Anonymous Referee #2

We thank the reviewer for taking the time and effort to review the manuscript and appreciate the comments. In this response, we aim to highlight more clearly the purpose, value, and novelty claimed in the manuscript.

General comment and recommendation

This manuscript presents a comparison of (correlations between) cloud properties and thermodynamic and dynamic fields derived from AIRS and MERRA data, with those derived in previous literature and derived in this paper from MODIS.

This paper is clearly not a comparison or a validation paper. Its purpose is to describe the synergistic use of previously validated data products from AIRS, MODIS, and CloudSat, together with MERRA reanalysis at the native temporal and spatial resolution, to investigate relationships between cloud microphysical and optical properties, and dynamical and thermodynamic fields. To our knowledge, at the time of submission, no attempt of this kind of approach has been made for investigating the marine boundary layer.

The manuscript comes across as rather unfocused, wandering between a variety of objectives, none of which end up convincingly presented.

This is a fair statement. We agree that we can tighten up the organization of the various components in the manuscript, and be more forthright with our conclusions and take home messages.

The manuscript appears to: a) evaluate AIRS and MERRA against MODIS and other cited products;

As stated above, this paper is clearly not a comparison or a validation paper. In fact, the only common field between the three instruments (AIRS, MODIS, CloudSat) and one reanalysis (MERRA) used in the manuscript is the vertical structure of relative humidity (Section 4.4 and Appendix A), which we will move to the Appendix in the revision per reviewer #1's suggestion. AIRS, MODIS, CloudSat, and MERRA each provide unique information that can be brought to bear on observing the subtropical MBL. We are not comparing common geophysical fields obtained between them. Our guiding philosophy is "Why not play to the strengths of each instrument?"

b) provide physical insights into what explains the transition by means of reflectance, optical depth, boundary layer depth and effective radius;

That is (partially) correct, although there are many other geophysical variables used, and various moments (mean, variance, skewness) that highlight certain aspects of MBL structure. However, we are not attempting to "explain the

transition" so much as rather present a new way to observe it.

c) present the AIRS cloud product that can best reflect the transition from stratocumulus to cumulus, where a variety of measures are tried out;

Yes. There are three cloud products in particular from AIRS that are used. (1) The AIRS cloud thermodynamic phase product is used to coarsely group together uniform closed cellular stratocumulus and broken, disorganized open cellular trade cumulus clouds. (2) The AIRS effective cloud fraction (ECF) is derived from infrared channels so it will have a different perspective of cloud cover compared to visible reflectance or optical thickness. (3) The AIRS visible channels are used to quantify the reflectance that is filtered by an AIRS visible cloud mask.

d) present different physical behaviours between four regions in which the transition between stratocumulus and cumulus occurs.

As this work progressed, it became apparent that the data revealed that each of the four regions has some subtle differences when contrasted against the other regions, but the NEA is the biggest outlier of the four. These differences might be reflected in dMSE, vertical velocity, liquid water cloud effective radius, so on and so forth. We did not attempt to explain why the four regions show such differences, which is well beyond the scope of this investigation and would undoubtedly require extensive numerical modeling experiments, and further investigation of several years of data.

Sometimes the lack of focus and a specific question of interest seems to shines through in the authors' writing, for instance, when they introduce new sections, which might represent choices that are "not optimal", but simply provide "a fresh look at available products",

It appears the reviewer is referring to line 13, page 8 for the quoted text. We offered specific reasons for why we chose the effective cloud fraction (ECF) variable from AIRS over the cloud fraction (CF) derived from the MODIS cloud mask, and also reflectance from AIRS over the optical thickness from MODIS. The reasons are described just above the quoted text:

lines 7-11, page 8: "The MBL depth exhibits clearer patterns in the ECF dimension rather than the cloud fraction dimension. The latter is more compressed and the gradients are weaker in both dimensions. The MBL depth is deepest for lower values of ECF, τ , and reflectance. In addition, the MBL depth also decreases for the most reflective clouds at a given value of ECF while this behavior is not observed for τ . We posit that an additional population of sub-pixel cumulus clouds is captured within the reflectance data that is not captured in τ data."

or when they describe that choosing which variables to plot in their joint pdfs is challenging.

It appears the reviewer is referring to line 2, page 8. The most honest way to go into this investigation is to ask what do we do when confronted with a choice from an enormous selection of available data? Dozens of geophysical variables are available from each instrument and reanalysis system. These variables can be plotted against each other in 1000s of combinations (or much more). The moments of these variables are also another dimensional choice, so to speak. On top of that, any field can be overlaid onto the two dimensions as done throughout the joint pdf figures. So where does one start? As we pointed out, our reasoning for starting where we did is found here:

Line 3, page 8: "Motivated in large part to link cloud and thermodynamic properties derived from infrared and visible bands…"

If the focus would be on presenting novel insights,

Since the reviewer is emphasizing multiple times the "novelty" of this work, after doing a word search we found only three instances of the word "novel" are used in the manuscript. Perhaps this word choice is unfortunate and we will remove accordingly in the revision.

I had expected that beyond abstract descriptions of behavior of different quantities in the joint pdf's the authors explain what this behaviour actually tells us about observed cloud fields.

The paper is not only about observed cloud fields. It is an attempt to describe a more holistic synthesis of the subtropical MBL from the point of view of A-train satellite observations and MERRA reanalysis built from native temporal and spatial resolution data. This includes winds, T/q/RH structure, the occurrence of light precipitation, vertical and horizontal motion, the depth of the MBL, and how they link to cloud properties. This paper is not about the physical causes of the stratocumulus to cumulus transition, but we do cite some of these papers in the Introduction.

If the focus would be on an evaluation of AIRS products, I would expect that the evaluation were more thorough and go beyond a comparison of seasonal averaged fields.

As stated above, this paper is clearly not a comparison, validation, or data evaluation paper for AIRS products. We have cited references throughout that point the reader to previous validation efforts that support the use of the data as shown in the manuscript.

What contributes to the wandering is that the authors use different data sets for different objectives as they present themselves.

The whole purpose of the paper is to use the different instruments and reanalysis

data sets as building blocks to construct a simultaneous point of view of the MBL, playing on the strengths of each instrument. The reviewer comment strongly suggests we can be much more clear and concise about our purpose. We will revise accordingly.

AIRS and MERRA are used for interpreting skewness measures and the transition between cloud types, along with comparisons of MODIS. Only MODIS is used for the purpose of evaluating effective radii in the two regimes, and what might physically or methodologically explain effective radii behaviour.

As far as we know, MODIS has the most useful, validated, tested, and investigated global retrieval of liquid water cloud effective radius available to the scientific community. AIRS does not provide one. CloudSat uses MODIS effective radius in its forward algorithm of retrieval products. MERRA is quite awful at clouds. MODIS is, for all practical purposes, the only ballgame in town.

In much of the authors assertions, previous literature is referenced, but often not explained.

We tried to be as comprehensive as possible with citing references for our statements. In the revision we will try and be as clear as possible as to why we are citing a particular work.

Because the manuscript does not present a novel insight and fails to convincingly argue for any method or the AIRS or MERRA datasets at providing novel insights, I recommend a rejection of the manuscript.

We would like to bring up that reviewer #1 had a different opinion: "This effort is commendable and valuable, combining these data at small scales can potentially reveal a lot about the relationships between clouds and their environment."

We hope that we will convince reviewer #2 the value of this work pending the revisions that we will make.

Specific comments

The title is unspecific (which satellite and reanalysis data sets?) and promises more insight than the paper offers.

There are plenty of papers that have general titles when there are numerous data sets described because the titles would be too long. (Also, for example, it is common that papers that use the CMIP archive do not reference particular models used.) The words "A satellite and reanalysis view" suggests that this is an observational study rather than a study that deduces the complex physical mechanisms at play in the MBL. We could consider changing it to "An A-train and MERRA view" if that helps.

The term cloud organization is not well-chosen, because the authors do not present results on cloud organization nor discuss what cloud organization means. The words organised and disorganised are repeatedly used throughout the manuscript, but mostly in reference to organised stratocumulus and disorganised trade-wind cumulus. Both stratocumulus and trade-cumulus can be organized and disorganized, depending on some definition of organization. Sometimes it seems the authors refer to homogeneous and heterogeneous, but mostly it seems that they refer to the two different cloud types.

In our revision we will be more concise. The particular choice of words was meant to follow on to the work of Muhlbauer et al. When we refer to cloud organization in the revision, we will be more specific and cite values of skewness rather than used 'organized' or 'disorganized'.

The word novel is repeatedly used, but seems an overstatement. In much of the manuscript, the authors confirm insights found in previous studies, and that citation list is long.

We found the word 'novel' was used only three times in the entire manuscript. We will select another word or rephrase accordingly.

We feel as though it is a particular strength of this approach that we are able to reaffirm a long list of previous findings and reference a large body of research.

One novelty that is argued for is the use of AIRS and MERRA datasets at their instantaneous native resolution. But to prove the suitability for these datasets for this kind of study, the authors qualitatively compare the morphology of the stratocumulus to cumulus transition from seasonal averaged AIRS and MERRA data with the morphology known from existing studies. I do not think a qualitative comparison of the seasonal mean transition tells us enough about how good AIRS and MERRA perform at their native resolution.

The seasonal averages were developed as a first order check on our methods and use of data. Since the seasonal averages agreed very well with previous research, that gave us confidence in moving forward with the joint pdfs. (Also, it was one of the only ways to compare with previous research since so few studies have looked at joint pdfs in the manner that we showed in figures 7-13.) In the revision we will be clear about why we start with the seasonal averages then proceed to the joint pdfs.

The authors make an argument for separating the cloud regimes stratocumulus and cumulus based on infrared-based thermodynamic phase (rather than by dynamical regime such as done in previous literature).

The motivation for this approach is found in the Methodology section. As this is a pixel based approach, we require that all ice cloud instances are removed, and we are confident that the AIRS phase product is more than sufficient.

Page 5, lines 6-7: "Removal of pixels containing mid- and high-level clouds helps to reduce ambiguities introduced by free tropospheric clouds and also a portion of the thermodynamic and dynamic variability associated with cloudy areas of synoptic-scale waves."

The dynamical approach is consistent with this approach in the sense that stratocumulus clouds show larger free tropospheric subsidence than the cumulus clouds. We will revise the manuscript accordingly to emphasize these points.

The thermodynamic phase provides information about whether just liquid or ice is present in the detected clouds.

We refer the reviewer to page 5, lines 11-12: "Jin and Nasiri (2014) showed that AIRS successfully identifies the presence of ice within the AIRS FOV in excess of 90% of the time when compared to CALIPSO thermodynamic phase estimates."

AIRS is an extremely radiometrically stable instrument with very strong sensitivity to cloud phase as discussed in Kahn et al. (2014) and Jin and Nasiri (2014) and citations within.

Based on a single scene in Figure 1 and 2 the authors argue that stratocumulus is well identified by those pixels that are detected as liquid, whereas trade-wind cumulus are those pixels that have an unknown thermodynamic phase. How do the authors know that this separation holds well for other scenes?

Please refer to above response. We have evaluated AIRS against CALIOP (Jin and Nasiri, 2014), and Kahn et al. (2015), J. Geophys. Res. in the case of MODIS phase. These evaluations were performed globally for large sets of observations.

After all, trade-wind cumulus are also made of liquid only, and it is unclear and not explained why they could not be identified as such in other scenes. It is also not clear for what purpose the two cloud types are separated here in this paper. Mostly this seems to be a proposition to use AIRS thermodynamic phase in future studies, but with insufficient evidence.

We agree that the delineation between stratocumulus (liquid) and cumulus (unknown) should be made clearer in the revision. Since the AIRS cloud phase is based on channel selection that exploits the differences in the index of refraction for liquid and ice, if the cloud amount in the AIRS pixel is small enough the spectral signature will be so small that it does not trigger a liquid test (see Jin and Nasiri, 2014). We do know that there is cloud in the pixel using the ECF field (validated using CALIPSO lidar, see Kahn et al., 2014), so what is happening is that none of the phase tests are triggered even though a small amount of cloud is there. These cases line up very well with trade cumulus in the four regions selected.

One aspect of the paper that prevents it from providing clear physical insights (if this were the main objective) is that the authors never explain what the skewness in reflectance or optical depth tells us about the nature of the cloud field that is observed (and this is true for many of the behaviours derived from the joint pdfs). The skewness measure has been used in previous studies, and can with some background of course be interpreted, but the authors never explicitly do. This makes the description of results rather abstract.

We will strengthen this aspect of the interpretation of the data. On the whole, the more skewed the reflectance is, the smaller the ECF is. When the reflectance is approximately Gaussian, the ECF is larger. The former is seen very clearly in the cumulus pdfs and the latter in the stratocumulus pdfs. Since there is such good separation between the two cloud types, they should be discussed separately. (This also should be considered as an independent confirmation of the sensitivity of the AIRS phase algorithm to cloud type.) Even for the same combination of reflectance and ECF in cumulus and stratocumulus pdfs for the MBL depth, the MBL depth is shallower for stratocumulus. The same is true for dMSE (more positive for stratocumulus than trade cumulus.) This is a really interesting result that shows there is cloud regime dependence even for the same value of ECF and reflectance, and that separation is facilitated by the AIRS phase algorithm categories liquid and unknown. We will revise the text accordingly to strengthen these discussion points.

The discussion in section 4.5 and the conclusions argue for both physical causes (precipitation) as well as retrieval-related biases (inhomogeneity) for the observed larger effective radius in cumulus clouds compared to stratocumulus. But whereas first is stated that (L24) "the observed increase in re is entirely consistent with environmental variability (winds/droplet growth/precipitation)", it is written further along that the greater inhomogeneity in such precipitating cumulus fields can cause assumptions used in retrievals to break down. Hence, should I trust the retrieved larger effective radii observed?

On lines 16-18 on the same page we state the following: "As these particular MODIS pixels are limited to successful retrievals only, we offer evidence that the increase in re is entirely consistent with environmental variability that is furthermore consistent with droplet growth and precipitation." Since precipitating retrievals might be more inhomogeneous than non-precipitating ones, that alone could be a cause of the increase in re using the MODIS look up table approach. We did not claim otherwise. We simply showed that these larger values of effective radius strongly correspond to occurrences of precipitation detected by CloudSat.

Given the comments of reviewers #1 and #2, we will investigate this further for the revision. We will attempt to map an inhomogeneity parameter onto the retrievals of effective radius that are both precipitating and not precipitating, for a range of wind speeds.

In the last paragraphs of section 4.6 and the summary, the authors argue two seemingly contradicting statements with which they end their manuscript. Namely, that three of the four regions studied show similar relationships and behaviours among cloud-related quantities and the (thermo)dynamic state, but also that the relationships are non-unique (can vary greatly), for which their datasets provide a good opportunity for further exploration. I understand the subtlety, but is this the best ending?

Good point. This ending needs some work and we will revise accordingly for the revision.