water budget, and also constrains the mass budget. So it would have helped me if the story of the paper was clearer and if the reader's expectation was a bit more guided.

Following the reviewer's comment, we have moved the discussion of cloud adiabaticity and retrieval discussion to the Appendix. We now also clarify better at the beginning that we are considering the total water budgets at Line 80-81 (vapor and liquid, including rain) and we use all marine cases without cloud selection.

In a similar spirit, I missed the coherence and connections between the different Sections. How do the monthly mixed-layer budget analyses connect to the analyses of the mesoscale variability? Could you construct mixed-layer budgets also on shorter timescales to connect to the monthly budgets and understand how the terms contribute differently at different timescales?

Thank you for raising this issue. The purpose of the analysis was to understand the role of moisture at the site at different temporal and hence spatial scales. Clouds and liquid water path exhibit a distinct annual and diurnal cycle as reported by previous studies (e.g. Ghate et al., 2021 JAMC). However, boundary layer water vapor only exhibits a distinct annual cycle, not a diurnal cycle. In addition to the annual cycle, water vapor also exhibits a mesoscale variability. Both the annual and mesoscale variability influence clouds and precipitation, the annual variability being driven by large scale advection and the mesoscale variability being driven by local processes. Both variabilities are hence explored in this article.

We added these considerations in the introduction Lines 73-76: "Water vapor, unlike LWP, does not exhibit a diurnal cycle, therefore the annual and mesoscale variability are the primary modes through which the interaction between boundary layer vapor and cloud processes can be examined. Both modes influence clouds and precipitation, the annual variability being driven by large scale advection, as shown later, and the mesoscale variability being driven by local processes."

About the second question, calculating the budgets on diurnal timescales would have required PBL height at hourly timescales that we lack due to the sparsity of radiosonde launches. However, we would like to note that, if suitable data were available, a study on diurnal budgets in different seasons would illuminate the role played by different processes in determining LWP, because LWP, rain rate and entrainment exhibit a distinct diurnal cycle.

I would also recommend the authors to highlight more explicitly what is new in the paper. How does the novel retrieval used, which better separates cloud and drizzle contributions to TLWP, affect the robustness of your analyses and conclusions? What are the novel insights gained with the ENA data here compared to previous ML budgets? In L541 it is mentioned that 'the results presented herein are useful for future observational and modeling studies on low clouds conducted at the ARM ENA site', but can you be a bit more explicit?

We thank the reviewer for this observation that gave us the opportunity to better discuss this aspect. The derived moisture budgets utilized high-quality retrievals that have low uncertainty thereby giving us confidence that the results represent physical processes rather than instrumentation issues. As shown in Table 1 the analysis utilized many instruments each requiring careful calibration, denoising, cleaning, and (some) retrieval development. The consistency of the results among the instruments and the physically meaningful scenarios emerging from the mixed layer budget analysis highlight the quality and low uncertainty of the data, and the feasibility of the procedure used to calculate the budgets. Among the insights allowed by the new retrievals are the ubiquitous presence of drizzle throughout the year even in seasons when the average LWP is low, and the relatively more pronounced drizzle annual cycle compared to the LWP cycle, pointing to the fact that LWP is only one of the factors influencing drizzle formation. The fact the drizzle LWP has an annual cycle that is distinct from the cloud LWP points also to the importance of looking at processes seasonally. We think these results will be very useful in guiding our next study when we will analyze the combined effects of aerosols and turbulence in the development of precipitation, and we hope they will be useful to others as well. To our knowledge this aspect hasn't been fully examined yet.

To our knowledge, the present analysis is also the first study to characterize and determine the controls of moisture variability at a *subtropical* marine location such as the ENA site. Previous studies performing moisture budgets either utilized reanalysis data (e.g Wood and Bretherton, 2006), or coarse satellite data (e.g., Kalmus et al., 2012), or were of limited duration (e.g Caldwell et al. 2005; Albright et al. 2022; Brost et al. 1982). Hence it is difficult to make direct comparison to these studies. However, our derived entrainment rates, and precipitation fluxes are in close agreement to those reported by Caldwell et al., 2005 and Kalmus et al., 2012.

Finally with regard to the last question, the current analysis will be useful for future modeling and observational studies in various ways. For example, future studies aimed at retrieving or simulating entrainment rates should expect higher values in the winter months compared to summer months. Studies aimed at simulating clouds at the ENA site using traditional Lagrangian LES framework should utilize thermodynamic profiles of parcels before it gets advected over the site as advection contributes ~50% to the total moisture budget at the site. Finally similar moisture budget analysis performed using output from an Earth System Model (ESM) will inform whether the ESM is accurately simulating water vapor in the region and its sources and sinks at different spatial and temporal scales.

We have added some of these considerations in the conclusions:

Lines 496-500 Unlike LWP, water vapor at the site doesn't present a diurnal cycle but presents an annual and mesoscale variability that is strongly connected with cloud and precipitation at the site. To our knowledge, the present analysis is also the first study to characterize and determine the controls of moisture variability at a subtropical marine location, such as the ENA site, using a long-term dataset of ground-based data.

Lines 539-544 "As a final consideration we spend a few words to highlight how the separation of cloud and drizzle water path in the new retrievals reveals the ubiquitous presence of drizzle throughout the year even in seasons when the average LWP is low. By looking only at the total LWP, only a weak annual variation appears, however the drizzle LWP shows a more pronounced seasonal variability, pointing to the fact that LWP is only one of the factors influencing drizzle formation. These results can be useful for future observational studies aimed at understanding the combined effects of aerosols and turbulence in the development of precipitation, pointing to the importance of looking at processes seasonally."

Lines 549-551: "Finally, the annual regional moisture budget resulting from an Earth System Model (ESM) will inform whether the ESM is accurately capturing the water vapor variability in the region and its sources and sinks at different spatial and temporal scales."

4. Uncertainty quantification: I missed a quantification of the uncertainty of the different terms in equations (1) and (3). Also, can you briefly say in L133 how the uncertainties of the retrievals are estimated?
We have added an explanation of the uncertainty of each term in section 4.1.

Lines 236-238: "Uncertainty in the estimation of PBL height from radiosondes is estimated to be 100-200 m (Sivaraman et al., 2013) this is a lower limit estimate in our case because the PBL height is kept constant for 12 hours."

Lines 247-249: "The uncertainty of this term is driven by the uncertainty of the microwave radiometer retrievals that is estimated 0.5 mm for water vapor and 15 gm^{-2} for liquid water path. This translates in an uncertainty of less than 0.25 g/kg for the average q_t."

Lines 259-261: "The uncertainty of the advection term is hard to estimate. From a comparison of ECMWF and radiosondes profiles at the ENA the mixing ratio uncertainty is expected to be around 15% with higher uncertainty near the top of the boundary layer, where the humidity gradient is located."

The SPARCL precipitable water vapor and liquid water path uncertainties are derived from the a posteriori covariance of the optimal estimation.
We added at line 140: "Uncertainties in the water vapor and liquid water path retrievals are estimated from the a posteriori covariance information obtained with the optimal estimation retrieval…"

Minor comments:

- Retrievals section 2.3: I think this section could be written a bit more concisely. From one paragraph to the other, it seems to jump from one algorithm (with unfamiliar acronym for me) to another. Also, I think it could be worthwhile to present the comparison of SPARCL and MWRRET (L141 onwards) in an appendix, not to depart from the main story of the paper too much.

We thank the reviewer for this suggestion and moved the comparison to the Appendix as suggested.

- Use of commas: I missed a lot of commas throughout the manuscript, e.g.:
    - L113: were derived, cases that ....
    - L134: For ten selected cases of weakly precipitating marine stratocumulus clouds, vertical profiles of water
    - L143: Traditionally, the total liquid water path retrieved by radiometers is assumed to represent the cloud water path. However, in the presence
    - ... Please check the entire manuscript carefully.

      We checked the manuscript- Thank you

- References: I'm not sure which program (if any) the authors use for the references, but it's not consistent and sometimes erroneous. E.g. Zhou and Bretherton 2019 is cited differently in L381 and L432. Also, the reference 'Shultz and Stevens, 2018 is not correct (L379), or (Zheng et al. Lamer et al. 2019; ...) in L44. Please improve throughout the manuscript.

We have reformatted the references at the end and revised and uniformed all citations in the text.

- L276: I understand that strong precipitation rates can introduce strong peaks, but this is the intermittent nature of rain, and I guess not a measurement error. Can you understand from your data how such strong precipitation rates are locally balanced? It would be very interesting to analyze this.

We agree with both reviewers on this. It is very difficult to estimate whether a very large spike in measured precipitation is due to splashing. For this reason, we have repeated the computations keeping all precipitation readings. The new Fig. 7 (now Fig. 5) shows the updated values, it can be seen that the contribution of precipitation to the moisture sink has increased especially in winter. As for the last question, I am not sure. I think looking at extreme precipitation may require the development of a more specific higher resolution dataset and probably an analysis of uncertainties of precipitation that uses all precipitation measurements at the site (perhaps adding the scanning precipitation radar?).

- L308: please specify what kind of filter is applied here.

Added "eliminating points beyond 2 standard deviations from the monthly mean" at Line 226.

L343: "The seasonality of the large-scale advection term is also the factor that determines the seasonality of the overall budget." --> Please clarify how you get to this statement. From Fig. 7 it seems that the seasonality in the LHF or precipitation terms is also very large.

The reviewer is correct. We corrected the statement: "The large-scale advection term is the largest moisture sink". At Line 332.

- L345ff: Does the magnitude of the monthly residuals depend on how much data was used per month? I.e. if only a few days of data could be used, it might not be surprising that the monthly budgets don't balance well. Please clarify.

It is likely that a smaller dataset would be more sensitive to the large noise. In our case with more than 5 years of data, after selecting marine cases and cleaning the dataset, there were between 700 and 1900 cases in each month. It is likely that a smaller dataset may require more attention to avoid biased results.

- L389: How do you interpolate 10-min profiles over 1 minute? Do you downsample the data?

It was just a linear interpolation between successive measurements. We are working on an improved high-resolution dataset of Raman lidar moisture profiles that will enable us to do a higher resolution analysis.

- Figures:
    - 11: Please specify what normalized height refers to in the right panel

      The height is normalized to the cloud base height. We now specify in the caption of Fig. 11 and 12.

    - 12: the figures are far too small and can hardly be read.

      We increased the font size of axis labels and legends.

- L511: I do not understand this sentence, please clarify: "The lower mid-tropospheric humidity during the winter months, together with turbulence (Ghate et al. 2021) point towards turbulence being the primary controlling factor rather than water vapor in determining cloudiness in the region."

We agree with the reviewer that the sentence was confused It was replaced with: "During winter months, turbulence, rather than water vapor, appears to be the primary controlling factor of cloudiness in the region (Ghate et al., 2021)." Line 500.

- L532: "Moist and dry patches present differences in vertical velocity with dry regions displaying more frequent downdrafts than moist regions immediately below the cloud base." ï from Fig. 12 I'd say the opposite, please clarify.

Figure 12 a) shows that the frequency of occurrence of downdrafts in dry regions (blue dashed line) is about 50% between the normalized height of 0.6 and 1 compared to a frequency of downdrafts in the moist region (red dashed line) of about 35% in the same height range.

We changed the text to: “Moist and dry patches present differences in vertical velocity with dry regions (dashed blue lines in Fig. 12 a) displaying more frequent downdrafts than moist regions (red dashed lines) immediately below the cloud base.”

Typographical suggestions:

L45: that that the water ...Done

L229 / L263: the total water mixing ratio is once written in normal and once in bold font – please harmonize. Done

L234: please use proper math formulation for the averaging brackets. Done

L279: maybe add after ‘... the entrainment rate **(see Sec. 4.2)**‘ Done

L303: ...balanced by local change in the boundary layer **height** (?) Done

L321: large-scale turbulence --> do you mean subsidence instead of turbulence? Yes, changed. And previously, largescale was written without ‘-‘, please make it consistent. We revised all instances

L322: that what reported --> than what was reported .. Done

L514: hear --> heat Done