Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-295-AC1, 2017

Geoscientific Model Development Discussions
Open Access

Interactive comment on "A multi-diagnostic approach to cloud evaluation" *by* Keith D. Williams and Alejandro Bodas-Salcedo

Keith D. Williams and Alejandro Bodas-Salcedo

keith.williams@metoffice.gov.uk

Received and published: 30 March 2017

11

Author's response to referee 1 on "A multi-diagnostic approach to cloud evaluation"

K. D. Williams and A. Bodas-Salcedo

March 30, 2017

1 Major comments from referee 1

1.1 Referee Comments

My primary concern is on the scientific focus of this study. The title of the paper seems to suggest that the aim of this work is to introduce "a multi-diagnostic approach to cloud evaluation". However, the paper has spent a lot of time on the inter-comparison of the two configurations of the UM model. I have no problem with whichever topic the study is designed to focus on, as both topics have their own values. However, since the study "tries" to cover two topics at a time, the discussions are somewhat lacking in depth. Therefore, the paper reads more like a report.

If the study is designed to focus on introducing a new multi-diagnostic approach, then a thorough introduction of this approach, including the developments of individual diagnostic methods (including necessary technical details), their merits and limitations, their applications in the literature, as well as a quantitative estimate of the uncertainties

Printer-friendly version

Discussion paper

of these methods, should be fully discussed. The authors have discussed some of the above mentioned aspects, but only to a very limited extent.

If the study is designed to focus on the evaluation of the simulations, then I have real trouble in understanding what have been done in the new configuration. Section 2a provides a general summary of the changes that have been made, but necessary details such as what processes or parameters have been added or changed in the parameterizations are lacking. Also, there is no dedicated case study to investigate the model performance in depth (except a snapshot in Figure 3 and Figure 8). As such, it is very difficult for a reader to appreciate what differences in the simulations can be considered as a real improvement. This is particularly true when considering the presence of new errors in the new configuration for some cloud properties.

1.1.1 Author's response

The aim of the paper is to show how a comprehensive approach to cloud evaluation can be valuable in developing and assessing a new model configuration. In order to achieve this aim we believe we need to show both the multi-diagnostic approach and how it can be used to assess the performance of a new model configuration against a control. However, in the revised manuscript we address the referee's concern that not enough depth of information is given. We have deliberately taken the approach of, wherever possible, using published methods, hence a technical discussion of these methods already exists in the published literature and we believe it would make the paper too cumbersome to repeat it all. The novel aspect here is that we draw the techniques together for the purpose of assessing a new model configuration as part of the model development process. Hence we have provided more detail of the parametrization changes made in the model development process and the relative merits of the different diagnostic approaches (e.g. why one observational dataset might be chosen over another to look at a particular aspect of the cloud simulation). We feel that this discussion of the diagnostic approach is best placed within the results sections to highlight the point that the chosen approach will vary depending on the particular characteristic being examined.

1.1.2 Manuscript changes

Section 2a has been considerably expanded with a more thorough description of the relevant parametrization changes. Within section 3, where possible, the text attributing changes in the errors to particular parametrization changes has been expanded (e.g. around the warm rain microphysics discussion) to discuss how the parametrization differences lead to the improvement and the physical processes operating. The results of two new simulations have been added to Figure 2 in order to clearly attribute the differences seen to particular parametrization changes.

The description of the observational datasets and, where relevant to the paper, their uncertainties has been expanded in Section 2b and in the results sections. In a number of places, we have enhanced the discussion of the value of the multi-diagnostic approach and the increased process-orientated understanding it can provide (e.g. around the mid-latitude cyclone RSW error and cloud errors over mid-latitude land).

1.2 Referee Comments

My second concern is on the comparison of model simulations against satellite observations (e.g. Figure 7, 9, 10, and similarly supplementary Figure 2-4). Many differences are discussed; however, these is no discussion on their statistical significance. How much of the difference is due to sample errors and how much is due to systematic errors in the model? In my view, a significance test should be applied to the analysis to insure that the differences discussed are meaningful. To do this you can use something simple such as a t-test or more appropriately a Monte Carlo method as applied in Booth et al. (2013).

1.2.1 Author's response

We have now conducted a t-test based on the inter-annual variability of the observations and the models for the figures the referee refers to (and several others where this could be considered an issue). As expected, all the systematic errors discussed in the paper are considerably larger than the inter-annual variability and so remain significant.

1.2.2 Manuscript changes

Figures 2, 4, 5 and 11 have been reprocessed with shading around the line plots to represent 5% significance. For Figures 7, 9 & 10, the region of $<$5% significance has been coloured white so that all coloured regions in these plots show statistically significant differences. The significance test is also referred to in section 2b.

2 Specific comments from referee 1

2.1 Comment

Line 62-63: that's fine, but you also have spent a lot of time on inter-comparison of the two configurations of the model.

2.1.1 Response & manuscript change

The purpose of the paper is to show how a comprehensive approach to cloud evaluation can be valuable in developing and assessing a new model configuration. The sentence has been altered in the revised manuscript to indicate this.

2.2 Comment

Line 66: “high”, “mid”, and “low” clouds need to be defined.

2.2.1 Response & manuscript change

Definitions have been added to the manuscript as low:>680hPa, mid:440hPa–680hPa, high:<440hPa.

2.3 Comment

Line 73: please define “NWP”.

2.3.1 Response & manuscript change

This was already defined on line 24.

2.4 Comment

Line 97-117: a summary of the changes is good, but what changes have actually been made? What processes or parameters have been added or modified in the parameterizations? For example, what has been changed in the auto-conversion scheme (line 101)? What does the change do in the new aerosol scheme (line 112)? What does the turbulent scheme do to the production of liquid water (line 110)? You have provided the references, but some necessary details would be appreciated by the readers and would help justify your argument of the model improvement.

2.4.1 Response & manuscript change

This section of the paper is considerably expanded in the revised manuscript with a more detailed explanation of the parametrization changes as the referee requests. It should also be noted that it is intended that the present paper will be included within a GMD special issue which will also include the GA7 model description paper containing a full documentation of all the parametrization changes.

2.5 Comment

Line 145 and 147: CloudSat and CALIPSO provide a “curtain view” of the clouds, which are not really 3-D.

2.5.1 Response & manuscript change

‘3D structure’ has been replaced with ‘hydrometeor profile’

2.6 Comment

Line 180: so how many years are used exactly?

2.6.1 Response & manuscript change

The following has been added to the revised manuscript “25 years for ISCCP, 12 years for CERES-EBAF and 5 years for CloudSat/CALIPSO”

2.7 Comment

Line 194: but you said “3-D” before

2.7.1 Response & manuscript change

The ‘3D’ on line 146 (now removed) referred to CloudSat. Here we are stating that CALIPSO provides the best 2D (latitude/longitude) estimate of total cloud fraction; this doesn’t preclude it from having useful information in the vertical as well.

2.8 Comment

Line 212: “this corrected in GA7” should be “this is corrected in GA7”.

2.8.1 Response & manuscript change

Revised manuscript has been corrected as reviewer suggests.

2.9 Comment

Line 219: it does appear to be the case in GA6 to me. Please clarify.

2.9.1 Response & manuscript change

Referring to the top left panel of Figure 2b we can see no evidence that the altitude of the cirrus with lower backscattering ratios (3-5) is any higher than the thicker cirrus (backscattering ratios 7-20) - if anything the reverse is true. This is unlike the panels for

Printer-friendly version

Discussion paper

GA7 and CALIPSO which show the cirrus with the lowest backscattering ratios to be higher. We really can't see how we can make this clearer and request that the referee looks again at the text and figure.

2.10 Comment

Line 224: in this case I see the model produces a lot of mid-top clouds (which seem to have moderate optical depth) whereas you argue earlier (line 191) that the model simulates too little of this type of cloud?

2.10.1 Response & manuscript change

The hydrometeor signal observed is likely to be the spurious large scale precipitation referred to in the discussion of Figure2c and result from thin large scale cloud which has formed in the moist air around the convective system. As they occur under the anvil of a deep-convective system they won't be seen by ISCCP, and most of them may not be seen by CALIPSO either due to full attenuation from the ice cloud above. In contrast, the mid-top cloud which is 'missing' should be visible to CALIPSO (almost certainly it is missing congestus-type cloud). An extra paragraph has been added to the manuscript discussing these points.

2.11 Comment

Line 230: how "cloud fraction" is defined in the simulation and in the observational data set, respectively? Is a direct comparison meaningful? Please clarify.

2.11.1 Response & manuscript change

The radar–lidar product has considerably higher along track resolution (nominally 1.7km) than the model (80km at the equator), hence regridding the combined radar-lidar data onto the model grid gives an observed cloud fraction to a precision of about 2%. The main assumption here is that the along-track cloud fraction is representative of the 2D grid box. Whilst this is a fair assumption when considering a large number of cases which the A-train will cross at random orientations, we acknowledge that there may be an error when considering a single case such as this. However it's unlikely to affect the key model errors discussed in the paper regarding the figure. These points and caveats have been expanded upon in the revised manuscript as the reviewer suggests.

2.12 Comment

Line 238-239: a lot of these "drizzling" clouds in the simulations have a reflectivity below -20 DBZ, which is very, very weak. It seems odd that these clouds are not picked up by the CALIPSO simulator at all.

2.12.1 Response & manuscript change

As we note in the paper, the rates are $<$0.005mm/hr which is consistent with the very weak signal. The concrete evidence given in the paper is that if large scale rain is not passed to the CloudSat simulator then the signal is removed. As we are below a thick anvil, the CALIPSO simulator signal is likely to have been attenuated and so not see cloud if present. However in GA7, once the spurious precipitation is removed, there is still a cloud signal in the CloudSat simulator just below the threshold of -40dBZ which suggests that the cloud is very thin.

2.13 Comment

Line 257-258 and relevant texts throughout the paper: care should be taken when drawing this conclusion. Previous studies (e.g. Chepfer et al. 2013) have shown that, due to the averaging issue, differences in the zonal cloud fraction retrieved in different CALIPSO products can be quite large (up to a factor of 2 for some regions). It is not unlikely that the GOCCP may have underestimated the cirrus extent.

2.13.1 Response & manuscript change

The reviewer is correct, and GOCCP probably underestimates the amount of cirrus. Chepfer at al. (2013) show that the averaging effect is sensitive to the length of the averaging and is higher for low-level, small-scale broken cloud. For high clouds, the differences between GOCCP and the CALIPSO cloud retrieval used by RL-GEOPROF are dominated by the SR detection threshold. The height-dependent SR detection threshold used in this study increases the sensitivity to high clouds (supplementary Figure 1). For cirrus clouds in the regions shown in Figure 2, the bias introduced by lack of averaging smaller than 0.05 (Figure 10 in Chepfer at al., 2013). This supports the interpretation that GA6, and to a lesser extent GA7, overestimates cirrus. This discussion has been added to the manuscript.

2.14 Comment

Line 298: this is a fairly big box. While I understand that this is a standard method used in previous studies, I am not convinced that it is appropriate for high-latitude regions, where cyclones (e.g. polar vortex) are generally smaller in size and the distance between individual cyclones can be a lot smaller (compared to mid-latitude cyclones).

2.14.1 Response & manuscript change

Throughout the paper we have tried to use published methodologies. As the referee notes, this is a standard size used in other studies looking at similar latitude bands. We acknowledge that cyclones come in a range of sizes and that this technique will just combine them all and hence smooth the signal (this point has been added to the revised manuscript), however we have tested different box sizes and the results are qualitatively similar and conclusions unchanged.

2.15 Comment

Line 331-322: This is a complicated case, with multiple fronts being diagnosed. Therefore it is hard for me to associate the cloud features discussed in the paper to the synoptic components shown on the MSLP chart. Further information such as latitude and longitude on the discussed cloud fields should help.

2.15.1 Response & manuscript change

Latitude references have been added to the revised manuscript as the reviewer suggests.

2.16 Comment

Line 332-334: I don't understand this sentence.

2.16.1 Response & manuscript change

This and adjacent sentences have been re-written to explain the point more clearly and a reference added as the result is consistent with previous experience.

2.17 Comment

Line 364 and relevant text later in the paper: I disagree. What I have seen is that the large RSW bias is present in some of the cold air side of the cyclone, but almost everywhere in the poleward side of the cyclone! Why? This is, to me, an important issue but no discussion has been made in the paper (or the referenced study). There is a lot of focus on the cold air side of the cyclone, but this is only part of the story revealed by the plot. Also, the bias on the cold air side of the cyclone does not explain the poleward increase of the radiative bias shown in Figure 13.

2.17.1 Response & manuscript change

We don't really understand the referee's distinction between the "cold air" and "poleward" side of the cyclone. On average, the poleward side of the cyclone is a subset of the cold air which also extends around the western side of the cyclone. By "cold air" we are referring to all of the region from the poleward side of the cloud head associated with the typical warm front position of about 3 o'clock on the southern hemisphere composite in figure 9, around the poleward and westward side to the poleward edge of the typical cold front position at 10-11 o'clock. In the revised manuscript we have clarified that cold air includes the poleward side of the cyclone.

2.18 Comment

Line 368: I don't think Figure 10 is necessary. It does not seem to provide any substantially different information than Figure 9.

2.18.1 Response & manuscript change

The key difference is that RSW in Figure 9 will depend on the insolation, hence for the same cloud albedo error, the RSW error will be larger in the summer than winter. Figure 10 shows the in-cloud albedos which do not depend on the insolation (this point has been explicitly added to the revised manuscript). Hence Figure 10 shows the interesting point that the in-cloud albedos are lower in the austral summer compared with the winter which, combined with the stronger insolation, leads to the larger negative RSW bias in Figure 9.

2.19 Comment

Line 369-370: why the cloud amount errors are not large enough to contribute significantly to the SW errors? Please explicate.

2.19.1 Response & manuscript change

We accept that this statement was probably too strong and requires correction which we have done in the revised manuscript. The cloud fraction errors for GA7 in Figure 7 and the supplementary material do not have the same spatial pattern (and are sometimes of the wrong sign) to explain the RSW error.

2.20 Comment

Line 373-374: again, the errors seem to be prevalent in the poleward side of the cyclone, too (which is also the case in the referenced study).

2.20.1 Response & manuscript change

See above response regarding “poleward” and “cold air”.

2.21 Comment

Line 411: why this appears to be an issue for the UM? Please explicate.

2.21.1 Response & manuscript change

Sentence has been revised in the manuscript to “This appears to be an issue for the UM in parts of the tropics where too little shallow cumulus is simulated and typically the model has cloud fractions of $<$3 oktas whereas fractions over this threshold are often observed and hence a cloud base height assigned”.

2.22 Comment

Line425-426: could some of these excess low clouds actually be precipitation not being detected by the instrument?

2.22.1 Response & manuscript change

Here we are comparing the model cloud (no simulator involved) with SYNOP observations so the presence, or not, of precipitation should be irrelevant.

2.23 Comment

Line 447-448: but now there seems to be too much red (for RSW) in the sub-tropics which was non-existent in HadGEM2-A?

2.23.1 Response & manuscript change

The sentence has be revised to read “The error in the sub-tropical cumulus transition regions of excess RSW has been removed and there is now a smaller negative bias in GA7”. We have also reproduced Figure 13 with a revised colour bar to make it clearer that the negative bias in GA7 is smaller in magnitude than the positive bias in HadGEM2-A.

2.24 Comment

Figure 2: (1) you use “equivalent reflectivity factor” in the plot but “reflectivity” in the caption (and the following figures). (2) What do the colour bars show? It should be indicated in the plots. (3) The lowest km should be masked in the CloudSat plot (as done in your following figures).

2.24.1 Response & manuscript change

The term "reflectivity" is now used throughout (including figures). Indication of what the colour bars show has been added to the Figure 2 caption and explanation of how the histograms are formed has been added to section 2. The lowest 1.2km has now been masked for CloudSat (the main instrument to suffer from near-surface issues).

2.25 Comment

Figure 3: What does "CloudSat/CALIPSO" mean in the top left plot while you only show reflectivity?

2.25.1 Response & manuscript change

The caption to the figure explains that "...situations in which the lidar detected cloud but the radar did not being included with a nominal value of -40dBZ (e.g. Mace and Wrenn, 2013)"

2.26 Comment

Figure 5: I don't think including the very cold SST ranges is necessary as they are quite rare and the plots show very similar features (i.e. the first four plots in the bottom panel).

2.26.1 Response & manuscript change

Although relatively rare spatio-temporally, these typically represent the subtropical stratocumulus regions which are widely regarded by the cloud feedback community to be critical for the cloud response to climate change, hence their separation by SST is relevant. In addition, the coldest SST bin shows a significant (i.e. still distinguishable following the significance tests discussed above have been applied) error in GA6 which has been improved in GA7.

2.27 Comment

Figure 8: is 64.32N (the right end of the cross-session plots) over land? I can see the topography-like feature at the surface in the bottom left plot, but why there are clouds produced underneath the surface in the simulations?

2.27.1 Response & manuscript change

Neither cross section is over land. The masked region in the lower-right corner corner of the observed cross-section is due to the reliability mask associated with the RL-GEOPROF product indicating high uncertainty in both CloudSat and CALIPSO for these bins.