Response to Reviewers: Variability of Bulk Water Vapor Content in the Marine Cloudy Boundary Layers from Microwave and Near-Infrared Imagery

We sincerely thank the reviewers for their thoughtful comments. Below are our responses in blue.

Response to Anonymous Referee #1

Reviewer comments

General: An interesting contribution on the synergetic use of different satellite measurements to monitor some PBL characteristics. I have a few major concerns, and various minor questions comments that I'd like to see addressed in an update. I rated the manuscript as Major Revision, mostly to leave the authors more time if any of my comments require larger work.

Major Issues:

- I am missing some more information on the systematically different sonde data vs. the RO data. Figure 2 and 3 seem to show different conditions, thus potentially sampling different areas, with RO data having CWV data up to 3cm, while the sonde data extends only up to about 1.5cm. Is the RO sharpness primarily identifying regions that are more humid while the sonde data is in less humid areas?

As the reviewer points out, the difference is mainly sampling different areas, that is, when the sharpness parameter is greater than 2.5 (top panel in figure 4) the comparison encompasses regions where a lot of cumulus clouds might be present (see Figure 3 for reference), hence, the BL-CWV is higher. However, when the comparison is made for sharpness parameters greater than 3 (bottom panel in figure 4) the comparison is mostly over stratus regions and the BL-CWV values resemble more the sonde values. The sondes on the other hand are always restricted to stratus regions since we only used robust inversions, that is, when the three different methods (explained in the text) agree within 200m.

The following sentences will be modified,
(P4 line 7 of the original manuscript): That is, those inversions where the boundary layer inversion height estimates of the tree methods agree within 200m, **which mostly occur in stratus regions**.

(P5 line 8 of the original manuscript): This improvement arises because when using a larger sharpness parameter, we are ensuring that most pairings are in the stratus regions **(see Figure 3 for reference)** where the AMSR-MODIS technique should work better. **A larger sharpness parameter also reduces the range of the BL-CWV comparison by excluding the high values found under cumulus regimes. This makes the comparison ranges (that is, in Figure 4-bottom) similar to the ones found in the sonde comparison where the sondes used are restricted to stratus regions by using the robust inversion criteria. That is, when the three different methods to find the inversion (explained in section 3.1) agree within 200m.**

Is there any chance to also compare RO to sonde to improve the understanding? After careful consideration, we feel that such comparison belongs to a GPS-RO validation paper and not here. However, we hope that our previous explanation will suffice for the reviewer.

Or to include more data?
We included as much radiosonde and GPS-RO data as we could find.

- Page 2, Line 31 (P2/L31): points to systematic issues in the December months. Just "removing" a month because it does not seem to fit, and use a another data version that fits, would need a much more substantiated justification. Thus I'd like to see more info on what might cause this issue and why the version 6.1 was used now. And then, why 6.1 is not used throughout.

The issue was a one-off coding error on the MODIS processing algorithm (per the personal communication with Richard Frey). We discover the issue after the whole dataset was produced with version 6.0 and a reprocess of the AMSR-MODIS using version 6.1 through-out the entire time period is currently outside our possibilities due to the large time involved in downloading the terabytes of MODIS data. We will modify the following sentence: Instead, version 6.1 was used for all December months **as recommended by the MODIS team. A full reprocessing of the AMSR-MODIS dataset using MODIS version 6.1 (or the latest MODIS version) is left for a future AMSR-MODIS version.**

Minor / Editorial Issues:

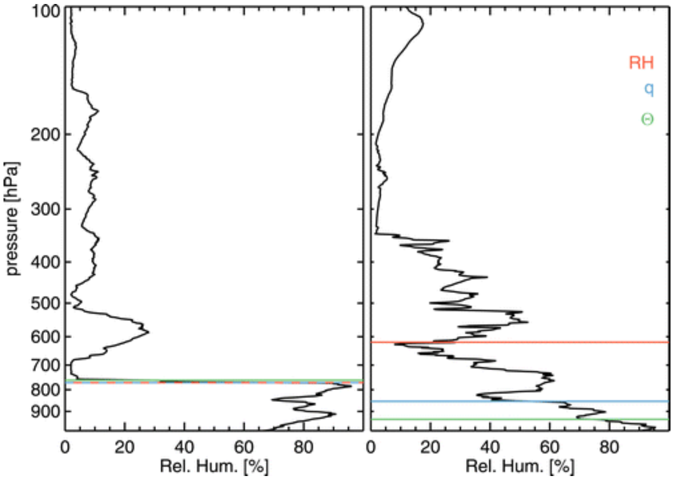
- Section 2: would it be worthwhile to point out that MODIS measures around 1:30pm? The following sentence will be modified to: All these instruments orbit in tandem measuring the same volume of air within minutes of each other, that is, by design, these measurements are collocated; their equatorial crossing time is ~1:30pm.

- P2/L27/L33: the estimated errors are mentioned here, could you include also whether this is a systematic or random uncertainty / error? Both of these estimates are random errors. This will be reflected in the text.

- P3/L11: "as the number of the number": correct? The sentence will be changed to: It also shows its associated standard deviation as well as the number of single observations (MODIS pixels) …

- P4/L7: "data below 200m or above 4km": I could of course look at the cited article, but maybe it can be better explained here. Does this mean if the inversion layer is below/above, the sonde is not used? Or is this data below 200m not used in the integral (which would be more of an issue)?
We only exclude this data in the BL height determination analysis, the sentence will be modified to: As in Millán et al. (2016), **during the inversion height determination**, we exclude all the data below 200 m or above 4 km **to avoid artifacts caused by temperature inversions near the surface as well as to avoid free-tropospheric features**. **Further**, we use only robust inversions, that is, …

- P4/L8: How much data is removed in these screening steps? And did you try to loosen the threshold of the three methods to be within 200m to increase the sample size?
We did not quantify how much data was screen by those steps, because such a stringent threshold is needed to ensure that the radiosondes used represent a well behave boundary layer profile. As shown in the following figure, relaxing such threshold could result in using sonde profiles where the boundary layer is not well defined.



Examples of relative humidity (RH) sonde measurements: (left) a robust inversion and (right) an unrobust inversion. Color lines display the boundary layer height determined using the location of the minimum vertical gradient of RH (red), the location of the minimum vertical gradient of specific humidity *q* (blue), and the location of the maximum vertical gradient of potential temperature Θ (green). Figure 3 from Millán et al (2016) doi: 10.1175/JAMC-D-15-0143.1

- P4/L24: The RO processing articles seem rather outdated, I assume that the processing uses some more recent algorithms, e.g. 1DVar (I might be wrong though). And where does the data come from, UCAR? The version 2.6 though is not a UCAR identifier as far as I can see. It would also be interesting to get more info on the humidity background if a 1DVar was used. Otherwise, RO will not provide very accurate info on humidity in the mid troposphere. The data comes from the JPL retrieval algorithm, for which those are the most current references. The following sentence will be modified to: In particular we use version 2.6 **of the JPL processing algorithm.**

- P6/26: When using ERA-I data, is exactly the same data at the same time used, or is this a larger data set? The following sentence will be added at the end of that paragraph: Note that for each region, ERA-Interim data from the nearest synoptic time (0,6,12,18 UT) to the measurement local time was used.

- Figure 1: Please use different color range for BL-CWV and its std dev, to allow more visualization of the std dev values. After careful consideration, we decided that it was best to leave the colorbar as it is. This way, the reader gets an immediate sense of the size of the standard deviation, as opposed to have to look at the ranges of two colorbars.

- Table 2: What is the last column? It seems not the number of obs. It is the correlation between BL-CWV and the number of observations. The caption of the table will be change to: Climatological (top) and interannual (bottom) correlation coefficients between BL-CWV and **SST, LTS BL-CTH and the number of observations**.

- Figure 4: There appear to be also negative CWV in this plot, is that found in certain regions/for low CWV values? Yes and no, that is, they do tend to occur more at low CWV values but not only at those places. Any reason why this is not excluded from the data set? Seems to point to the MODIS overestimation you mentioned. We believe that excluding these points could lead to a high bias of the BL-CWC. The following sentence will be added after the MODIS overestimation discussion: The overestimation of the MODIS CWV above the clouds could lead to negative values in the AMSR-MODIS dataset (as can be seen in Figure 4). However, we do not recommend that these negatives values are excluded of any analysis of the AMSR-MODIS dataset because some negative values will be due to the noisy nature of the MODIS measurements over cloudy pixels, and excluding those will lead to biasing high.

- Figure 5: These rectangular/boxes are from a publication in 1993. Some question on that: Why is the Australian not included? The Australian region is a relatively weaker SC region and hence is harder to interpret. Why not use all identified regions in that article(or at least the mid-lat marine stratus ones)? Again, because we focus in the robust SC regions. And last, is there a good reason to update the boxes with the latest data available? We have a better, higher resolution picture of our planet compared to 1993. Several studies have used these boxes and we wanted to compare against them without adding an extra level of complexity by changing the study regions, but the reviewer is correct, and a paper studying the impact of choosing different regions in the SC vicinities and its impact upon stratus amount and BL-CWV would be interesting.

- Figure 6, 7, 8, 9: my color print out shows almost exactly the same colors for Peruvian and Namibian. I cannot distinguish them at least. On screen it is okay. We do have high end color printers, thus I assume this might also happen for others printing it out. When zooming into Figure 5, I noted that the boxes have the same color, but I think that is not really necessary and limits your color options. The color for the Namibian region will be changed for a darker purple.

- Figure 6: Maybe I missed it somewhere in the discussion, but this has no NH/SH shift, but that is found in Figure 7 and others. In 6, it seems they all peak around July/August. The difference is not as visible in Figure 6 as it is in Figure 7, however, it is there. As stated in the text (P6 line 19 of the first version): The annual cycle us notably stronger in the Peruvian and Namibian regions with maxima during August *(as opposed to July)* and the peak lasting from June to November.

- Figure 6: Is the result better visible if it is normalized to the total in the area?
Below is the figure as requested by the reviewer, as shown is really similar to the Figure shown in the original paper. The reason is because the area of the regions is really similar (all regions shown are 10 by 10 degrees) which is roughly 1e6 KM2 (the number we used originally to scale them. In responding to this question, we realize that the y-label was wrongly shown as counts*1e6 when it should have been counts / 1e6. This will be updated in the new manuscript.

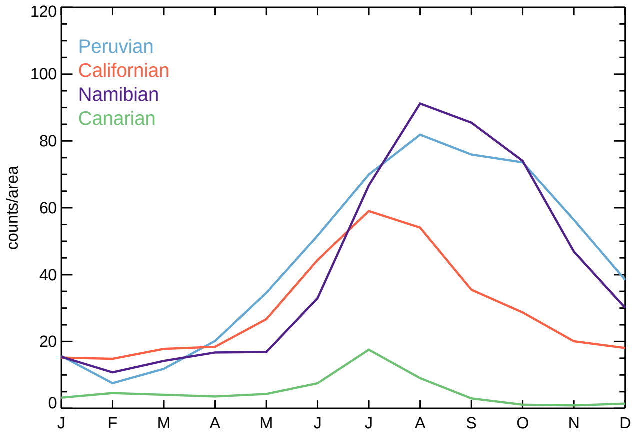


Figure: Annual cycle of the total AMSR-MODIS number observations (scaled by the total area of each region) for the regions delimited in figure 5 by the rectangular boxes

Also, these regions are different in size and cover different subsidence, does that have any impact?

As mention above, they are really similar in size, and while they do have a somehow different subsidence (see figure below) , the counts are not consistently correlated with subsidence (in this case using omega500 from ERA-Interim), as LTS is (as shown in Figure 7 of the manuscript).

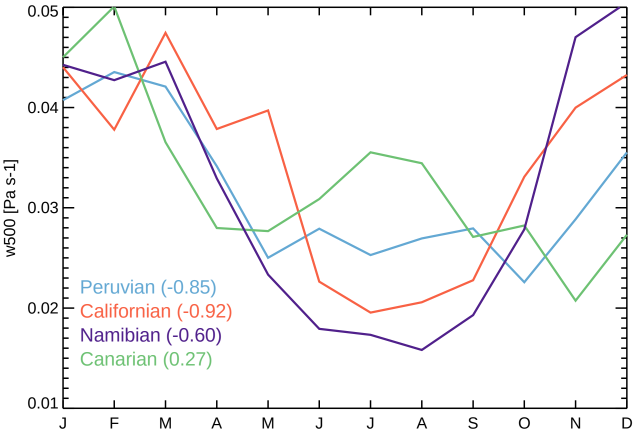


Figure: Annual cycle of w500 (from ERA-Interim) for the regions delimited in figure 5. The number of brackets are the correlation coefficient between the annual cycle of the number of observations and these w500 cycles.

We will add in the manuscript: Other parameters (MODIS CTP, AMSR SST, ERA-Interim w500 and ERA-Interim Surface Pressure) were analyzed in a similar manner but none of them were strongly correlated with the number of observations across the four regions used here.

- Generally, is there any impact visible of the different AMSR instruments used here? We didn't find any, that is, we analyze time series of the AMSR total CWV and SST globally as well as in the regions study and did not find any visible impact / discrepancies. The following sentence will be added in the paragraph describing the AMSR instruments: Note that, no discrepancies nor visible impacts were found in time series from these two instruments.

Response to Anonymous Referee #2

General comments:
This paper presents an assessment of boundary layer water vapor from satellite data. The method applied is based on a 16-year dataset of collocated near-infrared and microwave satellite observations.

In general, the paper is well structured, and provides some new interesting results. However, it needs some minor revisions before it can be published.

Specific comments:
p.2, l.4-5: you should mention that the datasets are derived from **satellite** observations
That sentence will be changed to: The aim of this study is to show results from a ~16 year boundary layer column water vapor (BL-CWV) dataset derived from the synergy of microwave and near-infrared **satellite** imagery.

p.2, l.30-31: Other months do not show this inconsistency? What are the reasons for that?
As explained to reviewer 1, the issue was a one-off coding error on the MODIS processing algorithm (per the personal communication with Richard Frey). We will modify the following sentence: Instead, version 6.1 was used for all December months **as recommended by the MODIS team. A full reprocessing of the AMSR-MODIS dataset using MODIS version 6.1 (or the latest MODIS version) is left for a future AMSR-MODIS version.**

p.2, l.34-35: Is this error (between 5 and 10 %) the error of the near infrared channels? Or the error of CWV? What about the error between cloudy and cloud-free cases? Is there a dependency on solar zenith angle?
This is the error of the CWV, the sentence will be changed to: In particular, we use the CWV estimated using near-infrared channels. These CWV values have an estimated random error between 5% and 10% [Gao et al 2013].

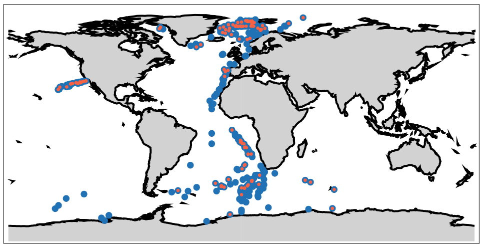
There is no literature describing any differences between cloudy and clear-sky cases nor any dependence for solar zenith angle for the near IR CWV product. We will add the following sentence which will follow immediately after the Gao et al 2013 citation: These errors may have a solar zenith angle dependence as found for other MODIS products [i.e., Horvath et al (2013) , Grosvenor et al (2014)] and may worsen under cloud conditions, as such, we assume the 10% error through-out.

p.3, l.6-7: Do you mean that you use only clouds that have been classified as “only liquid”? It is not clear here how you deal with mixed-phase clouds. The whole sentence should be rephrased for better clarity.
The sentence will be changed to: We only use the clouds which have been classified, by the cloud thermodynamic phase classification algorithm (Plattnick et al 2015), as liquid. This is a completely re-written algorithm which instead of using a linear sequential structure, as in version 5, uses a voting discrimination logic to identify the cloud thermodynamic phase as ice, liquid or undetermined (Marchant et al. 2016).

p.3, l.12-13: The monthly standard deviation of BL-CWV depends strongly on the variability of the boundary layer height (CTH). Have you checked this dependence?

Yes, there is a strong dependence between CTH and the BL-CWV as expected. In a future study we will exploit its dependence to explore different bulk BL-CWV characterization (i.e. Stephens 1990- cropped at the CTH, a well-mixed model, a piecewise model, etc).

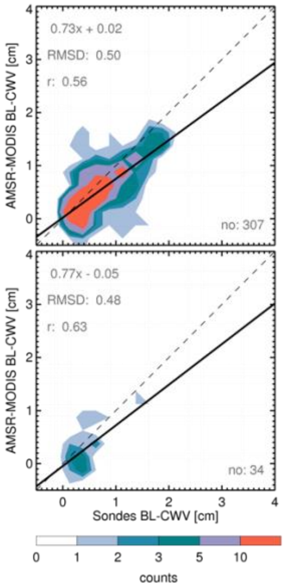
p.3, l.31: Did you only use Arctic/Antarctic radiosondes? Which latitude belts did you include?
The following figure will be added:



Caption: Map showing the geolocations of the radiosondes used in this study. Blue dots display the radiosondes that fulfill the criteria used in Figure 3-top while red dots display the subset that fulfill the criteria of Figure 3-bottom.

p.4, l. 10ff (and Fig. 2): I think that the variability within 6 hours is much larger than over 10 km. I guess most of the uncertainty reduction in the 1 hour/1km analysis comes from the shorter time range.
Note, that we do not have a 1hour/1km we display a 10km/6hours and a 1km/6hours.
Have you tried to keep 10 km (or even more) and reduce the temporal distance to 1 hour? In addition, 1 km drift of radiosondes is easily reached already within the boundary layer, therefore, I would suggest to neglect this “strong 1km” criterion and rather focus on temporal matching.

Below is a figure similar to figure 2 (in the original draft) keeping the 10km and reducing the temporal distance to 1 hour (bottom panel). As can be seen, these criteria result in only 34 matches and hence we prefer our previous one. Note that increasing the spatial threshold to 20km only results in 43 matches.



p.4, l. 20-23: It is known that GPS-RO data are missing some lower level inversions (especially below 1000 m above ground). How do you deal with this fact? Does it introduce a bias in your comparison?
As pointed by the reviewer in his/her next comment, we only use GPS above the inversion (when an inversion can be found) and subtract that estimate from the total CWV from AMSR.

p.5, l.3 (and Fig. 4): It is a bit misleading that you call the algorithms “AMSR-MODIS” and “GPS-RO”. This suggests that the GPS-RO algorithm is independent, however you use GPS-derived CWV above the inversion and then subtract it from AMSR total column. Therefore, you are not comparing independent data here. Please comment on that!
We will change the section name to AMSR - GPSRO to avoid misleading the reader. Also, the first sentence will read: As cross-validation, we use **AMSR - GPSRO** data. The **GPSRO** technique uses phase delays ...

Also, we will add the following sentence: As such, a comparison between AMSR-MODIS and AMSR - GPSRO, is, in essence, a comparison between MODIS water vapor above the clouds and the GPSRO water vapor above the BL inversion layer.

p.5, l.7: Although slope and RMS decrease, the correlation coefficient also decreases. Do you have an explanation for that?

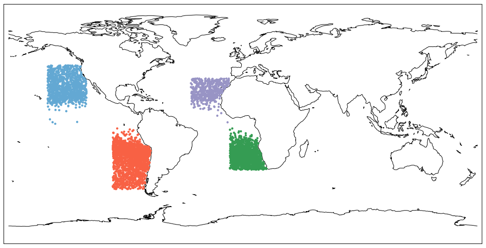
The RMSD depends on the values compared. If normalized, for example, by the mean of the AMSR-GPSRO values. The NRMSD are 0.50 and 0.49 for the sharpness parameter threshold of 2.5 and 3 respectively, that is to say, almost identical. We will update Figure 2 and 4 of the previous draft to use the NRMSD.

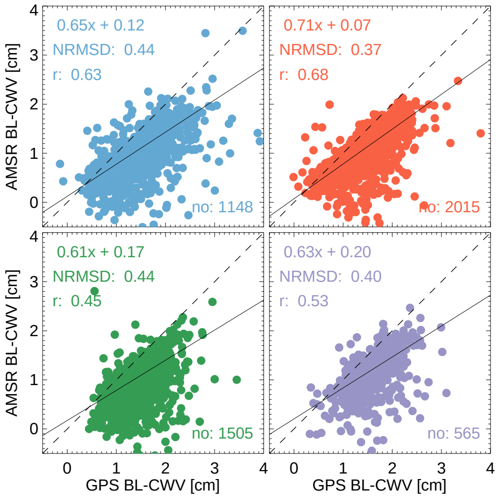
We will change the radiosonde comparison sentences to: The best-fit line has a slope of 0.73, a normalized (by the mean of the sondes values) root mean square deviation (NRMSD) of 0.69, and a correlation coefficient of 0.56 …
And to: By decreasing the coincidence criteria distance from 10 to 1 km (Figure 3-bottom) it is possible to improve these metrics (the best-fit line slope becomes 0.75, the NRMSD 0.59, and the correlation coefficient 0.71) but the total number of matches decreases from 307 to 124.

We will also change the GPSRO text to: By increasing the sharpness parameter requirement from 2.5 to 3.0 (Figure 5-bottom) the relationship between these two datasets improves with the best-fit line slope becoming 0.71 and the correlation coefficient 0.54. The NRMSD (in this case normalized by the mean of the AMSR-GPSRO values) remains nearly identical at ~0.5. However, the total number of matches decreases from ~23500 to ~750.

p.5, l.10-18: Is it possible that different viewing geometries or high solar zenith angles play a role in the uncertainties? If so, did you make separate analyses for different solar geometries or for different regions of the Earth?

As shown in the figures below (using a sharpness parameter of 2.5), there is some variation in the agreement between the two datasets per region, but not high enough to strongly indicate a viewing geometry/geolocation bias (at least in the stratus regions where most of the AMSR-MODIS observations are located, i.e. Figure 5 of the original draft).





p.6, l. 33-34 (and Fig. 7): What is the reason for the lower LTS in the Canarian region? Is it due to frequent advection of unstable air masses from the Saharan desert?

As discussed by Klein and Hartmann (1993), the SST for the Canarian region are about 3 to 5 degrees warmer than in the California region (this can be seen in figure 8 of the original manuscript) while the 700mb temperatures are really similar which results in a lower LTS.

p.7, l.3: Why did you reverse the order of the regions here (compared to p.7, l.1)? We will change the order to be the same as in p7 l.1.

p.7, l.21-25: What are the model constraints? Vertical temperature structure? CWV? CTH?

The model description will be changed to:
Figure 10 shows the measured annual cycle for BL-CWV, as well as the derived one from a simple well-mixed boundary layer model as the one described by Millan et al (2016), assuming a surface relative humidity of 80\% **and using the AMSR SST temperature, the MODIS CTH and ERA-Interim surface pressure as constraints**. These cycles were both normalized by their respective maximum values.

p.7, l.26-27: I cannot see an overestimation since you are plotting normalized values in Figure 9. It would be good to see absolute values from the model! Does the magnitude of the overestimation is in line with the findings in Figures 2 and 4?

The overestimation is not entirely explained by it, another reason is because the well mixed layer model is a oversimplification of the BL water profile, and it normally overestimates the BL-CWV. The following will be added: **and in part because of the simplistic representation of the boundary layer humidity profile by such a model.**

After consideration, we decided to show the figure with the normalized version to highlight that the water vapor observations can be, in general, be interpreted by a simple model as opposed to highlight the deficiencies of the model.

p.8, l.18: You are mentioning only here the restrictions of your method to homogeneous cloud fields during daylight. Does that affect the overall validity of your results? Do you expect a diurnal cycle?

The need of homogeneous cloud fields is mention in p3 l4: That is, we aim to identify homogenous fields of clouds in the MODIS data. Further, it is also mention, in section 4.1: High number of observations means that uniform liquid cloud fields were found consistently in such areas, and can be interpreted as …

The daylight restriction is implicit in P2 l5,6: Near infrared imagery provides the water vapor above the clouds (by measuring the solar radiation reflected near the 0.94-um water vapor band) while microwave imagery …

These restrictions do not affect our result because in regions where the boundary layer is well defined, the clouds tend to be homogenous, i.e. stratus clouds.

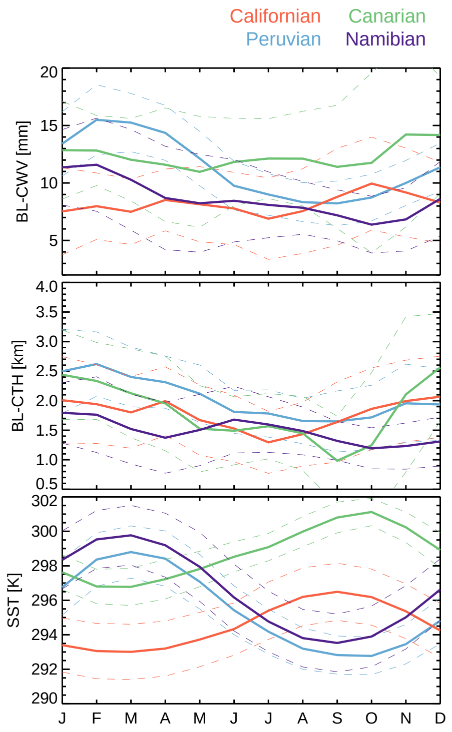
Although there is a boundary layer diurnal cycle, this does not affect our results, the AMSR-MODIS dataset simply captures the boundary layer state at around 1:30pm. This will need to be considered in future studies using the AMSR-MODIS dataset as with any other dataset measuring any atmospheric parameter with diurnal cycle.

The AMSR and MODIS equator crossing time will be mention in the following sentence: All these instruments orbit in tandem measuring the same volume of air within minutes of each other, that is, by design, these measurements are collocated; **their equatorial crossing time is ~1:30pm.**

p.8, l.23-25: This sentence (That is version2.0 (…) algorithm) is not necessary in the summary. The sentence will be deleted from the summary.

Figure 8: Please provide information on the monthly variation of BL-CWV and BL-CTH, e.g. showing error bars or box-and-whisker plots

Below is the figure showing the standard deviation.



To avoid cluttering, the following figure will be used with the following sentence in the caption: The numbers shown are the average standard deviation per region.

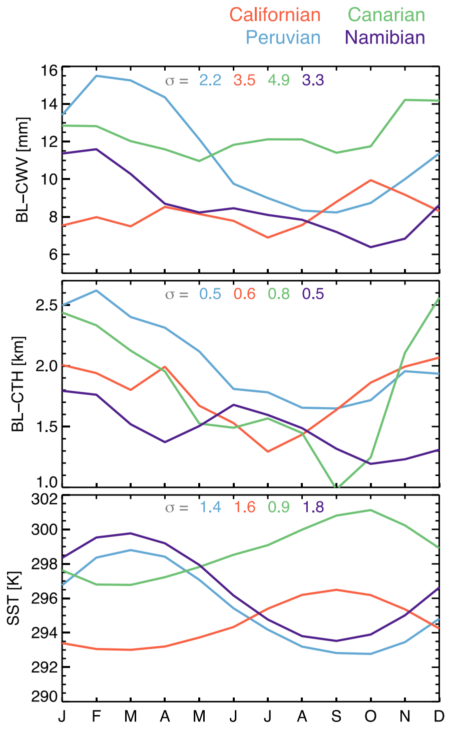


Figure 9: Since you plot normalized values, the unit [mm] is not correct! Thanks for spotting this, it will be changed to mm mm-1.

Technical corrections:

p.1, l.20: replace “are” by “is” Done
p.1, l.24: “processes” Done
p.3, l.22: “criterion”, not criteria We believe that criteria is the right option since we use a temporal condition, a spatial condition, and, in the case of GPS-RO, a sharpness parameter condition.
p.3, l.24: “gridding” Done
p.3, l.25: “represent” (not “represents”) Done
p.5, l.1: “coast” (not “cost”) Done
p.6, l.31, p.7, l.10: “4 K” _(without degree sign) Done
p.8, l.14: “remain”, not “remains” Done
p.8, l.32: “over”, not “on” Done
p.8, l.32: “Sc-Cu”: You never introduced these acronyms: we changed to Stratocumulus to Cumulus
Fig. 9 (caption): please correct: Normalize**d** (…) The number**s** (…) coefficient**s** (…) Done

References:

Horvath et al (2013) - doi:10.1002/2013JD021355
Grosvenor et al (2014) - doi:10.5194/acp-14-7291-2014
Marchant et al. (2016) – doi:10.5194/amt-9-1587-2016

Variability of Bulk Water Vapor Content in the Marine Cloudy Boundary Layers from Microwave and Near-Infrared Imagery

Luis F. Millán[1], Matthew D. Lebsock[1], and Joao Teixeira[1]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

Correspondence to: L. Millán (luis.f.millan@jpl.nasa.gov)

Abstract. This work uses the synergy of collocated microwave radiometry and near-infrared imagery to study the marine boundary layer water vapor. The Advanced Microwave Scanning Radiometer (AMSR) provides the total column water vapor, while the Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared imagery provides the water vapor above the cloud layers. The difference between the two gives the vapor between the surface and the cloud top, which may be interpreted as the boundary layer water vapor under certain conditions. As a by product of this algorithm, we also store cloud top information of the MODIS pixels used, a proxy for the inversion height, as well as the sea surface temperature and total column water vapor from the AMSR measurements. Hence, the AMSR-MODIS dataset provides several of the variables associated with the boundary layer thermodynamic structure. Comparisons against radiosondes, and GPS-Radio Occultation data demonstrate the robustness of these boundary layer water vapor estimates. We explore the annual cycle of the number of observations as a proxy for stratus cloud amount, in well known stratus regions; we then exploit the 16 years of AMSR-MODIS synergy to study for the first time the annual variations of the boundary layer water vapor in comparison to the sea surface temperature and the boundary layer cloud top height (equivalent to the inversion height) climatologies, and lastly, we explore the climatological behavior of these variables on stratocumulus-to-cumulus transitions.

1 Introduction

The boundary layer may be defined as the lower part of the troposphere that is directly influenced by the presence of the Earth's surface through turbulence. This layer mediates the exchanges of energy, momentum, water, carbon, and pollutants between the surface and the rest of the atmosphere and responds to surface forcing with a timescale of about an hour or less (Stull, 1988). Further, boundary layer processes are also intimately coupled with low clouds, such as stratocumulus. Stratocumulus is the most common cloud type covering around one-fifth of the Earth's surface (with mostly four-fifths of them located over the ocean) and thus have a profound impact on Earth's energy balance, primarily through solar radiation reflection (e.g., Wood, 2012). As such, boundary layer processes are crucial for understanding cloud-climate feedback mechanisms (e.g., Teixeira et al., 2011).

Despite their importance, boundary layer processes are still not well represented in weather and climate models. For example, differences in the response of low clouds to warming scenarios are responsible for most of the spread in model-based estimates of equilibrium climate sensitivity (Bony and Dufresne, 2005; Randall et al., 2007) and this spread appears to be attributable to how cloud, convective, and boundary layer processes are parameterized in such models (Boucher et al., 2013). However, one major issue in the development of accurate boundary layer parameterizations is the lack of global measurements.

The aim of this study is to show results from a ∼16 year boundary layer column water vapor (BL-CWV) dataset derived from the synergy of microwave and near-infrared satellite imagery. Near-infrared imagery provides the water vapor above the clouds (by measuring the solar radiation reflected near the 0.94-μm water vapor band) while microwave radiometry provides information on the total column water vapor (by measuring at the water vapor absorption line near 22 GHz). As shown by Millán et al. (2016), the difference between their water vapor information provides an estimate of the BL-CWV when the cloud top is capped at the boundary layer top.

Variability in the boundary layer water vapor plays an important role in the evolution of clouds and precipitation. Some field campaigns (e.g., Crum and Stull, 1987; Weckwerth et al., 1996, 2004) have provided some information about its temporal and spatial distribution in a few regions but its global variability and impact on clouds is still not properly understood. For example, subtle fluctuations in the vertical profile of water vapor appear to be associated with recurring stratocumulus and cumulus regimes (Betts and Boers, 1990). Further, several studies have shown that boundary layer water vapor is a critical quantity required for forecasting the initiation of convection (Crook, 1996; Ziegler and Rasmussen, 1998; Fabry, 2006; Martin and Xue, 2006). The combination of microwave and near-infrared imagery provides a unique capability to study the column water vapor in the planetary boundary layer.

2 Measurements

In this study, the AMSR-MODIS BL-CWV dataset version 2 is used. This dataset was produced merging passive microwave and near-infrared CWV measurements as part of a NASA Making Earth System Data Records for Use in Research Environments (MEaSURES) project. In short, BL-CWV was found by subtracting the CWV above the clouds estimated by the Moderate Resolution Imaging Spectroradiometer (MODIS) from the total CWV estimated by Advanced Microwave Scanning Radiometer (AMSR) instruments. In particular, we use AMSR-E, AMSR-2 and AQUA MODIS data which allow us to estimate the BL-CWV from 2002 to date; except for a gap between April 2011 and July 2012 when AMSR-E stopped operating and AMSR-2 became operational. Note that, no discrepancies nor visible impacts were found in time series from these two instruments.

The AMSR instruments are dual-polarized conically scanning microwave radiometers with channels measuring in between 6.9-89 GHz. They provide day and night estimates of total CWV over the oceans with an estimated random error of ∼0.6 mm (Wentz and Meissner, 2000). Through-out this study we used the Remote Sensing Systems (REMSS) CWV retrievals, in particular version 7, which aggregates these estimates to a quarter-degree spatial resolution. MODIS is an imaging spectroradiometer with 36 channels spread through-out the visible, near-infrared, and infrared. Here, we use version 6.0 except during December when cloud top height values were found to be unphysically large and inconsistent with the other months [R. Frey, Personal Communication]. Instead, version 6.1 was used for all December months as recommended by the MODIS team. A full reprocessing of the AMSR-MODIS dataset using MODIS version 6.1 (or the latest MODIS version) is left for a future AMSR-MODIS version. In particular, we use the CWV estimated using near-infrared channels. These CWV values have an estimated random error between 5% and 10% (Gao and Kaufman, 2003). These errors may have a solar zenith angle dependence as found for other MODIS products (i.e. Horváth et al., 2014; Grosvenor and Wood, 2014) and may worsen under cloud conditions, as such, we assume a 10% error through-out.

All these instruments orbit in tandem measuring the same volume of air within minutes of each other, that is, by design, these measurements are collocated; their equatorial crossing time is ∼1:30pm. The MODIS retrievals of above cloud water vapor have poor height registration when the cloud is either thin or broken. To alleviate these biases several flags as well as proximity tests are applied to remove pixels with intrapixel heterogeneity and/or high clouds as specified by Millán et al. (2016). That is, we aim to identify homogeneous fields of liquid clouds in the MODIS data. Version 2 is the second public release of the AMSR-MODIS data. The only difference against version 1 is that high clouds are masked out using the cloud phase optical properties. We only use the clouds which have been classified, by the cloud thermodynamic phase classification algorithm (Platnick et al., 2015), as liquid. This is a completely re-written algorithm which instead of using a linear sequential structure, as in MODIS version 5, uses a voting discrimination logic to discriminate the cloud thermodynamic phase into ice, liquid or undetermined (Marchant et al., 2016). AMSR-MODIS Version 1 instead screened only pixels where cirrus or aerosols were detected using the 1.38-μm high-cloud flag (MYD35).

During the processing, the algorithm uses the MODIS level 2 products in their native grid (i.e. MODIS pixels with a 1 km size at nadir) before binning the data into a 1° by 1° grid. We produce daily and monthly files. Figure 1 shows an example of a BL-CWV daily as well as a monthly composite. It also shows its associated standard deviation as well as the number ~~of the number~~ of single observations (MODIS pixels) used in each grid. Note that, as a by product of the BL-CWV algorithm, we also save the cloud top height (BL-CTH), the cloud top pressure (BL-CTP) and the cloud top temperature (BL-CTT) of the MODIS pixels used, as well as the sea surface temperature (SST) and total CWV from AMSR in the same grid. As such, the AMSR-MODIS dataset provides several of the variables associated with the bulk boundary layer thermodynamic properties. Monthly files were constructed aggregating the daily files neglecting pixels which daily standard deviation was greater than 0.2 cm. This threshold mostly rejects pixels in the intertropical convergence zone (ITCZ) where the boundary layer is not well defined.

3 Comparisons with other observations

In this section the accuracy of the AMSR-MODIS V2 BL-CWV measurements is assessed through comparisons with radiosondes and Global Positioning System Radio Occultation (GPSRO) measurements. For these comparisons, we consider only observations that are collocated geographically and temporally. The coincidence criteria used varies and is stated in each subsection below. Note that throughout these comparisons we use the AMSR-MODIS level 2 data (that is, we use the data before gridding it), to allow a better comparison. In analyzing these comparisons, it is important to bear in mind that each of the observations used is sampling different volumes; sondes are precise in-situ measurements which represent conditions at a local point, AMSR-MODIS level 2 product estimates the boundary layer conditions within a pixel size of 1 km at nadir, while GPSRO samples through the limb of the atmosphere, averaging over large horizontal distances of ∼200 km. Hence, geophysical variability will inevitably complicate the interpretation of such comparisons.

3.1 Radiosondes

In the comparison shown here we used sondes from two field campaigns: (1) the Alfred Wegener Institute (AWI) Polarstern laboratory campaign with more than 50 expeditions to the Arctic and the Antarctic (König-Langlo and Marx, 1997) since 1982 and (2) the Marine Atmospheric Radiation Measurement (ARM) GPCI Investigation of Clouds (MAGIC) campaign with approximately 20 round trips between Los Angeles and Honolulu during 2012-2013 (Kalmus et al., 2014; Zhou et al., 2015). Figure 2 shows the location of the radiosondes used.

To compute the BL-CWV from these measurements, we first identified the boundary layer inversion height and then integrated the specific humidity profile from that height to the surface. We use three different methods to find the inversion: the location of the minimum vertical gradient of specific humidity, the location of the minimum vertical gradient of relative humidity, and the location of the maximum vertical gradient of potential temperature. As in Millán et al. (2016), during the inversion height determination, we exclude all the data below 200 m or above 4 km to avoid artifacts caused by temperature inversions near the surface as well as to avoid free-tropospheric features. Further, we use only robust inversions, that is, those inversions where the boundary layer inversion height estimates of the three methods agree within 200 m, which mostly occur in stratus regions.

Figure 3-top shows the scatter between AMSR-MODIS and radiosonde BL-CWV within ±10 km and ±6 h. The best-fit line has a slope of 0.73, a normalized (by the mean of the sondes values) root mean sqaure deviation (NRMSD) of 0.69, and a correlation coefficient of 0.56, which suggests a reasonable but imperfect agreement between the two datasets. By decreasing the coincidence criteria distance from 10 to 1 km (Figure 3-bottom) it is possible to improve these metrics (the best-fit line slope becomes 0.75, the NRMSD 0.59, and the correlation coefficient 0s.71) but the total number of matches decreases from 307 to 124. Despite the scatter and the bias between the datasets, we find these results encouraging. The scatter was to be expected due the inherently noisy nature of the AMSR-MODIS product and because we do not know the extent to which the sonde measurements are representative of the average BL-CWV in the MODIS pixel.

3.2 AMSR - GPSRO

As cross-validation, we use AMSR - GPSRO data. The GPSRO technique uses phase delays in the GPS signals collected from a receiver on board of a low Earth orbiting satellite to derive profiles of refractivity. From these profiles, humidity in the middle and lower troposphere can be derived. In particular we use GPSRO data from the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) constellation. A description of the measurements and the retrieval technique can be found in Kursinski et al. (1995), Kursinski and Hajj (2001), and Hajj et al. (2002). The accuracy of these measurements is around 10 to 20% below 7 km and 5% or better in the boundary layer (Kursinski et al., 1995). In particular we use version 2.6 of the JPL processing algorithm.

To compute the BL-CWV from GPSRO we follow a similar methodology as in the AMSR-MODIS dataset. First, we match-up the GPSRO measurements with AMSR. As coincidence criteria we assume a match when any GPSRO lands within an AMSR footprint and ±6 hours. Then, following Ao et al. (2012), we identified the boundary layer inversion height as the minimum vertical gradient of the refractivity, which corresponds to the height where the refractivity changes most rapidly, and integrate the humidity profile from that height *upwards* to compute the CWV above the inversion height. Lastly, we subtract these estimates from the AMSR total CWV to compute the BL-CWV. As such, a comparison between AMSR-MODIS and AMSR - GPSRO, is, in essence, a comparison between MODIS water vapor above the clouds and the GPSRO water vapor above the BL inversion layer.

As an additional constraint we use the sharpness parameter, defined as the minimum refractivity gradient relative to the RMS value of the gradient averaged over the bottom 6 km of the atmosphere (see Ao et al. (2012) for more information), to identify regions where the BL inversion is well defined. As discussed by Ao et al. (2012), we found that the sharpness parameter is largest over the eastern subtropical oceans where stratocumulus occur (see Figure 4), with maximum average values of around 2.7 near the coast of Chile. The smallest sharpness parameters can be found in the ITCZ where the boundary layer is not well defined.

Figure 5-top shows the scatter between AMSR-MODIS and GPSRO BL-CWV using as coincidence criteria ±10 km and ±6 h and a sharpness parameter value greater than 2.5. Again, despite a fair amount of scatter and bias, the degree of agreement between the two datasets lends confidence in the usefulness of the AMSR-MODIS BL-CWV. By increasing the sharpness parameter requirement from 2.5 to 3.0 (Figure 5-bottom) the relationship between these two datasets improves with the best-fit line slope becoming 0.71 and the correlation coefficient 0.54. The NRMSD (in this case normalized by the mean of the AMSR-GPSRO values) remains nearly identical at 0.5. However, the total number of matches decreases from ∼23500 to ∼750. This improvement arises because when using a larger sharpness parameter we are ensuring that most pairings are in the stratus regions (see Figure 3 for reference) where the AMSR-MODIS technique should work better. A larger sharpness parameter also reduces the range of the BL-CWV comparison by excluding the high values found under cumulus regimes. This makes the comparison ranges (that is, in Figure 4-bottom) similar to the ones found in the sonde comparison where the sondes used are restricted to stratus regions by using the robust inversion criteria. That is, when the three different methods to find the inversion (explained in section 3.1) agree within 200m.

Through these comparisons, a consistent picture emerges suggesting either an underestimation of the AMSR-MODIS BL-CWV or an overestimation of the radiosonde and GPSRO BL-CWV. An underestimation of the AMSR-MODIS BL-CWV has two possible reasons, an underestimation of the total CWV by AMSR and/or an overestimation of the MODIS CWV above the clouds. We found an excellent agreement between the AMSR total CWV versus the radiosondes measurements (not shown), with a strong correlation coefficient (0.94), a best-fit line slope of 1.06 and an RMS deviation of 0.28 cm. This suggest that there may be an overestimation of the MODIS CWV above the clouds. The retrieval of BL-CWV above clouds is complicated by the fact that the near IR radiation penetrates the cloud layer. The multiple scattering of the light within the cloud increases the optical path length of the cloud and should result in an overestimate in water vapor above the clouds. The MODIS algorithm does not account for this effect and as a result the cloudy pixels are flagged with marginal quality assurance. The overestimation of the MODIS CWV above the clouds could lead to negative values in the AMSR-MODIS dataset (as can be seen in Figure 5-top). However, we do not recommend that these negatives values are excluded of any analysis of the AMSR-MODIS dataset because some negative values will be due to the noisy nature of the MODIS measurements over cloudy pixels, and excluding those will lead to biasing high.

We believe that a consistent overestimation of the radiosonde and GPSRO BL-CWV is unlikely due to the sharp gradients associated with the boundary layer inversion but we do suspect that uncertainties in determining such inversion are one likely culprit causing some of the scatter shown in Figures 3 and 5. In some cases, it is difficult to determine the boundary layer inversion height in the radiosonde and in the GPSRO data because several alternating dry and moist layers may be present in the measurements. In those cases, there is no guarantee that the algorithms chosen will identify the correct height, choosing instead a residual layer or a dry intrusion, which will lead to an overestimation or underestimation, respectively, of the BL-CWV estimated by the radiosondes or GPSRO data. von Engeln and Teixeira (2013) have shown that using different methods to estimate the boundary layer inversion height can lead to significantly different results even when using the same original datasets. For example, a consistent overestimation of the boundary layer inversion height (at least in the radiosonde cases) might be possible because as shown by Seidel et al. (2010) finding the inversion using the location of the minimum (maximum) vertical gradient of relative humidity (potential temperature) consistently yield higher PBL height estimates than other methods. Nevertheless considering the boundary layer geophysical variability (for example, the short response time of the boundary layer), the different sampling volumes associated with each technique, and the uncertainties in determining the boundary layer inversion height, we conclude that AMSR-MODIS BL-CWV, sondes, and GPSRO BL-CWV measurements are in good agreement.

4 Results

4.1 Climatology of stratus amount

Figure 6 shows the total number of observations found throughout the AMSR-MODIS dataset from 2002 to 2017. High number of observations means that uniform liquid cloud fields were found consistently in such areas, and can be interpreted as a qualitative proxy for stratus cloud fraction amount. Overlaid on this map are contours displaying the mean vertical velocity at 500 hPa (ω_{500}) from ERA-Interim (Dee et al., 2011) showing regions of large scale subsidence and convective regions. As expected, high number of observations are found in subtropical eastern oceans, in regions where stratocumulus clouds frequently occur (e.g. Klein and Hartmann, 1993; Wood, 2012). These subtropical regions are characterized by relatively cold sea surface temperature, strong subsidence, and well defined temperature inversions at the boundary layer (see for example, the high values of the sharpness parameter shown in Figure 4). High number of observations can also be found in regions where stratus clouds frequently occur (e.g. Teixeira, 1999) like over the arctic, over the southern ocean, and off east coast of the continents in the northern hemisphere. The lowest number of observations are found in the deep tropics, particularly in convective regions where the presence of non-uniform cumulus and also obscuring high clouds associated with deep convection decreases considerably the probability of finding uniform liquid cloud fields. Hence, the observations in this tropical region, where the boundary layer is not well defined, are not particularly reliable.

Climatological annual cycles of the number of observations for the regions shown in Figure 6 are shown in Figure 7. These regions are subtropical stratus locations taken from Klein and Hartmann (1993) and listed in table 1 for clarity. The annual cycles in the Californian and Canarian regions are similar with maxima during July and the peak lasting from June to August, however, the Canarian region has far fewer observations (i.e. unobscured stratus clouds). The annual cycle is notably stronger in the Peruvian and Namibian regions with maxima during August and the peak lasting from June to November. Overall, the annual cycle of the number of observations is in good qualitative agreement with the climatology of marine stratus compiled from ship-based weather observations by Klein and Hartmann (1993) or the climatology of low clouds derived from 5 years of CloudSat and CALIPSO data by Muhlbauer et al. (2014).

Previous studies have suggested that the seasonality of this type of clouds largely follows the lower tropospheric stability (LTS) (Klein and Hartmann, 1993; Richter, 2004; Wood and Bretherton, 2006; Richter and Mechoso, 2006). Figure 8 shows the annual cycle of LTS taken from the ERA-Interim reanalysis. LTS is defined as the difference between potential temperature at 700 hPa and the temperature at the surface. The LTS relation can be theoretically derived from the energy balance equation for the boundary layer (Chung et al., 2012) and can be thought of as a proxy for the strength of the inversion capping the boundary layer; in principle, a strong inversion is more effective at trapping humidity in the boundary layer, which will gradually accumulate and reach saturation, hence, enhancing cloud cover. As displayed, the Canarian LTS annual cycle is similar to the Californian one but ∼4 K lower throughout the year, which as suggested by Klein and Hartmann (1993) may result in the significantly reduction of stratus in such region. These LTS annual cycles are similar to the ones shown or described by Klein and Hartmann (1993). More interestingly, Figure 8 also shows the correlation coefficient between the number of observations and LTS in each of these regions. As expected, relatively high values can be found in most regions (0.77, 0.8, 0.91, and 0.93 for the Californian, Canarian, Namibian, and Peruvian regions respectively). Interannual correlations, that is, correlations based upon the monthly time series as opposed to the climatological data, also display relative high correlations,with values of 0.76, 0.77, 0.87, and 0.85 for the Californian, Canarian, Namibian, and Peruvian regions, respectively. Note that for each region, ERA-Interim data from the nearest synoptic time (0,6,12,18 UT) to the measurement local time was used, further, other parameters (MODIS CTP, AMSR SST, ERA-Interim ω_{500} and ERA-Interim Surface Pressure) were analyzed in a similar manner but none of them were strongly correlated with the number of observations across the four regions used here.

4.2 Climatology of BL-CWV

Figure 9 shows the annual cycle for BL-CWV, SST, and BL-CTH taken from the AMSR-MODIS dataset. Only the Peruvian and Namibian region display a significant BL-CWV annual cycle with a maximum to minimum differences of 8 and 6 mm, respectively; displaying a clear sinusoidal signature (specially in the Peruvian region) with maxima in February and minima during the fall. In the other regions, the maximum to minimum BL-CWV difference is only 3 mm throughout the year with no well defined minima or maxima. All regions display a clear SST annual cycle, with maximum to minimum differences close to ∼4°K. As with the LTS annual cycles shown in Figure 8, these SST annual cycles agree with the ones shown or described by Klein and Hartmann (1993). The BL-CTH annual cycles display a lot of variability, with no clear discernible pattern among the regions. The Canarian and Peruvian regions show the greatest maximum and minimum differences with 1.5 and 0.9 km respectively.

Table 2 shows the climatological and interannual correlation coefficients between the BL-CWV annual cycle and the ones found for BL-CTH, SST, LTS, and the number of observations. Only the Peruvian and Namibian regions display high correlation coefficient (that is, $|r| > 0.7$), at least in the climatological correlations, between these parameters. In those two regions the seasonal cycle strongly follows a cycle of modulation of the SST, which is negatively correlated with the LTS, and positively correlated with boundary layer depth, and bulk boundary layer water vapor content. This pattern is also true with weaker correlation in the Californian and Canarian regions which may be due to the smaller seasonal amplitude of the cycles in these regions.

Figure 10 shows the measured annual cycle for BL-CWV, as well as the derived one from a simple well-mixed boundary layer model as the one described by Millán et al. (2016), assuming a surface relative humidity of 80% and using the AMSR SST, the MODIS CTH and the ERA-Interim surface pressure as constraints. These cycles were both normalized by their respective maximum values. The modeled BL-CWV does resemble the BL-CWV measured one, particularly in the Peruvian and Namibian regions, where the correlation coefficients between the modeled and measured BL-CWV are 0.99 and 0.95 respectively. This suggests than in the most robust subtropical stratocumulus regions key properties such as water vapor content can be represented by a simple mixed-layer model. Note, however, that the well-mixed model consistently overestimates the measured BL-CWV in part due to the underestimation of the AMSR-MODIS product as shown by Figure 3 and Figure 5 and in part because of the simplistic representation of the boundary layer humidity profile by such a model.

4.3 Stratocumulus to Cumulus transitions

To further analyze the data, we focused on typical Stratocumulus-Cumulus transects. In these transects, stratiform clouds typically reside above relatively cold waters near the coasts, below subsiding air, in shallow and normally well mixed boundary layers capped by a strong temperature inversion. As trade winds advect air toward the equator, the subsidence weakens and the sea surface gradually warms leading to an increase in heat and moisture fluxes and a rising and weakening of the inversion, resulting in trade wind shallow convective clouds and eventually in deep convective clouds (e.g., Teixeira et al., 2011).

Figure 11 displays the transects used. These transects were taken from Sandu et al. (2010), in particular the ones constructed using gridded mean climatological meteorological fields. Figure 12 shows the climatological SST, BL-CWV and BL-CTH along these transects. The Californian and Canarian transects display data from June, July, and August while the Peruvian and Namibian transects for September, October, and November. These months correspond to the ones used by Sandu et al. (2010) during their trajectory analysis. These are the periods where Klein and Hartmann (1993) found the highest cloud fraction in the stratocumulus region on each oceanic basin.

The Californian and Canarian, transects display the expected behavior with warmer temperatures towards the equator resulting in a systematic deepening and moistening of the boundary layer. The boundary layer cloud top height starts as shallow as 1.4 and deepens up to 2.4, or 2.5 km in the Californian and Canarian transects respectively. Similarly, the boundary layer column water starts as dry as 7 or 11 and moistens up to 22 or 25 mm, respectively. On the other hand, the Namibian and Peruvian transects do not display this “canonical” picture. Notably these southern hemisphere transects each cross the equator. In the Namibian transect, despite a clear increase in SST along it, BL-CTH remain constant, at around 1.5 km, throughout its entire length. On the other hand, BL-CWV shows a systematic moistening, starting as dry as 7 and going as high as 20 mm. In the Peruvian transect, despite a clear increase in SST, BL-CTH and BL-CWV remains constant (with values of 1.9, km and 10 mm) up to 2500 km into the transect; only deepening and moistening steeply due to a sharp jump in the SSTs as the transect crosses the ITCZ.

5 Summary

The synergy of AMSR and MODIS measurements provides the opportunity of estimating for the first time the column of water vapor inside the marine boundary layer, although the technique is limited to homogeneous cloud fields during daylight. The boundary layer water vapor information results from combining AMSR estimates of total column water vapor, which are unaffected by clouds, with those derived from MODIS near-infrared channels using solar radiation reflected by clouds, which estimate the water vapor above the clouds. In this study we discussed results from the second public release of the AMSR-MODIS dataset. The AMSR-MODIS dataset is available in daily and monthly composites with a 1° by 1° resolution. Monthly files were constructed aggregating the daily files but disregarding daily pixels with standard deviation greater than 0.2 cm. This threshold mostly rejects pixels in the ITCZ where the boundary layer is not well defined. As a by product of the BL-CWV algorithm, the AMSR-MODIS dataset also provides the BL-CTH, BL-CTP, and the BL-CTT of the MODIS pixels used, as well as the associated SST and total CWV from AMSR. As such, the AMSR-MODIS dataset provides many of the variables of interest for boundary layer studies.

We exploited 16 years of collocated AMSR and MODIS measurements to study the behavior of the number of observations as well as the behavior of the BL-CWV over well known stratus regions. Further, we also study the Stratocumulus to Cumulus transitions. The main findings can be summarized as follows:

- Comparisons between AMSR-MODIS BL-CWV against radiosondes and AMSR-GPSRO data were undertaken. A consistent picture emerges suggesting an underestimation of the AMSR-MODIS BL-CWV measurements most likely due to an overestimation by the water vapor column above the clouds by MODIS. However, considering the geophysical variability of the boundary layer, the different sampling volumes of each technique, as well as the uncertainties associated with determining the inversion height in the sondes and AMSR-GPSRO boundary layer estimates, we believe that the comparisons demonstrate the skill of the AMSR-MODIS boundary layer water vapor estimates to detect variability.

- In well know stratus regions, the annual cycle of the number of observations (a qualitative proxy for stratus cloud fraction amount) is in good qualitative agreement with the climatology of marine stratus compiled from ship-based weather observations by Klein and Hartmann (1993) and the climatology of low clouds derived from 5 years of CloudSat and CALIPSO data by Muhlbauer et al. (2014). Furthermore, as previous studies have suggested, in all the stratus regions the number of observations is well correlated with lower tropospheric stability showing the inclination of stratus (homogeneous clouds fields) to form under a strong capping inversion layer.

- In the most robust subtropical stratocumulus regions key properties such as water vapor content can be represented by a simple mixed-layer model.
- The Californian and Canarian stratocumulus to cumulus transitions displayed the “canonical” view of these transects with a gradual deepening and moistening of boundary layer as the sea surface temperature warm up towards the equator. On the other hand, the Namibian and Peruvian transects do not display this canonical behavior.

In summary, these results demonstrate that the AMSR-MODIS dataset provides useful information regarding the marine boundary layer, particularly over stratus regions. Further, the multi-sensor nature of the analysis demonstrates that there exists more information on boundary layer water vapor structure in the satellite observing system than is commonly assumed when considering the capabilities of single instruments.

Data availability. The AMSR-MODIS dataset can be found on the NASA Goddard Space Flight Center Earth Sciences (GES) Data and Information Services Center (DISC) website (http://disc.sci.gsfc.nasa.gov/) with “10.5067/MEASURES/AMDBLWV2” and “10.5067/MEASURES/AMMBLWV2” digital object identifiers for the daily and monthly data respectively. The data is stored in netcdf version 4 format. ERA-Interim reanalysis fields can be found at the ECMWF website (http://apps.ecmwf.int/datasets/).

Author contributions. Luis Millán wrote the AMSR-MODIS algorithm, and carried out the analyses. Matthew Lebsock and Joao Teixeira provided scientific expertise throughout all stages of the research.

Acknowledgements. The research described in this study was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

Ao, C. O., Waliser, D. E., Chan, S. K., Li, J.-L., Tian, B., Xie, F., and Mannucci, A. J.: Planetary boundary layer heights from GPS radio occultation refractivity and humidity profiles, Journal of Geophysical Research: Atmospheres, 117, doi:10.1029/2012JD017598, d16117, 2012.

Betts, A. K. and Boers, R.: A Cloudiness Transition in a Marine Boundary Layer, Journal of Atmospheric Sciences, 47, 1480–, doi:10.1175/1520-0469(1990)047<1480:ACTIAM>2.0.CO;2, 1990.

Bony, S. and Dufresne, J.-L.: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models, Geophysical Research Letters, 32, doi:10.1029/2005gl023851, 2005.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S., Sherwood, S., Stevens, B., and Zhang, X.: Clouds and Aerosols, book section 7, p. 571–658, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.016, www.climatechange2013.org, 2013.

Chung, D., Matheou, G., and Teixeira, J.: Steady-State Large-Eddy Simulations to Study the Stratocumulus to Shallow Cumulus Cloud Transition, Journal of the Atmospheric Sciences, 69, 3264–3276, doi:10.1175/jas-d-11-0256.1, https://doi.org/10.1175/jas-d-11-0256.1, 2012.

Crook, N. A.: Sensitivity of Moist Convection Forced by Boundary Layer Processes to Low-Level Thermodynamic Fields, Monthly Weather Review, 124, 1767–1785, doi:10.1175/1520-0493(1996)124<1767:somcfb>2.0.co;2, 1996.

Crum, T. D. and Stull, R. B.: Field Measurements of the Amount of Surface Layer Air versus Height in the Entrainment Zone, Journal of the Atmospheric Sciences, 44, 2743–2753, doi:10.1175/1520-0469(1987)044<2743:fmotao>2.0.co;2, 1987.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, doi:10.1002/qj.828, 2011.

Fabry, F.: The Spatial Variability of Moisture in the Boundary Layer and Its Effect on Convection Initiation: Project-Long Characterization, Monthly Weather Review, 134, 79–91, doi:10.1175/mwr3055.1, 2006.

Gao, B.-C. and Kaufman, Y. J.: Water vapor retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared channels, Journal of Geophysical Research: Atmospheres, 108, doi:10.1029/2002JD003023, 2003.

Grosvenor, D. P. and Wood, R.: The effect of solar zenith angle on MODIS cloud optical and microphysical retrievals within marine liquid water clouds, Atmospheric Chemistry and Physics, 14, 7291–7321, doi:10.5194/acp-14-7291-2014, https://doi.org/10.5194/acp-14-7291-2014, 2014.

Hajj, G., Kursinski, E. R., Romans, L. J., Bertiger, W. I., and Leroy, S. S.: A technical description of atmospheric sounding by GPS occultation, Journal of Atmospheric and Solar-Terrestrial Physics, 64, 451 – 469, doi:10.1016/S1364-6826(01)00114-6, 2002.

Horváth, Á., Seethala, C., and Deneke, H.: View angle dependence of MODIS liquid water path retrievals in warm oceanic clouds, Journal of Geophysical Research: Atmospheres, 119, 8304–8328, doi:10.1002/2013jd021355, https://doi.org/10.1002/2013jd021355, 2014.

Kalmus, P., Lebsock, M., and Teixeira, J.: Observational Boundary Layer Energy and Water Budgets of the Stratocumulus-to-Cumulus Transition, Journal of Climate, 27, 9155–9170, doi:10.1175/jcli-d-14-00242.1, 2014.

Klein, S. and Hartmann, D. L.: The Seasonal Cycle of Low Stratiform Clouds , Journal of Climate, 6, 1587–1605, doi:10.1175/1520-0442(1993)006<1587:TSCOLS>2.0.CO;2, 1993.

König-Langlo, G. and Marx, B.: The meteorological information system at the Alfred-Wegener-Institute, Climate and Environment Database Systems, M. Lautenschlager and M. Reinke, Eds. Kluwer, pp. 117–126, 1997.

Kursinski, E. R. and Hajj, G. A.: A comparison of water vapor derived from GPS occultations and global weather analyses, Journal of Geophysical Research: Atmospheres, 106, 1113–1138, doi:10.1029/2000JD900421, 2001.

Kursinski, E. R., Hajj, G. A., Hardy, K. R., Romans, L. J., and Schofield, J. T.: Observing tropospheric water vapor by radio occultation using the Global Positioning System, Geophysical Research Letters, 22, 2365–2368, doi:10.1029/95GL02127, 1995.

Marchant, B., Platnick, S., Meyer, K., Arnold, G. T., and Riedi, J.: MODIS Collection 6 shortwave-derived cloud phase classification algorithm and comparisons with CALIOP, Atmospheric Measurement Techniques, 9, 1587–1599, doi:10.5194/amt-9-1587-2016, https://doi.org/10.5194/amt-9-1587-2016, 2016.

Martin, W. J. and Xue, M.: Sensitivity Analysis of Convection of the 24 May 2002 IHOP Case Using Very Large Ensembles, Monthly Weather Review, 134, 192–207, doi:10.1175/mwr3061.1, 2006.

Millán, L., Lebsock, M., Fishbein, E., Kalmus, P., and Teixeira, J.: Quantifying Marine Boundary Layer Water Vapor beneath Low Clouds with Near-Infrared and Microwave Imagery, Journal of Applied Meteorolgy and Climatology, 55, 213–225, doi:10.1175/JAMC-D-15-0143.1, 2016.

Muhlbauer, A., McCoy, I. L., and Wood, R.: Climatology of stratocumulus cloud morphologies: microphysical properties and radiative effects, Atmospheric Chemistry and Physics, 14, 6695–6716, doi:10.5194/acp-14-6695-2014, https://www.atmos-chem-phys.net/14/6695/2014/, 2014.

Platnick, S., King, M. D., Meyer, K. G., Wind, G., Amarisinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Ridgway, B., and Riedi, J.: MODIS Cloud Optical Properties: User Guide for the Collection 6 Level-2 MOD06/MYD06 Product and Associated Level-3 Datasets, 1, 1–145, 2015.

Randall, D., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R. J., Sumi, A., and E., T. K.: Climate models and their evaluation., book section 8, pp. 589–662, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Richter, I.: Orographic influences on the annual cycle of Namibian stratocumulus clouds, Geophysical Research Letters, 31, doi:10.1029/2004gl020814, https://doi.org/10.1029/2004gl020814, 2004.

Richter, I. and Mechoso, C. R.: Orographic Influences on Subtropical Stratocumulus, Journal of the Atmospheric Sciences, 63, 2585–2601, doi:10.1175/jas3756.1, https://doi.org/10.1175/jas3756.1, 2006.

Sandu, I., Stevens, B., and Pincus, R.: On the transitions in marine boundary layer cloudiness, Atmospheric Chemistry and Physics, 10, 2377–2391, doi:10.5194/acp-10-2377-2010, https://www.atmos-chem-phys.net/10/2377/2010/, 2010.

Seidel, D. J., Ao, C. O., and Li, K.: Estimating climatological planetary boundary layer heights from radiosonde observations: Comparison of methods and uncertainty analysis, Journal of Geophysical Research, 115, doi:10.1029/2009jd013680, 2010.

Stull, R. B., ed.: An Introduction to Boundary Layer Meteorology, Springer Netherlands, doi:10.1007/978-94-009-3027-8, 1988.

Teixeira, J.: Simulation of fog with the ECMWF prognostic cloud scheme, Quarterly Journal of the Royal Meteorological Society, 125, 529–552, doi:10.1002/qj.49712555409, 1999.

Teixeira, J., Cardoso, S., Bonazzola, M., Cole, J., DelGenio, A., DeMott, C., Franklin, C., Hannay, C., Jakob, C., Jiao, Y., Karlsson, J., Kitagawa, H., Köhler, M., Kuwano-Yoshida, A., LeDrian, C., Li, J., Lock, A., Miller, M. J., Marquet, P., Martins, J., Mechoso, C. R.,

v. Meijgaard, E., Meinke, I., Miranda, P. M. A., Mironov, D., Neggers, R., Pan, H. L., Randall, D. A., Rasch, P. J., Rockel, B., Rossow, W. B., Ritter, B., Siebesma, A. P., Soares, P. M. M., Turk, F. J., Vaillancourt, P. A., Engeln, A. V., and Zhao, M.: Tropical and Subtropical Cloud Transitions in Weather and Climate Prediction Models: The GCSS/WGNE Pacific Cross-Section Intercomparison (GPCI), Journal of Climate, 24, 5223–5256, doi:10.1175/2011jcli3672.1, 2011.

von Engeln, A. and Teixeira, J.: A Planetary Boundary Layer Height Climatology Derived from ECMWF Reanalysis Data, Journal of Climate, 26, 6575–6590, doi:10.1175/jcli-d-12-00385.1, 2013.

Weckwerth, T. M., Wilson, J. W., and Wakimoto, R. M.: Thermodynamic Variability within the Convective Boundary Layer Due to Horizontal Convective Rolls, Monthly Weather Review, 124, 769–784, doi:10.1175/1520-0493(1996)124<0769:tvwtcb>2.0.co;2, 1996.

Weckwerth, T. M., Parsons, D. B., Koch, S. E., Moore, J. A., LeMone, M. A., Demoz, B. B., Flamant, C., Geerts, B., Wang, J., and Feltz, W. F.: An Overview of the International H2O Project (IHOP_2002) and Some Preliminary Highlights, Bulletin of the American Meteorological Society, 85, 253–278, doi:10.1175/bams-85-2-253, 2004.

Wentz, F. J. and Meissner, T.: AMSR ocean algorithm, version 2, Remote Sensing Systems Tech. Rep. 121599A-1, p. 66p, availablefromhttp://eospso.gsfc.nasa.gov/sites/default/files/atbd/atbd-amsr-ocean.pdf.], 2000.

Wood, R.: Stratocumulus Clouds, Monthly Weather Review, 140, 2373–2423, doi:10.1175/mwr-d-11-00121.1, 2012.

Wood, R. and Bretherton, C. S.: On the Relationship between Stratiform Low Cloud Cover and Lower-Tropospheric Stability, Journal of Climate, 19, 6425–6432, doi:10.1175/jcli3988.1, https://doi.org/10.1175/jcli3988.1, 2006.

Zhou, X., Kollias, P., and Lewis, E. R.: Clouds, Precipitation, and Marine Boundary Layer Structure during the MAGIC Field Campaign, Journal of Climate, 28, 2420–2442, doi:10.1175/jcli-d-14-00320.1, 2015.

Ziegler, C. L. and Rasmussen, E. N.: The Initiation of Moist Convection at the Dryline: Forecasting Issues from aCase Study Perspective, Weather and Forecasting, 13, 1106–1131, doi:10.1175/1520-0434(1998)013<1106:tiomca>2.0.co;2, 1998.

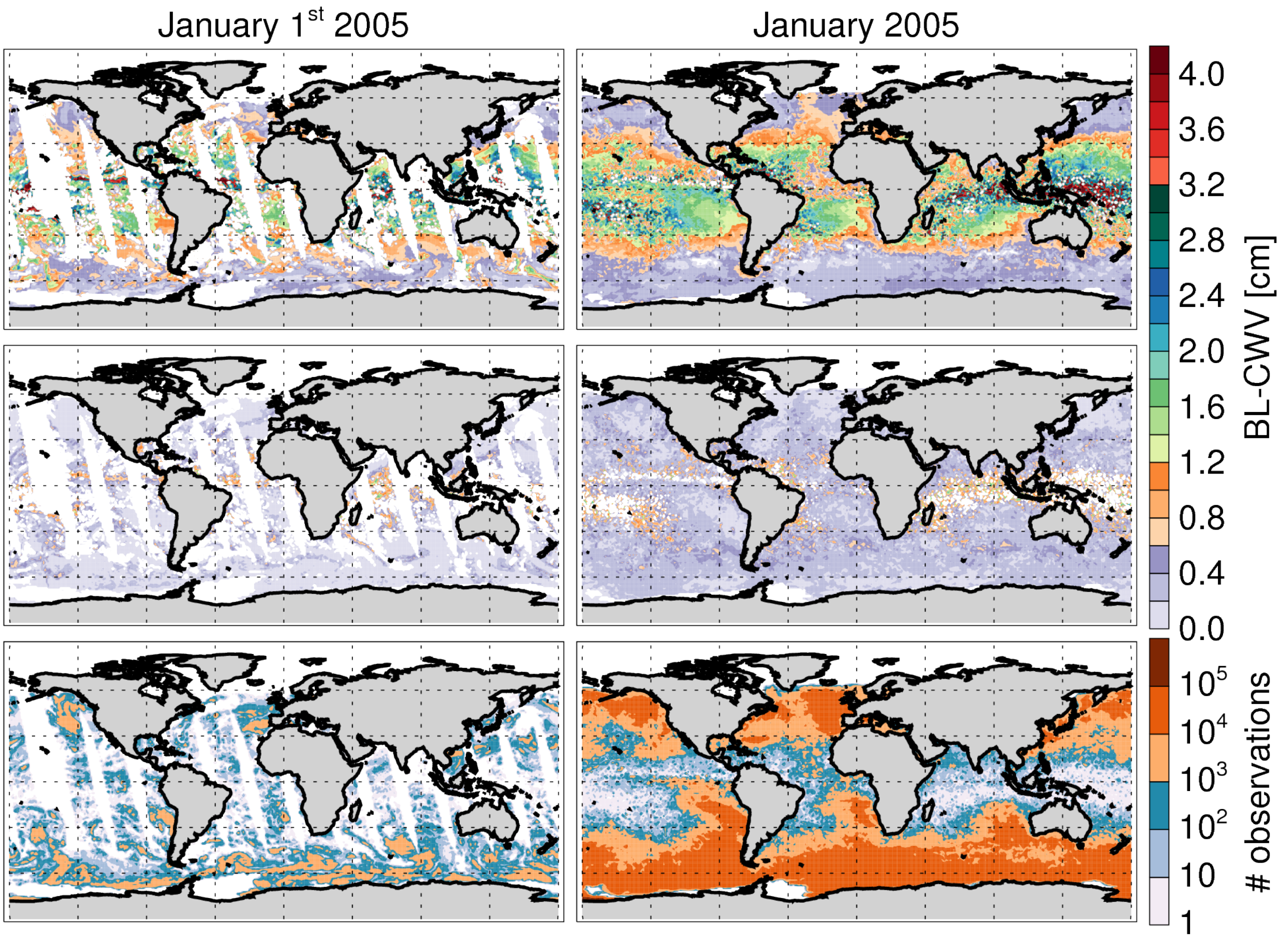


Figure 1. Example of daily (January 1st, 2005, left) and monthly (January 2005, right) composites of BL-CWV (top), its standard deviation (middle), and the number of observations used (bottom).

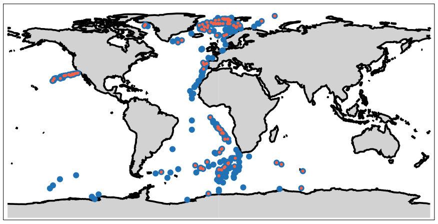


Figure 2. Map showing the geolocations of the radiosondes used in this study. Blue dots display the radiosondes that fulfill the criteria used in Figure 3-top while red dots display the subset that fullfill the critera of 3 3-bottom.

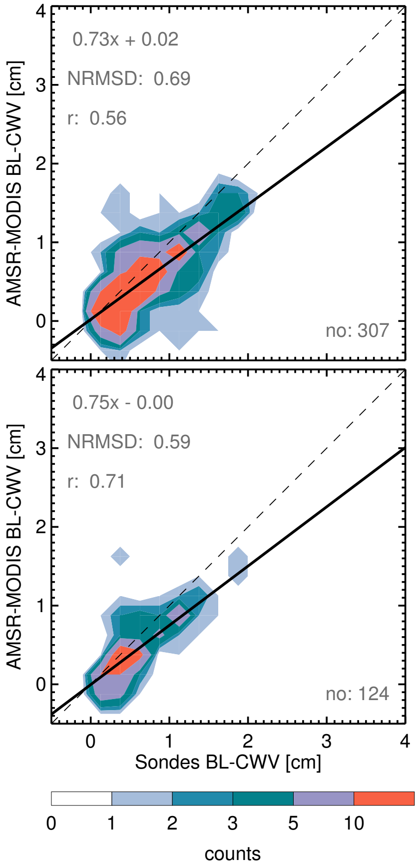


Figure 3. Sondes BL-CWV measurements scattered against the AMSR-MODIS BL-CWV estimates using ±10 km and ±6 h (top) and ±1 km and ±6 h (bottom) as coincidence criteria. The dashed black line is the one-to-one line. The solid black line displays a linear fit. The normalized root mean square deviation, the linear fit equation, the correlation coefficient R, and the total number of matches are shown.

Table 1. Geographical extent of the regions used in this study.

Region	Geographical boundaries
Peruvian	10°-20° S, 80°-90° W
Namibian	10°-20° S, 0°-10° E
Californian	20°-30° N, 120°-130° W
Canarian	15°-25° N, 25°-35° W

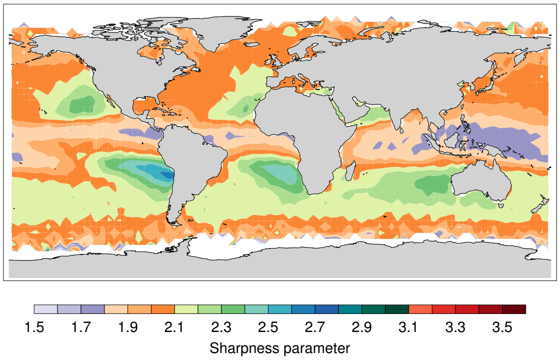


Figure 4. Sharpness parameter (relative minimum refractivity gradient) from 9 years (2006-2014) of the COSMIC data used on a 4° by 4° grid.

Table 2. Climatological (top) and interannual (bottom) correlation coefficients between BL-CWV and several other variables. Bold text indicates a high correlation coefficient (|r|>0.7)

Region	SST	LTS	BL-CTH	Number of Observations
Peruvian	0.95	-0.95	0.96	-0.98
Namibian	0.81	-0.76	0.81	-0.72
Californian	0.37	-0.27	0.48	-0.36
Canarian	0.06	-0.55	0.72	-0.26
Peruvian	0.95	-0.88	0.86	-0.92
Namibian	0.75	-0.63	0.82	-0.61
Californian	0.44	-0.31	0.63	-0.27
Canarian	0.12	-0.22	0.61	-0.10

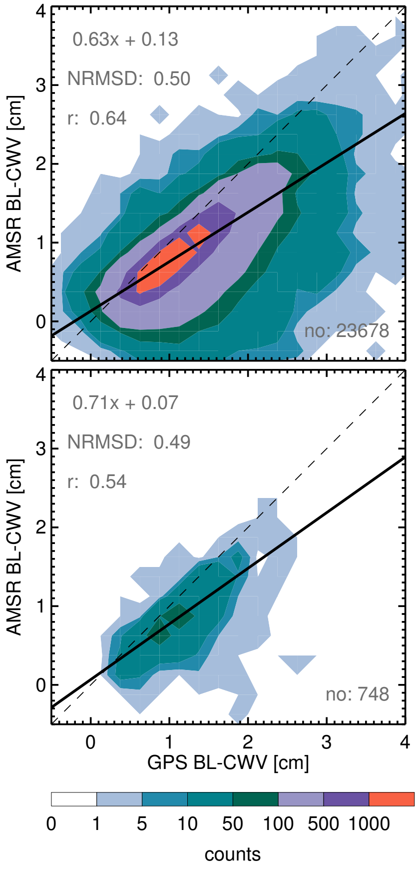


Figure 5. AMSR-GPSRO BL-CWV measurements scattered against the AMSR-MODIS BL-CWV measurements using ±10 km, ±6 h, and a sharpness parameter greater than 2.5 (top) and ±10 km, ±6 h, and a sharpness parameter greater than 3 (bottom) as coincidence criteria. The dashed black line is the one-to-one line. The solid black line displays a linear fit. The normalized root mean square deviation, the linear fit equation, the correlation coefficient R, and the total number of matches are shown.

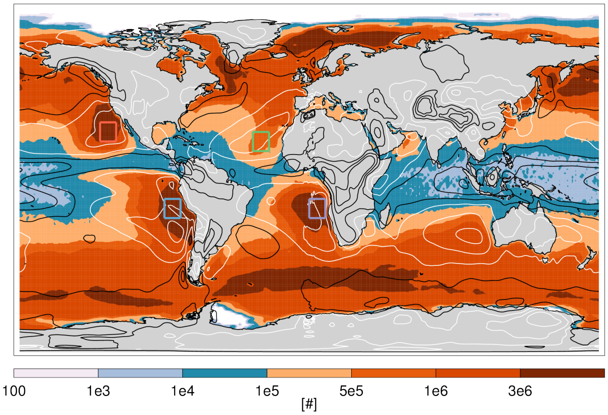


Figure 6. Number of observations found in the AMSR-MODIS dataset over 2002 to 2017. Overlaid contours display air vertical velocity at 500 hPa (ω_{500}) from ERA-Interim, with white contours at 0.01,0.03, 0.05 Pa s^{-1} denoting sinking of air and black contours -0.05,-0.03,-0.01 Pa s^{-1} denoting rising of air. A 2D smoothing has been applied to the ω_{500} fields. Color rectangular boxes identify regions with high amount of stratocumulus clouds. These locations are adopted from Klein and Hartmann (1993).

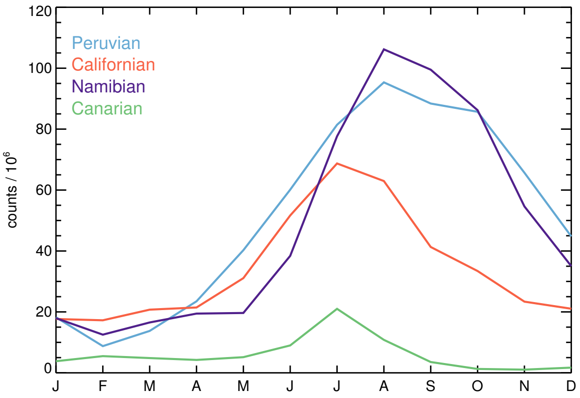


Figure 7. Annual cycle of the total AMSR-MODIS number of observations for the regions delimited in Figure 6 by the rectangular boxes.

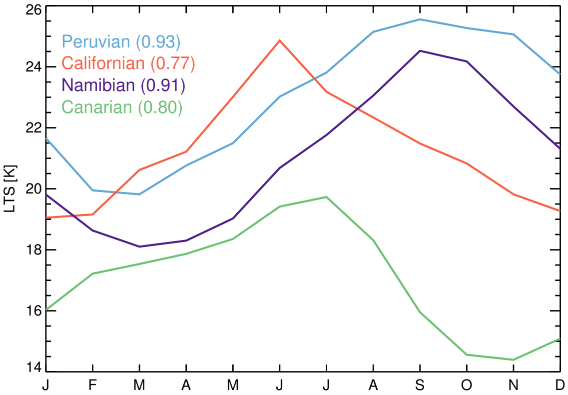


Figure 8. Annual cycle of LTS for the regions delimited in Figure 6 by the rectangular boxes. The number in brackets are the correlation coefficient between the annual cycle of the number of observations (shown in Figure 7) and these LTS cycles.

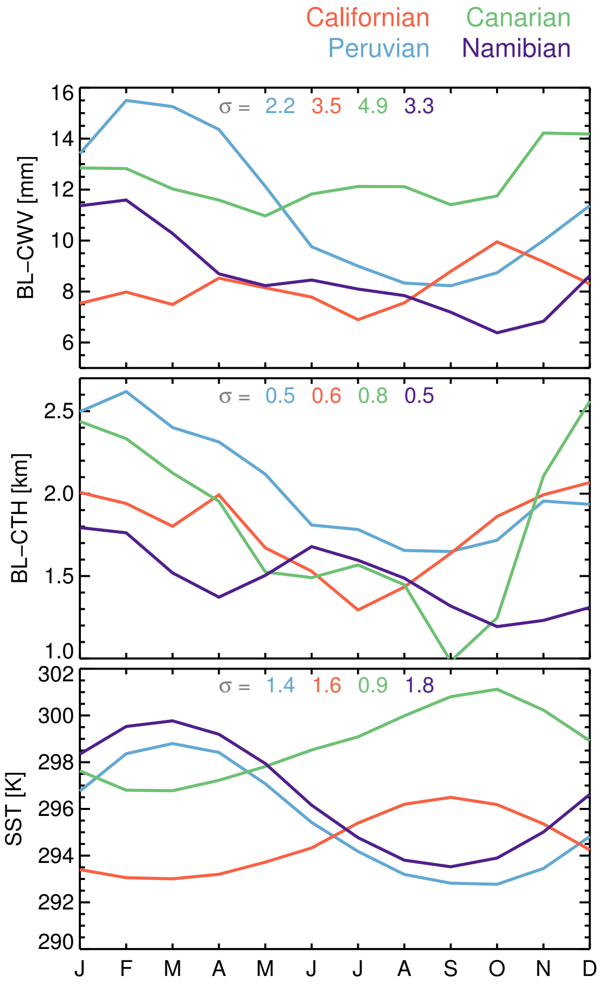


Figure 9. Seasonal cycle of BL-CWV, BL-CTH, and SST, for the regions delimited in the Figure 6 by the rectangular boxes. The numbers shown are the average standard deviation per region.

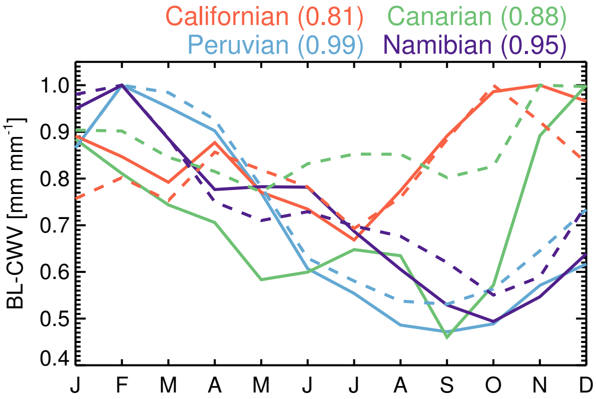


Figure 10. Normalized measured seasonal cycle of BL-CWV (solid lines), as well as derived from simple mixed layer model (dash lines) for the regions delimited in the Figure 6 by the rectangular boxes. The numbers in brackets are the correlations coefficient between the measured BL-CWV and the modeled one.

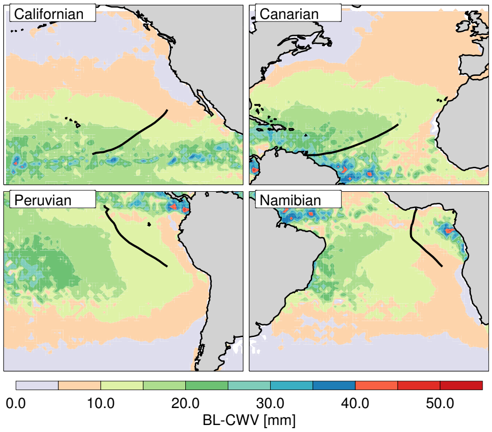


Figure 11. Transects along the climatological streamlines used in this study (taken from Sandu et al. (2010)). The contours show the climatological composite for all the AMSR-MODIS BL-CWV data available.

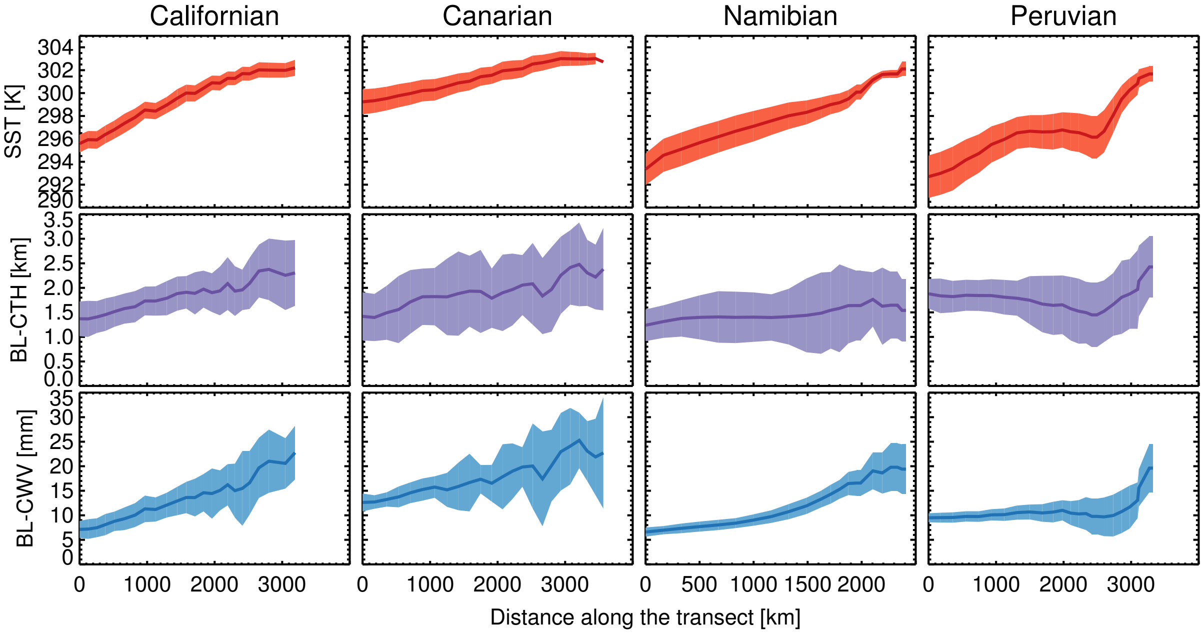


Figure 12. Climatological SST, BL-CTH and BL-CWV along the transects shown in Figure 11. The envelopes display the standard deviation.