Interactive comment on “Ship and satellite observations over the ocean for verification of the shortwave cloud radiative effect in climate models” *by* T. Hanschmann et al.

Reply to the review of Anonymous Referee #2

First of all, we thank anonymous referee #2 for his/her comments on our paper.
We will provide our response to each comment in the following. The comments of the reviewer will be shown in regular letters and our reply will be shown in italic red letters.

1) Abstract: “In this study the accuracy of the radiative transfer scheme of the ECHAM climate model for reproducing the shortwave cloud radiative effect (SWCRE) at the sea surface has been investigated.”
The paper is actually more an evaluation of the different data sources used as input to the radiative transfer scheme than of the radiative transfer scheme itself. For example, the 6 experiments in the closure study differ by whether the calculations are specified using surface observations or satellite observations, and there is a lot of emphasis on the value of including satellite retrieved cloud effective radius in the calculations. There does not appear to be any actual evaluation of the accuracy of the radiative transfer scheme itself. Please clarify.

We agree with this comment and have re-written the corresponding part of the abstract accordingly. The abstract now emphasizes more clearly the goal of the study, which is the evaluation of the model input data by means of radiative flux comparisons. Furthermore, following the comment above we changed the title of the paper to: “Evaluation of the shortwave cloud radiative effect over the ocean by use of ship and satellite observations.” This captures the actual scope of the paper better and clarifies the misleading formulation on evaluating the GCM-RTM.

2) p.3 (line 28): “Average differences between measured and modeled DSR of 35.7% were found for single layer water clouds using the Rapid Radiative Transfer Model for GCMs ...”
Does the 35.7% really correspond to bias error or an RMS error? It seems too large to be a bias error. Was the error due to sampling errors in mountainous terrain? Also, are 3 significant digits necessary (why not 36%)?

The error of 35.7% reported in the paper of Ebell et al. (2011) refers to the average difference normalized by the observed fluxes, also called the relative bias, between calculated minus observed fluxes for all single layer water clouds observed during the whole COPS campaign (9 months). This value does correspond to the bias because the differences between modeled and measured fluxes are not squared. However, to be more comparable to our data the bias itself might be the more appropriate quantity. Ebell et al. (2011) found -39.1W/m^{-2} for the bias. The RMS difference in their comparison was 143.1Wm^{-2}. If we consider their category of multi-layer water clouds, as it also was observed during our cases, the bias in Ebell et al.(2011) is even larger (-56.4 Wm^{-2}). Ebell et al. (2011) explained their findings for the bias by broken cloud conditions, which do not match the plane parallel assumption of the model. We took the original numbers from the paper by Ebell et al. (2011) and would not like to average them to fewer-digit accuracy.

Reference:
Ebell, K., Crewell, S., Loehnert, U., Turner, D. D., and O’Connor, E. J.: Cloud statistics and cloud radiative effect for a low-mountain site, Q. J. Roy. Meteor. Soc., 137, 306–324, doi:10.1002/qj.748, 2011. 17745

3) p. 4 line 7: remove the word “nowadays”.

The evaluation of the representation of cloud properties over the ocean in GCMs does not necessarily refer to recent GCMs. The word has been removed.

4) p. 4 lines 7-11: The representation of Cloud radiative effects in GCMs has been evaluated far more often using ERBE or CERES satellite data, not ISCCP.

Our statement refers to the usage of satellite-based cloud properties as input to GCMs and their further comparison with measurements. Neither ERBE nor CERES provide cloud property data. However, they do provide radiative fluxes to validate model data. To clarify this statement we changed the text to:

"Until now, the assessment of model-based cloud properties over the ocean (e.g. as estimated by ISCCP; Rossow and Schiffer, 1991) and their associated cloud radiative effect (e.g. as observed by ERBE (Barkstrom, 1984) & CERES (Wielicki et al., 1995)) have mostly been based on satellite data sets, or on data of voluntary observing ships(e.g. Bedacht et al., 2007)."

5) p. 4 line 13: Start a new paragraph with "The aim of this study…". Also, as noted above (see comment about Abstract), the study doesn't actually evaluate the ECHAM-5 radiative transfer scheme, but rather evaluates the inputs to the radiative transfer scheme in the context of how accurately cloud radiative effects at the surface can be calculated. Please revise.

In agreement with this comment, we have rewritten this part of the manuscript. We emphasise the evaluation of the best-matching cloud properties for representing the SWCRE modelled with ECHAM-5 RTM in a new paragraph.

5) p. 10 line 1: What does "PS" stand for?

PS stands for POLARSTERN, which is the research vessel where the ground-based measurements have been performed. An explanation has been added to the manuscript.

5) p. 11 lines 3-7: It is not necessary to list this detail in the main text. Please use figure caption for that.

These experiments are the central part of our study. We maintain that the description of the experiments is essential for the understanding of the closure study and, hence, we would like to keep this in the main text.

7) p. 11 line 16: How can a positive SW CRE occur? It implies the cloudy flux at the surface exceeds the clear flux at the surface. Could it be that the clear-sky flux is artificially low due to shadowing by adjacent clouds? Please explain.

A positive shortwave cloud radiative effect occurs when the cloudy flux exceeds the clear sky flux. This happens frequently during broken cloud conditions. An alternative explanation could be an underestimation of the reference clear-sky irradiance. In our work, we used the clear-sky irradiance parameterization of Kalisch and Macke (2008). This parameterization underestimates the clear sky flux only for very dry atmospheres, but by no more than 50 Wm^{-2} at noon. Furthermore, sky images from the measurement campaigns confirm that broken cloud conditions caused the excess radiation.

Reference:

Kalisch, J. and Macke, A.: Estimation of the total cloud cover with high temporal resolution and parameterization of short-term fluctuations of sea surface insolation, Meteorol. Z., 17, 603–611, 2008. 17745, 17747, 17755, 17768

8) p. 12 lines 12-15: These sentences ("From the observed clear sky…error of the measurement devices") repeat what has already been stated in lines 9 and 10. Please revise.

We agree that this statement is repetitive; therefore we have rewritten the text as follows:
"For the clear sky case a slightly larger DSR is observed, compared to the modeled clear sky value. This deviation is most likely the result of two competing effects: the shadowing of the forward scattered solar radiation and the reflection of solar radiation by the ship's superstructure. The observed diurnal cycle of the clear sky flux (not shown here) indicates that the reflection effect is slightly stronger than the shadowing effect. However, the difference between the two artifacts is less than 3 Wm^{-2}, which is well within the accuracy of the pyranometer. Different aerosol loads can also be the source of the deviation. The small deviation for the clear sky case shows that in the absence of clouds the RTM can well reproduce the diurnal cycle of DSR at the surface."

9) p. 12 line 18: Please define what constitutes "good" and "bad".

We agree that qualifying the results as "good" or "bad" would require an objective measure. Our results do not clearly answer which experiment performs best. Consequently, our results will be compared with those from former studies. We will modify this in the revised manuscript and will compare the results with other independent data to verify the accuracy of our findings. First, we compute the mean bias with its standard deviation and the mean absolute deviation of the bias for each experiment over the seven different cases. These values are shown in table 1 and will be added to table 3 in the *manuscript. We add them to help verify and compare our results. Compared to e.g. Ebell et al. (2011), the mean absolute deviation of the bias in observed and modeled fluxes is mostly lower in our comparison. Especially for the experiments using the observed effective radius, we find a clear reduction in mean absolute deviation. In a further comparison we used the CM-SAF specifications, which provide a target accuracy for the surface incoming shortwave radiation product of 20 Wm^{-2} for daily mean (see the CDOP Product Requirements Document, CM-SAF (2011)). If we compare our results to their target accuracies we find a larger mean absolute deviation of the bias.*

Reference:
CMSAF Technical Team, CDOP Product Requirements Document, SAF on climate monitoring, Doc. No.: SAF/CM/DWD/PRD/1, Issue: 1.8, Date: 22. November 2011

10) p. 12 lines 16-19: "Summarizing..." This statement is incorrect. See results for Exp PS-RSAT and PS-R10 in Table 3. Differences are positive for all conditions except the clear case.

"Summarizing the results in Table 3, no systematic differences are found between using ship-based or satellite-based cloud properties for describing the observed shortwave cloud radiative effect." This comment refers to the different signs some experiments show for different cases. However, there are different results for the different experiments depending on the considered case. As the reviewer mentioned, some experiments show only overestimation of the modeled DSR and others show mostly underestimation. Because we compare daily mean values the reasons for under-/overestimations are diverse. From the results shown in table 3 we cannot detect a systematic difference between the first three and the second three experiments, although we found an overestimation in the first three experiments and an underestimation in the last three experiments in the mean experiment bias. Hereby, we note that for all experiments the standard deviation exceeded the mean. In short, we found large absolute differences for both cases. However, we did observe a decreasing mean absolute deviation in bias if the satellite based effective radius was included in the experiments.

11) p. 27 Table 3: It would be helpful to have some sense of the solar zenith angle range. That is, what constitutes "low sun" and "high sun"?

According to this comment we have changed the titles of these two cases to their geographical location. High sun is related to tropical conditions and low sun to midlatitude conditions.

Table 1: The table shows the mean bias (with standard deviation) and the mean absolute deviation of the bias for each experiment over all considered cases in table 3 of the manuscript. These two lines should be added to table 3 in the manuscript.

	EXPERIMENT PS	*EXPERIMENT PS-RSAT*	*EXPERIMENT PS-R10*	*EXPERIMENT CMSAF*	*EXPERIMENT CMSAF-NORSAT*	*EXPERIMENT CMSAF-R10*
mean bias	*1.54 ± 42.78*	*28.43 ± 29.81*	*31.90 ± 38.14*	*-6.47 ± 37.67*	*-32.14 ± 58.51*	*-7.11 ± 50.44*
mean absolute deviation of the bias	*31.68*	*29.14*	*32.61*	*26.24*	*50.09*	*39.77*