Response to reviewer comments, “Identifying meteorological influences on marine low cloud mesoscale morphology using deep learning classifications”, Mohrmann et al.

We thank both reviewers for their time and helpful comments on the manuscript. Based on their feedback, we have the following revisions to be manuscript:

Response to reviewer #1:

Comment (1), also (6): The reviewer raises a very good point regarding the confusing wording of the manuscript title, and in hindsight it is very understandable that a reader could be misled by it to think that deep learning was used in identifying meteorological influences, not merely in creating the classification dataset used. A title change is warranted, and we have altered the new title to be “Identifying meteorological influences on marine low cloud mesoscale morphology using satellite classifications”. As the reviewer points out, the description of the classification dataset is primarily carried out in Yuan et al. (2020), and so no emphasis on it needs to be given in the title of this manuscript. This manuscript is primarily focused on the meteorology, and so we did not spend too many words on describing the methods behind the classification or its accuracy, but we agree with the reviewer’s comment (both comments (1) and (6)) that the accuracy of the deep learning and a brief description of its method and accuracy is warranted. The following language has been added to section 2.1, paragraph 1, in the description of the classification dataset:

Old text: [list of scenes] …These scenes are then used to train a convolutional neural net, which in turn is run near-globally on such MODIS oceanic scenes., and a detailed description of the classification dataset and training can be found in Yuan et al. (2020).

New text: [list of scenes]. These categories were chosen by examining the category climatologies in Muhlbauer et al. (2014) and studying regions where there was little variability in category (primarily the tropics, where disorganized MCC dominated), and identifying additional commonly occurring cloud morphologies. These (clustered and suppressed Cu) were then added to the pre-existing cloud categories, along with homogeneous stratiform category initially used in Wood and Hartmann (2006). Examples of these types can be found in Figure 2.

The scenes are then used to train a convolutional neural net (CNN). The CNN input data is the image of scene visible reflectance. A full description of the machine learning training and model evaluation can be found in Yuan et al. (2020); the main results are that average model precision evaluated on a test set was approximately 93% across all types. Open-MCC had the lowest precision, most likely because it was the lowest-frequency category. The largest source of model confusion was between disorganized MCC and clustered Cu, which is unsurprising as these categories have similar appearance.

Comment (2): Please see the first paragraph added above, where we have added some more explanation about the origins of these cloud types. Additional information on the types is found in Yuan et al. (2020) and we do not think it would be beneficial to repeat too much of that discussion in this work.

Comment (3): We have added in the appropriate citation for CERES (Doelling et al., 2013)

Comment (4): We have added in the appropriate citation for CERES (Gelaro et al., 2017)

Comment (5): We have re-ordered the figures such that they are now in the order first mentioned in the paper.

Comment (6): Addressed in response to comment (1).

Comment (7): The following line has been added to section 3.1: "Panel (a) shows the fraction of scenes covered by the dominant cloud type for that grid box" to address this omission, thank you.

Response to reviewer #2:

Comment on l101, l133, l202: By mass continuity, the horizontal divergence at 700 hPa would only give us the *local* vertical divergence dw/dz (or dω/dp in pressure coordinated) at 700 hPa, and while we do not expect very strong gradients of large-scale divergence with height in the marine lower troposphere, this value will be somewhat sensitive to the level chosen. Using w_{700}/z_{700} results from the integration of dw/dz from the surface to 700mb, and so represents the mean horizontal divergence over that layer, making it more suitable as an estimate of large-scale divergence. To make this point clearer, we have amended this paragraph as follows:

"Note that this large-scale divergence is not the horizontal divergence at 700 hPa, but rather the mean divergence from the surface to the 700 hPa level; this follows from the mass continuity equation by considering a column of air from the surface (where vertical motion is 0) to 700 hPa. The terms large-scale divergence and 700 hPa subsidence are used interchangeably throughout; divergence is plotted instead of subsidence to allow for more straightforward comparison with surface divergence. As surface pressure varies with time, the second equality is only approximate."

Comment on l218: Corrected typo, thank you.

Comment on l218-220: The reviewer correctly interpreted this sentence, and are perhaps wondering if there was anything more to this point; we have added a parenthetical "(as expected)" to this sentence to indicate there is not deeper point being made.

Comment on l234: Corrected by adding reference.

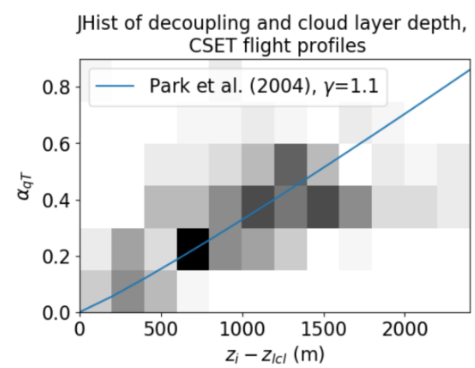
Comment on S3.5: regarding alpha_q, we have added the following text: . The parameter α_q is a measure of how much the upper boundary layer moisture resembles the lower FT as compared to the lower boundary layer; a value of 0 would indicate a perfectly well-mixed boundary layer, while a value of 1 would indicate a perfectly decoupled boundary layer where the upper BL moisture was equivalent to the lower FT moisture.

$$\alpha_{qT} = \frac{q_T(upper\ BL) - q_T(lower\ BL)}{q_T(lower\ FT) - q_T(lower\ BL)}$$

For a given profile, the thermal inversion height is estimated using the maximum in lapse rate, with the inversion being the layer where the lapse rate deviation from a moist adiabat was >25% of maximum deviation (this was tuned to agree with a visual assessment of the inversion layer and worked well for all profiles). Upper and lower BL in the q_T equation are taken as the top and bottom 25% of the BL depth, while the lower FT is taken as the 500m above the inversion top. While this method may not be the most precise in individual more complex cumulus cases with more spatially and vertically heterogeneous moisture profiles, we use it for consistency and reproducibility. We also note that a joint histogram analysis of α_q vs cloud layer depth (not shown) produced consistent results to Wood and Bretherton (2004) and Park et al. (2004).

Regarding question 2 on this section, this is clarified in the text added above; the inversion used for diagnosing MBL depth is always the strongest thermal inversion, which tended to occur at the top of the remnant cloud layer in trade-like shallow Cu profiles. The result is that the boundary layers appeared highly decoupled as this layer has been subject to ample dry entrainment during the transition. There were some cases where the upper layer was thoroughly erased by dry entrainment that the diagnosed inversion was much shallower. A sentence has been added acknowledging that this method may not be ideal for some cases and justifying our use.

Regarding question 3 on this section, we reprocessed the profiles to include a surface mixed layer depth using the (near-)surface-derived LCL, consistent with Wood and Bretherton (2004). This allows for a more apples-to-apples comparison with the figures referenced by the reviewer showing model results of the same quantities. We include the fit from Park et al, 2004 shown in all 3 previous papers (though we note that Neggers et al. used a lower gamma value). It certainly seems that our profiles are consistent with the results in Wood and Bretherton, though we do not have the sample size to say anything more insightful regarding the slight disagreements between the LES results and the plotted fit in de Roode et al. 2016. We have added a sentence (see above) stating the consistency with Wood and Bretherton (2004) / Park et al (2004).



Regarding the comparison to Lock (2009), we did briefly attempt to validate those results, but could not find a strong correlation between kappa and low cloud fraction. The most likely explanation for this is that in our observations, both kappa and cloud fraction are taken roughly simultaneously, whereas the simulations in Lock allow for cloud adjustment. It is also possible that kappa is sensitive to noise resulting from messy real-world profiles, and so we cannot say anything one way or another about its role in controlling cloud fraction.

Comment on Fig 2: added "Image scale is roughly 100 km across." to caption.

For the discussion of clustered vs disorganized MCC, the following text has been added to section 2.1: "The primary difference between these two types is that disorganized MCC represents a regime with cellular convection at some characteristic scale, though not obviously organized into open- or closed-cell regimes, while clustered Cu represents aggregated convection at a variety of scales within a scene. For distinguishing between these two types during manual labelling, scene large-scale context proved helpful."

Comment on Fig 5: This may be a monitor issue, though we do agree that for Figure 5, the colors are not as easy to distinguish, though as the reviewer notes it is a moot point as no disorganized MCC or clustered Cu occurs in this scene. The colors are selected for consistency with the rest of the plots, where they do not present an issue, and so we will keep them as is.

Regarding the differing color scales between the two divergence plots, we have updated the plots so that both ASCAT and MERRA surface divergence are on the same color scale.

Comment on l342: typo corrected, thank you.

Additional changes to the manuscript:

Added code/data availability, added author contributions.