

Response to referee #2 for manuscript:

X. Ceamanos, D. Carrer, and J.-L. Roujean, Improved retrieval of direct and diffuse downwelling surface shortwave flux in cloudless atmosphere using dynamic estimates of aerosol content and type: application to the LSA-SAF project, Atmos. Chem. Phys. Discuss., 14, 8333-8392, doi:10.5194/acpd-14-8333-2014, 2014

Dear Editor,

Please find below the response to all comments made by referee #2, Prof. Lucien Wald. Prof. Wald is one of the authors of the method McClear, which has been used in our study to evaluate the performances of the proposed algorithm SIRAMix. While referee #2 seems to agree on the main innovation of SIRAMix - the consideration of several aerosol types and its benefits in estimating DSSF -, he shows some concern on the comparison between both radiation retrieval methods. More specifically, referee #2 underlines the differences between the scores obtained for McClear in our manuscript and those reported in the McClear's article (Lefèvre et al., ATM, 2013). In our response below, we thoroughly address these concerns (see P3-P7) by clarifying certain aspects of the reported experiments in order to prove their rightness. Likewise, we propose to add some additional information in the revised version of the manuscript to clarify this matter to the potential readers of our article.

Referee #2 also suggests to perform an additional evaluation of SIRAMix based upon the clearness index (KT) (see P2 and P5 below), as it was done for McClear in (Lefèvre et al., 2013). Nonetheless, we have decided not to do such an evaluation due to the following non-exhaustive reasons: (i) the main goal of SIRAMix is to estimate DSSF (and not KT) to be used as forcing in climate and weather forecast models, (ii) given that KT is directly proportional to DSSF, an additional evaluation based on the former would not lead to different conclusions than those drawn from the DSSF-based evaluation included in our manuscript, and (iii) our article includes a series of comprehensive experiments that validate SIRAMix based upon estimated DSSF evaluation, similar to what is done in most alike papers in the literature.

We would like to thank here the work done by the reviewers and the editor, which has proved to be very helpful in improving the current manuscript.

Best regards,

The authors

Interactive comment on “Improved retrieval of direct and diffuse downwelling surface shortwave flux in cloudless atmosphere using dynamic estimates of aerosol content and type: application to the LSA-SAF project” by X. Ceamanos et al.

L. Wald (Referee)

lucien.wald@mines-paristech.fr

Received and published: 8 April 2014

The concept of SIRAMix is sound. This is an innovation that may be likely adopted in several operations. As underlined by the authors, the diffuse DSSF is strongly dependent on the type of aerosols. Having more than one type produces more accurate results on the diffuse. This is demonstrated in the comparison between SIRAMix and McClear diffuse DSSF.

Dear Prof. Wald,

Thank you very much for your sensible suggestions and comments. Please find our response to all your concerns below.

P1 - The authors state that the difference in diffuse is only due to aerosol types. I'd like to see a discussion why ruling out the possible influence of the ground albedo.

In fact, diffuse DSSF depends (among many parameters) on surface albedo, which determines the multiple scattering between the surface and the lower layers of the atmosphere. The use of different surface albedo products in SIRAMix (i.e., SERIVI-derived) and McClear (i.e., MODIS-derived) may contribute in obtaining distinct estimates of diffuse DSSF. In fact, both satellite-derived products present some small differences despite having proved to be of high quality (Geiger et al., 2008; Schaaf et al., 2011; Carrer et al., 2010). These differences could be one of the reasons for the improved diffuse DSSF provided by SIRAMix. Unfortunately, this thesis is difficult to prove due to the lack of reference measurements of surface albedo for the ground stations considered in the experiments. Furthermore, this hypothesis would mean that the LSA-SAF albedo is generally more accurate than MODIS's. This fact has not been observed in the literature. This discussion will be included in the revised version of the manuscript.

To further investigate this point, we propose to illustrate the dependency of diffuse DSSF on surface albedo. This is done by including an additional sensitivity study in the new version of the manuscript. We have considered the surface albedo as a free parameter in Experiments 1 and 2 (see sections 4.1.1 and 4.1.2), similar to what is done with other physical quantities. Accordingly, we suggest to replace Figures 4 and 5 by Figures A and B (see below), respectively. As it can be seen, Fig. A underlines the robustness of the DSSF estimated by SIRAMix with regard to simulations obtained by the radiative transfer code libRadtran (DSSF error lower than 1%) for a large range of surface albedos (from 0 to 0.4). On the other hand, Fig. B illustrates the significant impact of using a biased surface albedo to calculate diffuse DSSF with SIRAMix (up to 4%).

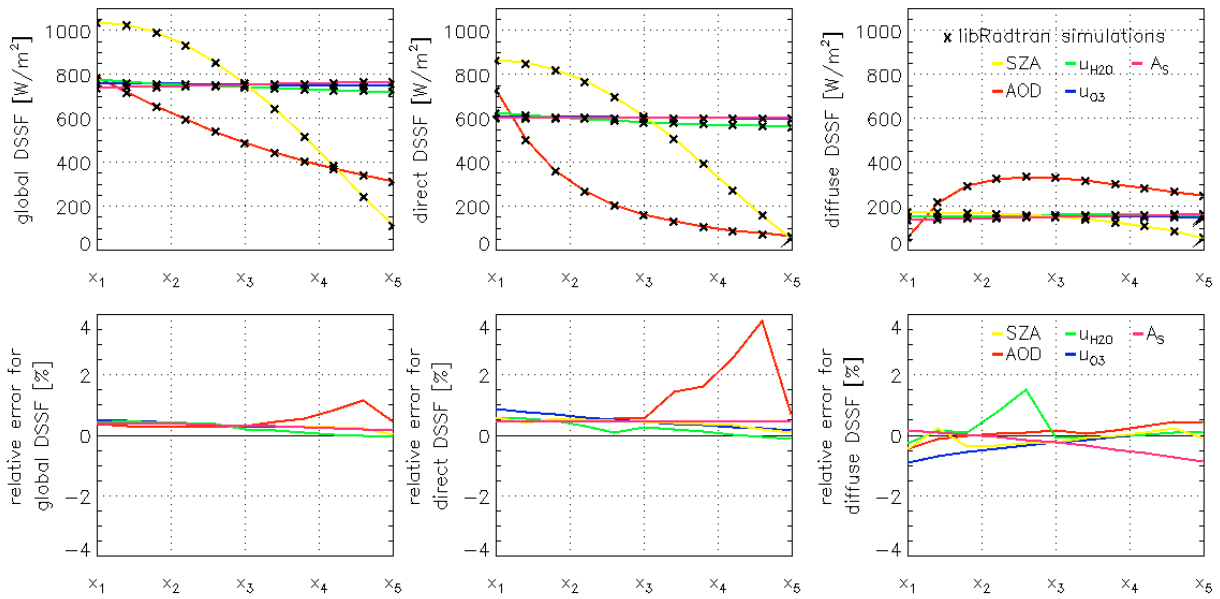


Figure A - Improved version of Figure 4 after including the surface albedo (A_s) as a free parameter - (top): Global, direct, and diffuse DSSF values calculated with SIRAMix according to varying AOD (red color), ozone content (blue color), water vapor concentration (green color), SZA (yellow color), and surface albedo (purple color). Coincident DSSF simulations with libRadtran are shown with black crosses. (bottom): Relative error for global, direct, and diffuse DSSF values when compared to libRadtran simulations. Horizontal axis ticks (x_1, x_2, x_3, x_4, x_5) correspond to (0,1,2,3,4) for AOD, to (100, 200, 300, 400, 500) in Dobson units for ozone, to (1.0, 2.0, 3.0, 4.0, 5.0) in g/cm^2 for water vapor, to (0, 20, 40, 60, 80) in degrees for SZA, and to (0.0, 0.1, 0.2, 0.3, 0.4) for surface albedo.

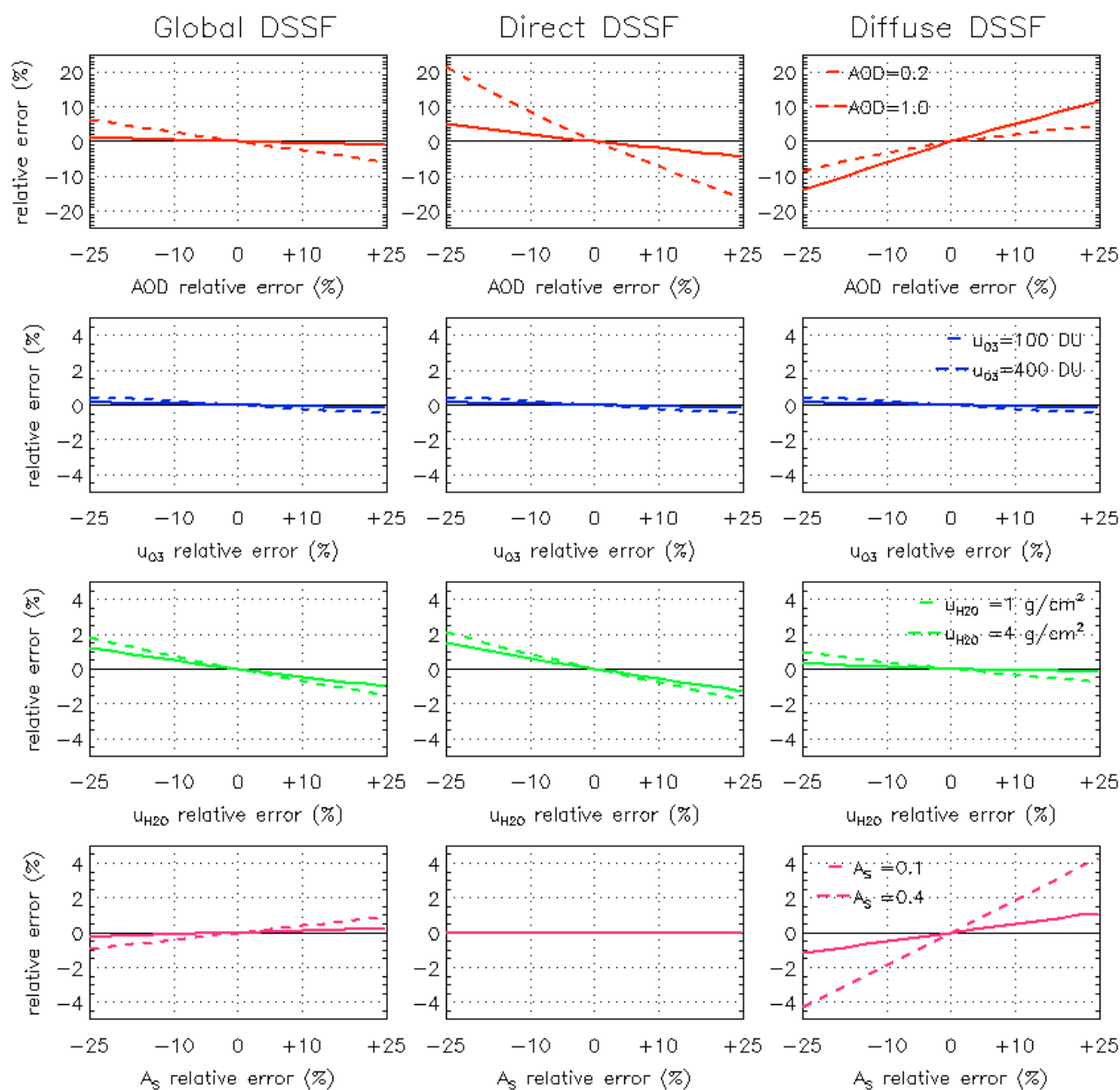


Figure B – Improved version of Figure 5 after including the surface albedo (A_s) as a free parameter - Relative error on global, direct, and diffuse DSSF calculated with SIRAMix caused by uncertainties in terms of relative error affecting AOD (1st line figures. Note different vertical scale), ozone content (2nd line figures), water vapor concentration (3rd line figures), and surface albedo (4th line figures). Two cases corresponding to different values of the parameter under study are considered (see plain and dashed lines). Note the different vertical scale for each input under study.

P2 - The clearness index is a valuable means to better understand the performance of a model. The clearness index, also known as the atmospheric transmissivity, is the ratio of the DSSF (global, direct or diffuse) to the similar quantity but at the top of the atmosphere. The change in solar radiation at the top of atmosphere due to changes in geometry, namely the daily course of the sun and seasonal effects, are usually well reproduced by models and lead to a de facto correlation between observations and estimates of DSSF hiding potential weaknesses. Clearness indices are stricter indicators of the performances of a model regarding its ability to estimate the optical state of the atmosphere. Though the clear sky indices are not completely independent of the solar zenith angle as they decrease as this angle increases, the dependency is much less pronounced than in DSSF. Accordingly, I'd like to see tables with clearness indices like Tables 4, 5 and 5 for DSSF as it will help to better identify the potentials of SIRAMix.

We thank referee #2 for his suggestion. However, we prefer not to include an additional study based on KT based on the following reasons:

- The main goal of SIRAMix is the estimation of global, direct, and diffuse DSSF. These physical quantities are used as forcing in most of surface, atmospheric and weather forecast models (e.g., Szypta et al., 2012, Carrer et al., 2012, Quintana et al., 2010; Habets et al., 2008). The information provided in our article on the DSSF uncertainties will be essential to determine the impact of using our DSSF product on these models. As far as we know, KT is seldom used in these climate models.
- Most of articles proposing or assessing methods for DSSF retrieval are exclusively based on DSSF evaluation (e.g., Geiger et al., 2008; Mueller et al., 2009; Deneke et al., 2005; Gueymard et al., 2003; Liang et al., 2013; Moreno et al., 2013; Psiloglou et al., 2007; Yang et al., 2006; Yoshida et al., 2013).
- The clearness index is directly proportional to the DSSF (i.e., $KT = DSSF / E_0$, where E_0 is the downwelling shortwave flux at the top of the atmosphere). This is the reason why Lefèvre et al., (2013) show that scores obtained by the method McClear are almost the same when evaluating DSSF or KT. For example, bias and RMSE scores in % of estimated global and direct DSSF - see Tables 2 and 3 from Lefèvre et al., (2013) - are almost identical to the same quantities for estimated KT - see Tables 4 and 5 in Lefèvre et al., (2013) -. The only notable difference among scores happens for the correlation coefficient. However, Lefèvre et al., (2013) state that “*The squared correlation coefficients for the clearness index KT (...) should be discussed with care as the limited range of values tends to decrease the correlation coefficient without denoting poor performance (...)*”, indicating that correlation coefficients may not be reliable.
- The reference data used for validation in our article are ground measurements of DSSF. Reference values of KT are not directly available and they should be obtained dividing the reference DSSF measurements by estimates of E_0 . We think that this indirect validation would be less reliable than validating SIRAMix based upon directly-obtained DSSF.
- We remember that we have performed an exhaustive evaluation of SIRAMix based on radiative transfer simulations, ground-based measurements, and other DSSF retrieval methods. This has resulted in more than ten pages, three tables and six figures.

In conclusion, we believe that the experiments that we have carried out to assess the performances of SIRAMix are already very comprehensive in the current form of the manuscript. In our opinion, an additional study based on KT would not lead to different conclusions and thus would result redundant.

P3 - Clear-sky detection is not discussed at all. This should be performed as it is particularly important when comparing to a clear-sky model like McClear. The algorithm used should be described.

We agree that the determination of clear sky conditions is fundamental in our study. In this matter, we would like to remind that SIRAMix is also a clear-sky model that exclusively works for cloudless situations (i.e., atmosphere composed of aerosols and gases only). However, SIRAMix and McClear use different strategies to determine clear sky instants. Please find below the details of these two methods:

1. McClear uses two filters to retain reliable clear-sky instants. The first filter rules out all DSSF estimates whose matching ground measurements do not satisfy the condition $E_{diff}/E_{glo} < 0.3$ (being E_x the global or diffuse DSSF). The second filter retains only periods with enough measurements that have passed the first filter in order to avoid the presence of broken clouds.
2. SIRAMix selects clear-sky instants based upon the cloud mask provided by the SAF-NWC (<http://www.nwcsaf.org/HD/MainNS.jsp>). This product is derived every 15 minutes from SEVIRI/MSG infrared observations using the method from Derrien and Le Gléau, (2005) and has proved to be highly accurate in many studies (Carrer et al., 2010a and 2012). Only instants of time flagged as “cloud free” in the SAF-NWC cloud mask are suitable to be processed by SIRAMix. For extra precaution, the cloud mask is “dilated” in time, ruling out any “cloud free” instant of time if the previous (-15 min) or next (+15 min) time slots are not also flagged as “cloud free”. This second step is mainly meant to avoid broken clouds.

The strategy used in our article (number 2 above) is explained in Section 2.3.1 of the original manuscript except for the mask dilation. This additional step will be fully detailed in the revised version of the manuscript.

P4 - I suspect that the clear sky cases may not be all clear and this may prevent an accurate comparison between SIRAMix and McClear as the latter is only for clear sky conditions. Several stations (Carpentras, Sede Boqer, Tamanrasset) are appearing in this paper as well as in the paper from Lefevre et al. (2013) describing McClear. Periods are not similar but one may expect to find similar figures for McClear in both papers. It happens that the mean values (called "averages" in Tables 4, and 6) in direct and global McClear DSSF presented in Lefevre et al. are significantly greater than those given in this paper. For direct DSSF, Carpentras: 465 (this paper) versus 505 W/m² (Lefevre et al.); Sede Boqer: 527 versus 667 W/m²; Tamanrasset: 511 versus 653 W/m². This may be explained if the data set in this paper contains not-so-clear cases.

Thanks for your comment. Indeed, different average DSSF values are observed for some common stations in our manuscript and in (Lefèvre et al., 2013). In our opinion, the main reason for such differences is not the faulty selection of clear-sky instants in our manuscript but the different filtering of non-clear-sky instants carried out in the two studies.

For the experiments detailed in our manuscript, we followed the strategy based on the SAF-NWC cloud mask (see method 2 in answer to P3). This applies to all DSSF data sets, that is, SIRAMix, the current LSA-SAF algorithm, the ground measurements, and also McClear's. The latter filtering of non-clear-sky instants is possible since the McClear data that we downloaded from <http://www.soda-pro.com/free-web-services/radiation/mcclear> are available for all instants of time independently of cloud coverage. The data set is described by the authors as "A time-series of solar radiation that should be received on a horizontal plane at ground level if the sky were clear", meaning that the McClear method has been run for instants of time flagged as "clear" and "cloudy". The filtering of McClear data in our experiments following the SIRAMix strategy results in a data set of DSSF values that is not the same than the one that would have resulted by following the clear sky detection detailed in (Lefèvre et al., 2013; see method 1 in answer to P3).

We suspect that the average DSSF values obtained for Carpentras, Sede Boqer and Tamanrasset are lower in our manuscript than those in (Lefèvre et al., 2013) because in the latter article many clear sky instants of time that are related to a mild or high AOD were ruled out from experiments. Indeed, the condition defined in (Lefèvre et al., 2013) (i.e., $E_{diff}/E_{glo} < 0.3$, see method 1 in answer to P3) may be too restrictive in some cases. This is especially the case for instants of time with mild/high AOD values or high SZA values, which might be wrongly classified as "cloudy". In order to illustrate this point, we have generated Figure C (see below) showing the daily evolution of the ratio E_{diff}/E_{glo} for different aerosol loads over Carpentras. Simulations were run with the radiative transfer code libRadtran considering a standard cloud-free atmosphere for the summer and winter solstices. As it can be seen, the condition used in (Lefèvre et al., 2013) to detect clear sky instants of time is often too strict. Indeed, many clear sky instants are classified as cloudy, even when AOD is low. This clear-sky selection becomes quite unsuitable especially during the winter, when a clear-sky day with an AOD=0.2 would be wholly discarded (see Fig. C right). It is interesting to remark here that according to ground measurements from AERONET (<http://aeronet.gsfc.nasa.gov>) the average AOD in 2011 (the period of study considered in our manuscript) over Carpentras, Sede Boqer and Tamanrasset was 0.16, 0.20, and 0.32.

In conclusion, the consideration in our article of clear sky instants of time that are mildly or highly polluted by aerosols -and likely ruled out in (Lefèvre et al., 2013)- may explain the lower average DSSF values obtained for the stations referred to by referee #2. Indeed, the higher the AOD, the

lower the radiation that reaches the ground. This reasoning is in agreement with the greater difference in average DSSF values for stations with occurrence of higher AOD values (i.e., Sede Boqer and Tamanrasset) than the less aerosol-affected station in Carpentras.

The issue of comparing different ensembles of clear sky instants of time in our manuscript and in (Lefèvre et al., 2013) will be discussed in the new version of the manuscript.

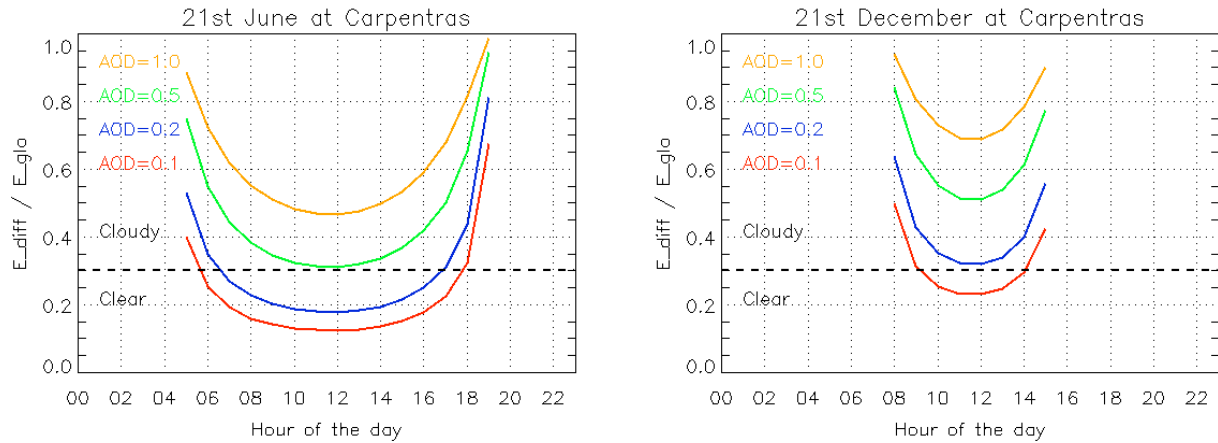


Figure C – Daily evolution of ratio E_{diff} / E_{glo} simulated by libRadtran over Carpentras for (right) the summer solstice and (left) the winter solstice. A U.S. Standard cloudless atmosphere was chosen with four different values of AOD: 0.1 in red, 0.2 in blue, 0.5 in green, and 1.0 in orange. The condition used in (Lefèvre et al., 2013) to detect clear sky moments is plotted with a dashed black line. All time instants over this line would be discarded in the experiments of (Lefèvre et al., 2013).

P5 - This is supported by a further analysis of the McClear RMSE in both papers. If one focus on the direct DSSF, one may expect a decrease in McClear RMSE for 30 min data in this paper compared to the McClear RMSE for 1 min given in Lefevre et al. This is not the case at all and McClear RMSE given in this paper are similar or greater than those in Lefevre et al. An example is Tamanrasset where the RMSE for direct DSSF is 48 W/m2 in Lefevre et al. and 82 W/m2 in this paper. Again, this may be explained by the inclusion of not-so-clear cases in the data set and this casts shadow on the conclusions. Using clearness indices will clearly help in this matter. Maybe the authors should restrict their selection of clear-sky conditions and perform the same comparison with this restricted data set.

Thank you for your comment. First of all, it is important to note that we do not work with 30-min averaged DSSF estimates, as referee #2 seems to suggest. SIRAMix provides instantaneous values of radiation in W/m², similar to the operational LSA-SAF method and the ground stations. Original 1-min averaged DSSF estimates from McClear were converted from Wh/m² to W/m² units. Then, the experiments of our manuscript only consider the instantaneous DSSF values at minutes 00 and 30 for all DSSF data sets. This choice has a historical reason, as the operational LSA-SAF product provides instantaneous DSSF every 30 minutes (corresponding to one SEVIRI/MSG slot out of two). For information, this choice was made according to the needs of the user community who was inquired by the LSA-SAF team. In this context, the RMSE values presented in our manuscript for all methods can be compared in terms of temporal resolution and units. Similarly, the results presented in (Lefèvre et al., 2013) can be directly compared to ours. These particularities regarding the time resolution of the different radiation data sets will be clarified in the new version of the manuscript.

Again, higher RMSE values in our manuscript with regard to (Lefèvre et al., 2013) come from the different selection of clear sky instants (see answer to P4 above). We obtain higher RMSE values for McClear since we consider clear-sky aerosol-polluted instants of time that were likely discarded in the experiments of (Lefèvre et al., 2013). We remember that the estimation of DSSF under mild or high AOD becomes more challenging than under low AOD, thus naturally resulting in higher RMSE values. Note that the RMSE differences are quite significant only in Tamanrasset (for McClear-estimated direct DSSF, 82.3 W/m² in our manuscript and 48 W/m² in Lefèvre et al. 2013) and, to a lesser extent, in Sede Boqer (71.8 W/m² and 62 W/m², respectively). This is in agreement with the occurrence of higher AOD values in these two stations. For Carpentras, however, we obtain similar results than in Lefèvre et al., (2013), thus supporting our thesis (for McClear-estimated direct DSSF, 33.6 W/m² in our manuscript and 35 W/m² m² in Lefèvre et al. 2013). This information will be included in the revised version of the manuscript.

Another reason to explain the differences in terms of RMSE is the fact of considering different periods of time. For example, the average AOD over Tamanrasset according to AERONET records (<http://aeronet.gsfc.nasa.gov/>) was higher in 2011 (AOD=0.32) -the study period of our manuscript-, than in the 2005-2008 period (AOD=0.24) -the period considered in (Lefèvre et al., 2013). The higher aerosol loading in our case may likely result in higher average errors.

Finally, we do not think that a study based upon clearness indexes (KT) would be useful here (see answer to P2).

P6 - In Table 6, the authors discuss mainly the RMSE (p. 8368, line 3). The RMSE is a quadratic combination of the bias and the standard-deviation. Given the low bias compared to McClear, and the fairly similar RMSE between McClear and SIRAMix, this would mean that the standard-deviation is greater for SIRAMix than for McClear. This may be discussed.

We do not agree with referee #2 on this point. First, Table 6 shows that the bias averaged for all stations is not lower but quite similar, that is, 7.4 W/m² for SIRAMix and 6.8 W/m² for McClear. We also find this equality by looking at results station by station: average bias for SIRAMix in comparison with McClear is greater (meaning a larger absolute bias) for three stations (Granada, Sede Boqer, and Tamanrasset), very similar for four stations (Burjassot, Evora, Palma de Mallorca, and Toravere), and lower for two stations (Cabauw and Carpentras). As for RMSE, it averages 23.6 W/m² for SIRAMix and 26.5 W/m² for McClear. Given the equality in terms of bias and the lower RMSE of SIRAMix, the standard deviation is also somewhat lower for SIRAMix than for McClear. Although these scores have not been showed in the manuscript to avoid redundancy, a comment on this matter will be included in the revised version.

P7 - In addition, I'd like to see a discussion on the correlation coefficient for the diffuse component (Table 6) which is greater for McClear than SIRAMix, except Toravere. It is possibly linked to the greater standard-deviation.

We agree that Table 5 shows a slightly lower correlation coefficient of SIRAMix for some stations (6 out of 8) with regard to McClear. However, this is not explained by the standard deviation, as it is lower for SIRAMix than for McClear (n.b., RMSE and bias are lower for SIRAMix, so is the standard deviation). One possible reason might be the use in SIRAMix of the individual AOD

values from MACC-II for each aerosol species. While this information helps to reduce the average bias and the RMSE of the diffuse DSSF, it seems to introduce some uncertainty in terms of the temporal coherence of the AOD series. This may result in lower correlation coefficients for SIRAMix, but not for McClear, which only uses the total AOD from MACC-II. However, this is difficult to prove as we do not have reference measurements of individual type-related AOD estimates.

In any case, results for diffuse DSSF in terms of squared correlation coefficient must be interpreted with precaution. According to (Lefèvre et al., 2013), "*a limited range of values (such as the low values of diffuse DSSF) tends to decrease the correlation coefficient without denoting poor performance*". A strong evidence supporting this thesis is that in spite of the higher individual correlation coefficients for McClear station by station, the average correlation is higher for SIRAMix (0.67) than for McClear (0.65). This discussion will be introduced in the new version of the manuscript.

Best regards,

The authors