<u>Response to referee #1 for manuscript:</u>

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 the concerns raised by referee #1. We have addressed the main points (P1-P9) by detailing several aspects of the proposed method SIRAMix and by clarifying some information given in our manuscript. Also, we have run additional sensitivity studies to answer the question of referee #1 regarding the impact on the DSSF estimation of the vertical profile assumed for aerosol particles (see P6). We have also added some of this additional information in the revised version of the article. Finally, we explain how minor comments made by referee #1 (see P10) have been addressed in the improved version of the 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.

Anonymous Referee #1
Received and published: 7 April 2014

"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" describes and assesses in detail a novel algorithm to derive surface shortwave radiation parameters from satellite observations, taking into account a combination of aerosols. This work is very significant, as direct and diffuse down-welling surface radiation fluxes are very useful in a number of meteorological and climate modelling activities. The method presented here appears to bring a large improvement in the quality of remotely observed DSSF, and the distinction between direct and diffuse fluxes, even though it is more sensitive to the aerosol load and errors are thus higher, are a definite plus. These results will be useful and used in a number of observing and modelling activities. They are also a novel and intelligent way of making use of the MACC-II reanalysis and near-real time aerosol products. The new products that are proposed (direct and diffuse surface down-welling fluxes, atmospheric radiative forcing of specific gases or aerosol species) are extemely appealing.
The SIRAMix algorithm is clearly explained, and the paper is overall well organized and easily readable. No major scientific shortcoming were spotted. The validation method (vs simulated observations, observations and state of the art algorigthms) is sound. The sensitivity study brings more insight to the product and also to its limitations. A few remarks, comments and questions are listed below. A (non-exhaustive) list of minor corrections is also added.

Dear Referee #1,

Thank you for your constructive comments and sensible suggestions. Please find our response to your comments below.

P1 - The proximity of this work to the modelling of the aerosol direct effect could be mentioned or discussed somewhere. This could place this work in a broader perspective and make more explicit the link between aerosol modelling and the method presented here. This could also provide more validation tools.

Thanks for your comment. This suggestion was partly addressed in Section 4.3 of the original manuscript, which investigates the capabilities of SIRAMix to monitor the direct radiative forcing (or direct effect) caused by aerosols and other atmospheric components. The physical parameterization at the core of SIRAMix provides the means to quantify this radiative quantity at the surface level given a set of atmospheric inputs. According to the suggestion of referee #1, we have completed this part of our study by adding a more complete discussion on this complementary asset of SIRAMix. This has been done throughout the article, but specially in Sections 5 and 6. This places our work in a broader perspective. We have stressed in the revised version of the article that the validation of the direct forcing products has not been carried as comprehensively as it has been done for the DSSF estimates since we wanted to emphasize the main objective of the paper: the estimation of diffuse and direct DSSF.

P2 - Line 27 : "This outcome (ie the fact that the MACC-II aerosol forecasts are less accurate than the analysis) will be taken into account in the forthcoming implementation of SIRAMix in the operational production chain of the LSA-SAF project". How do the authors propose to address this issue?

At the end of the conclusions of the revised manuscript we now state: "*While forecast MACC-II data will be used for the operational near real time LSA-SAF processing chain, reanalyzed MACC-II data will be used for the periodical reanalyses done in the LSA-SAF project.*". Indeed, the LSA-SAF products are associated to a timeliness of three hours, meaning that all land surface products are generated three hours after a given SEVIRI/MSG observation is acquired. In this context, we will firstly use forecast MACC-II data to generate the DSSF products in an operational manner. On the other hand, periodical reanalysis are carried out in the LSA-SAF project to update existent surface products, either using new versions of the algorithms or improved inputs (this new information has been included in Section 3.1.2). The use of MACC-II aerosol reanalyses will be considered in this case. We believe that this way of proceeding provides a good trade-off between accuracy and the near real time constrain required by the LSA-SAF project.

P3 - In section 2 line 194 : does there exist any reference for this formula?

Equation number 16 is a classical formula used in atmospheric radiative transfer. The reference (Sobolev, 1972) has been added.

Sobolev, V. V.: Light scattering in planetary atmospheres (translated as Light scattering in planetary atmospheres, Pergamon Press, Oxford, 1975), Nauka, Moscow, 1972.

P4 - Line 293 : MACC-II provides forecasts up to 5 days ahead. There are 11 aerosol pronostic variables (and not 9) : for both OM and BC both the hydrophilic and the hydrophobic components are takin into account

We agree with referee #1 and acknowledge our mistake. Indeed, the MACC-II system generates 11 AOD variables and not 9. We have corrected this mistake in the new version of the manuscript.

In this matter, we would like to remark that the AOD for the different bins of DU, SS, OM, BC are not accessible from the MACC-II data server (or at least from the download website that we used: http://apps.ecmwf.int/datasets/data/macc_reanalysis/). Only the total AOD corresponding to each one of the five major aerosol species (i.e., Dust, Sea Salt, Organic Matter, Black Carbon and Sulphate) was available. This information has also been included in the improved manuscript (see Section 2.3.3).

Also, we have performed the following modification concerning the nomenclature used to refer to the different aerosol species. First, the acronyms SU, DU, SS, OM, BC are used only for MACC-II aerosol species, in agreement with the MACC-II paper (Morcrette et al., 2009). Second, the acronyms WASO, INSO, SOOT, SSALL, and MIALL are used to refer to the GADS-based aerosol components defined in Appendix A and used by SIRAMix. The adoption of this new nomenclature is meant to avoid confusion, for example, between the sea salt components in MACC-II and GADS.

P5 - Lines 303-305 : the forecast time(s) concerned here should be mentioned

The following information has been included in the improved version of the manuscript (Sec. 2.3.3)

"For example, Cesnulyte et al. (2014) quantified the bias of forecast MACC-II AOD estimates from the fdmj experiment for a series of ground stations to be 0.02 in average but to range between -0.20 (26% of total AOD in the often polluted urban area of Xianghe) and 0.12 (36% of total AOD in the dusty Solar Village in Saudi Arabia). To cover a full day of AOD values, Cesnulyte et al. (2014) took hourly forecast AOD from time steps 1 to 12 h from forecast base times 00:00 UTC and 12:00 UTC."

P6 - Lines 318 - 325 : What is the impact of this height correction on AOD and on direct, diffuse and global DSSF? While this step is necessary due to the wide difference in resolution between the MACC-II products and the satellite pixel, the vertical repartition of aerosols that is used in this algorithm (ie exponential decrease with height), while certainly adequate on a climatolofical range, can at times differ a lot to the observed or analysed aerosol load. This height correction algorithm could be a source of errors for the whole proposed method, for regions with a marked orography, since it was shown in section 3 that results are very sensitive to AOD. As the authors already use MACC-II products in their algorithm, wouldn't it be possible to also use the same products to assess dynamically the aerosol decrease as a function of height above the ground? This point would perhaps deserve more discussion.

As explained in Section 2.3.3, the height of a given ground station may not be the same than the altitude considered in the corresponding MACC-II grid pixel. Since aerosols are not homogeneously distributed along the vertical, AOD values analyzed by MACC-II may not be adequate to be used directly in SIRAMix. In order to overcome this issue, SIRAMix adjusts the MACC-II AOD estimates to the station actual height. For that purpose, the vertical distribution profile of the aerosol particles must be known. In our study, we have assumed an exponential distribution of the aerosol concentration, with a concentration maximum at the surface level (see Appendix B2). This simple yet realistic distribution is adopted from (Hess et al., 1998), who provides different vertical profiles for different aerosol species. A similar strategy is used in other DSSF retrieval algorithms such as the McClear method (Lefèvre et al., 2013).

Sensitivity studies were run to evaluate the impact on DSSF estimation of assuming an exponential vertical distribution under the occurrence of a different vertical profile. Figure A shows two different aerosol vertical profiles that have been used in SIRAMix to simulate DSSF. The red line corresponds to a classical exponential vertical profile used for continental aerosols (Hess et al., 1998) while the black line assumes the presence of an aerosol cloud between 1 km and 2 km.

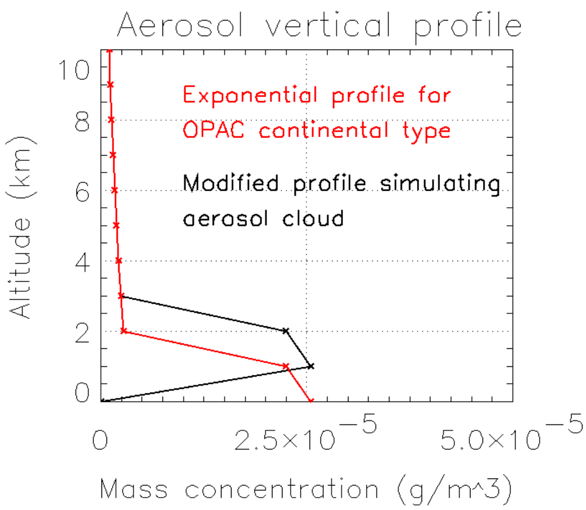

Figure A – Vertical profile of the mass concentration of aerosol particles considered to simulate DSSF with the proposed method SIRAMix. Red line corresponds to the vertical profile of a typical continental aerosol type (Hess et al., 1998). Black line is the result of modifying the previous profile to simulate an aerosol cloud not touching the ground.

Black lines in Figure B shows the diurnal evolution of the bias on global, direct, and diffuse DSSF calculated by SIRAMix due to the use of the exponential vertical profile (red line in Fig. A) instead of the configuration with the aerosol cloud (black line in Fig. A). Two aerosol contents (AOD=0.2 and AOD=1.0) are considered (see left and right columns in Fig. B, respectively). In addition, we have also plotted the bias resulting of using the correct vertical profile but an AOD that is affected by a ±10% bias (see red and blue lines). As it can be seen, results show that inaccuracies of using a wrong vertical aerosol profile are greatly lower (<5 W/m2) than those resulting from the AOD inaccuracies (up to 30 W/m2). This is especially true for standard aerosol conditions (AOD=0.2).

Given that MACC-II aerosol data may be biased by 10% or more, we concluded that the aerosol vertical profile is not a first-order parameter in the estimation of DSSF. In this context, we decided that the exponential vertical distribution was accurate enough for the purposes of SIRAMix. Besides this point, comprehensive experiments would be necessary to assess the use of MACC-II aerosol vertical profiles in SIRAMix, as uncertainties are supposed to exist in these analyzed data. This issue may be addressed in the future when MACC-II estimates will be improved.

Some of the previous information has been included in the improved version of the manuscript (see Apendix B2). However, Figures A and B in this document have not been included.

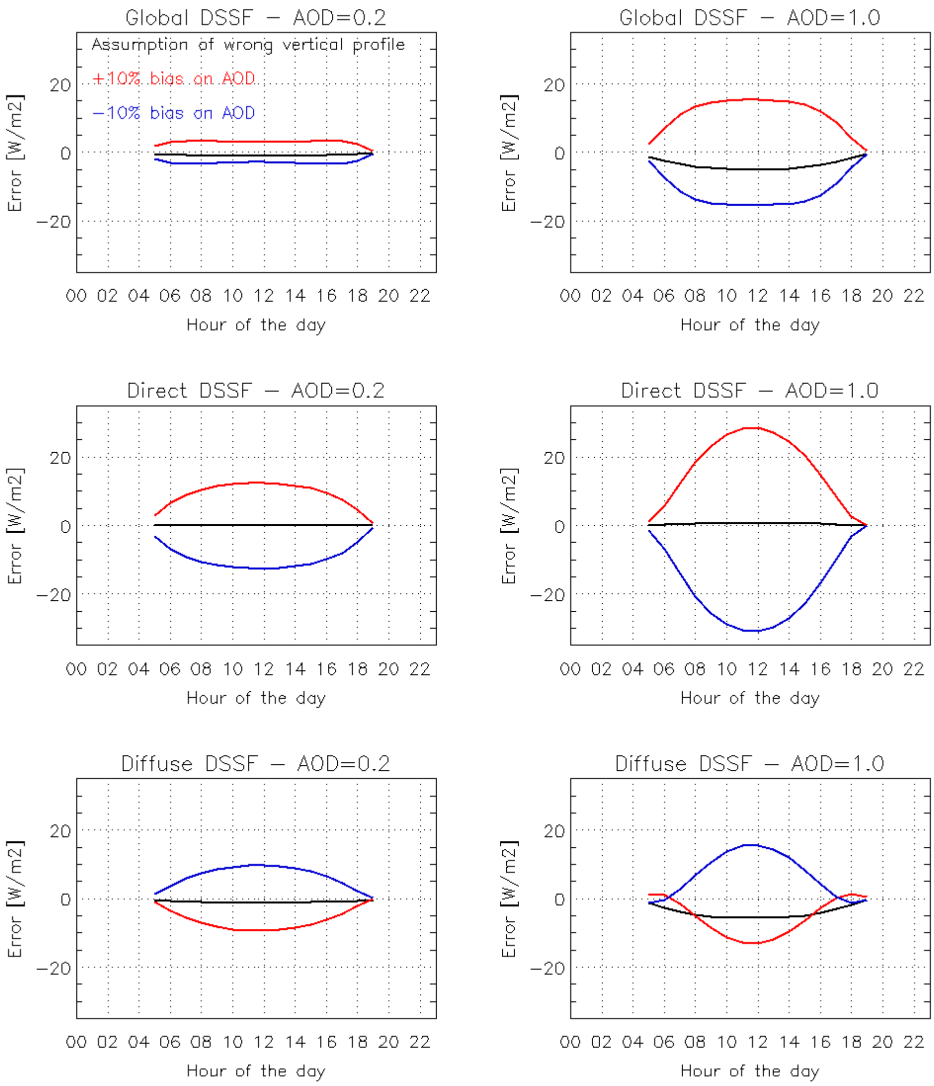

Figure B – Diurnal evolution of bias on global, direct, and diffuse DSSF when using (black line) a wrong aerosol vertical profile, (red line) a +10% biased AOD, or (blue line) a -10% biased AOD. For this experiment, we have selected the SZA corresponding to the station of Carpentras on June 21st and a typical mid-latitude continental atmosphere.

P7 - line 450 : the physical link between the errors on diffuse and direct fluxes should perhaps be more clearly explicited. Also, it appears that the positive bias (as compared to in situ observations and Mc Clear values) for the SIRAMix diffuse DSSF is more important than the negative bias of the SIRAMix direct DSSF, resulting in an overall small positive bias for the SIRAMix global DSSF. Is there any explanation for this?

First, we have added in Section 4.2 of the revised version of the manuscript a few sentences detailing the physical link between the errors on direct and diffuse DSSF. For instance, we now say "*The observed negative bias (e.g., -12.6 W/m2 for SIRAMix) may likely come from the overestimated aerosol load from MACC-II. Indeed, a lower AOD would result in a diminution of aerosol scattering that would increase the direct radiation reaching the ground.*" in Section 4.2.1 and "*The positive bias affecting all methods (e.g., 18.0 W/m2 for SIRAMix) may likely come from the overestimation of the AOD by MACC-II. The consideration in SIRAMix of a too high aerosol load would result in an increase of the atmospheric scattering and therefore of the diffuse radiation.*" in Section 4.2.2..

Second, Experiment 5 shows indeed an average positive error for global DSSF coming from (i) a moderate negative error affecting direct DSSF and (ii) a larger positive error from diffuse DSSF. Although this may be surprising given that diffuse DSSF is lower than direct radiation for average AOD values, we identify four reasons explaining this result:

1. Diffuse DSSF is more sensitive to inaccuracies affecting AOD when aerosol content is relatively small. This can be seen in Figure 5 middle-right (see plain line), which shows bias larger than ±10% for diffuse DSSF when AOD is 0.2 and the associated bias is ±25%. Under the same configuration, direct DSSF suffers only from a bias not greater than ±5%.
2. The physical parameterization used in SIRAMix is probably less accurate for diffuse DSSF, as the modeling of this radiation component has been historically more challenging than the direct one. However, experiments proved that it is more accurate than state-of-the-art methods such as McClear.
3. Likewise, the aerosol look up table generated with libRadtran is likely less accurate in terms of diffuse transmittances due to the same reasons given in point 2.
4. The accuracy of measurements taken by ground stations is generally lower for diffuse radiation, as the measuring technique is more complex than for the direct component.

The previous discussion has been included in section 4.2.2 of the improved manuscript as follows:

"*Despite the better estimation of diffuse DSSF by SIRAMix with regard to other methods, the mean RMSE represents a 41% of the average diffuse radiation (see Table 5). This is greater than for the direct DSSF, whose RMSE represents only the 12% of the average radiation (see Table 4). This difference is originated in the higher sensitivity of diffuse radiation to AOD inaccuracies under typical aerosol contents. This can be seen in Figure 5, which shows a bias larger than ±10% for diffuse DSSF when AOD is 0.2 and the associated bias is ±25%. Under the same configuration, direct DSSF suffers only from a bias not greater than ±5%. Also, the physical parameterization used in SIRAMix is probably less accurate for diffuse DSSF, as the modeling of this radiation component is more challenging than the direct one. Likewise, the aerosol look up table generated with libRadtran is likely less accurate in terms of diffuse transmittances due to the same reasons. Finally, the accuracy of measurements taken by ground stations is generally lower for diffuse radiation, as the measuring technique is more challenging in this case.*"

P8 - line 520 : a possible explanation of the higher RMSE for Tamanrasset and Sede Boqer (in addition to the fact that the aerosol load is generally much higher than other stations) is that the aerosol assimilation system of the MACC-II products uses MODIS observations of total AOD which are not available over deserts. This can affect the quality of the aerosol analysis over desertic areas.

Thank you for your comment. The following sentences have been included in Section 5 the improved manuscript: "*It is important to remark here that MODIS observations of total AOD are usually not available over deserts, therefore making impossible the aerosol assimilation carried out in the MACC-II system. This can affect the quality of the aerosol analysis over Sede Boqer and Tamanrasset.*"

P9 - line 580 : the agreement between the global estimate of aerosol SRF provided by SIRAMix and study that provides local values over North Korea is not enough to validate the aerosol SRF product.There exists many papers and studies trying to quantify and model the aerosol direct effect that could help provide a broader validation.

We agree with referee #1. According to his/her suggestion, we have added some references giving average values of aerosol SRF (SARF) in different places across the world for different periods of time. These values are in agreement with the average aerosol SRF obtained by SIRAMix for the year 2011. We have included the following paragraph in Section 4.3:

"*The aerosol direct forcing obtained by SIRAMix is in agreement with the values found in the literature for several regions of the world. For example, Péré et al. (2011) determined the mean aerosol SRF over the Mediterranean Basin in August 2003 to range from -10 to -30 W/m2 using a chemistry-transport model coupled with a meteorological model. On the other hand, ground measurements of DSSF allowed di Sarra et al. (2013) to quantify the average aerosol SRF on September 2005 as -24 W/m2 in the Mediterranean station of Lampedusa. A similar average was obtained by the MILAGRO (Megacity Initiative-Local and Global Research Observations) campaign in March 2006 over Mexico (Schmidt et al., 2010), which resulted in an average aerosol SRF of -22 W/m2. Likewise, Bush and Valero (2003) quantified the aerosol SRF over a mid-latitude region such as South Korea to be between -11 and -52 W/m2 (average of 30 W/m2). Finally, Mallet et al. (2006) and Roger et al. (2006) used measurements of microphysical and optical aerosol properties obtained during the ESCOMPTE (ExperimentS to COnstrain Models of atmospheric Pollution and Transport of Emissions) campaign to simulate average values of aerosol SRF equal to -(24-47) W/m2 for the southeast of France in June-July 2001.*"

P10 - Specific correction propositions (not exhaustive) :

- line 6 constant instead of unchanging

- line 10 : composed instead of constituted

These modifications have been done in the revised manuscript.

- line 11 : "real" is optimistic. Observed or analysed would maybe fit better?

We have replaced "*to match real aerosol conditions*" by "*to reproduce real aerosol conditions as best as possible*".

- line 19 : improve rather than decrease

- line 26 : issue instead of outcome

- line 35 : absence rather than lack

- line 36 : The latter particles

These modifications have been done in the revised manuscript.

- line 39-40 : "On the other hand" is not necessary

We have removed this from the revised article.

- line 59 : In particular usually doesn't start a sentence. "A static climatological aerosol load doesn't match the variability of aerosols in space and time"

- line 61 : "...doesn't describe accurately enough the usual mixture..." instead of "...is not correct in front of the usual mixture..."

These modifications have been done in the revised manuscript.

- line 66 "In front of" is not necessary. This sentence should be rewritten

Sentence "*In front of the poor knowledge on aerosols at broad scale, however, the description of aerosol properties had to be necessary simplified.*" has been replaced by "*However, the description of aerosol properties had to be necessarily simplified due to the poor knowledge on aerosols at broad scale.*".

- line 94 : existent instead of existing

- line 95 : "The upgrade consists of..."

- line 97 : "the abundance of which may vary with time and space". This part is not strictly necessary

- line 97 : "As explained"

- line 112 : Even though I see what you mean by "horizontally" (it refers to the surface and not the radiative flux), the term is misleading and should probably be removed.

- line 116-117 : this sentence should probably be move to the beginning of the section

These modifications have been done in the revised manuscript.

- line 132-133 : ", on mu0..., and on the factor v(t)". Why not write cos(theta0) in the formulae?

We prefer to keep using the Greek symbol *mu* for the cosine of the solar zenith angle for historical reasons.

- line 170 : computed instead of calculated, single scattering rather than singly scattered

- line 173 : idem

- line 269 : to monitor the aerosols

- line 291 : there was also the MACC project in between GEMS and MACC-II

- line 299 : made available

- line 301 : delayed mode

- line 350 : "minor" should probably be removed

These modifications have been done in the revised manuscript.

- lines 391-392 : "assesses the sensitivity of SIRMix to the variability of inputs" fits perhaps better

We would like to point out that Experiment 2 does not investigate the impact of the inputs variability but their uncertainties (or quality). The suggestion of referee #1 could make the reader think that we are considering the physical intra-variability of the inputs (the range of values within they naturally span), which is not the case. In this context, we prefer to keep the text as it is.

- lines 400-405 : while the difference between experiments 5 and 6 is clear when you go to the corresponding sections, it is not very clear as explained in these two bullet points. The sentence describing experiment 6 should probably be rewritten

We have further clarified this sentence by saying the following "*Experiment 6 in Section 4.3 investigates the capabilities of SIRAMix in quantifying the direct radiative forcing caused by aerosols and other atmospheric components. This feature of SIRAMix is made possible thanks to the accurate modeling of the downwelling atmospheric path done in the SIRAMix parameterization.*".

- line 413 : "shows a comparison" instead of "compares"

- line 460 : "As can be seen"

These modifications have been done in the revised manuscript.

Best regards,

The authors

<u>Response to referee #2 for manuscript:</u>

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 have added 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 has been included in Section 5 (see 3rd paragraph) of the revised version of the manuscript.

To further investigate this point, we have illustrated the dependency of diffuse DSSF on surface albedo. This has been 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 have updated 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%).

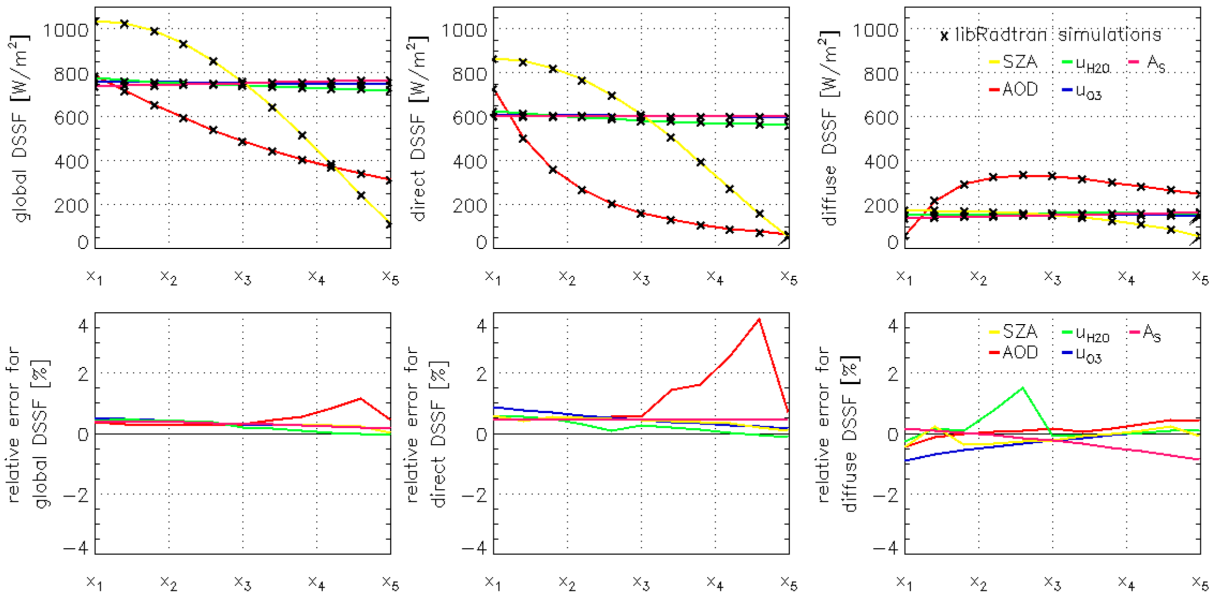

Figure A - *Improved version of Figure 4 after including the surface albedo (As) as a free parameter* - (top): Global, direct, and diffuse DSSF values calculated with SIRAMix according to varying AOD (red color), ozone content (blue color), water vapor concentration (green color), SZA (yellow color), 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 (x1, x2, x3, x4, x5) 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/cm2 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.

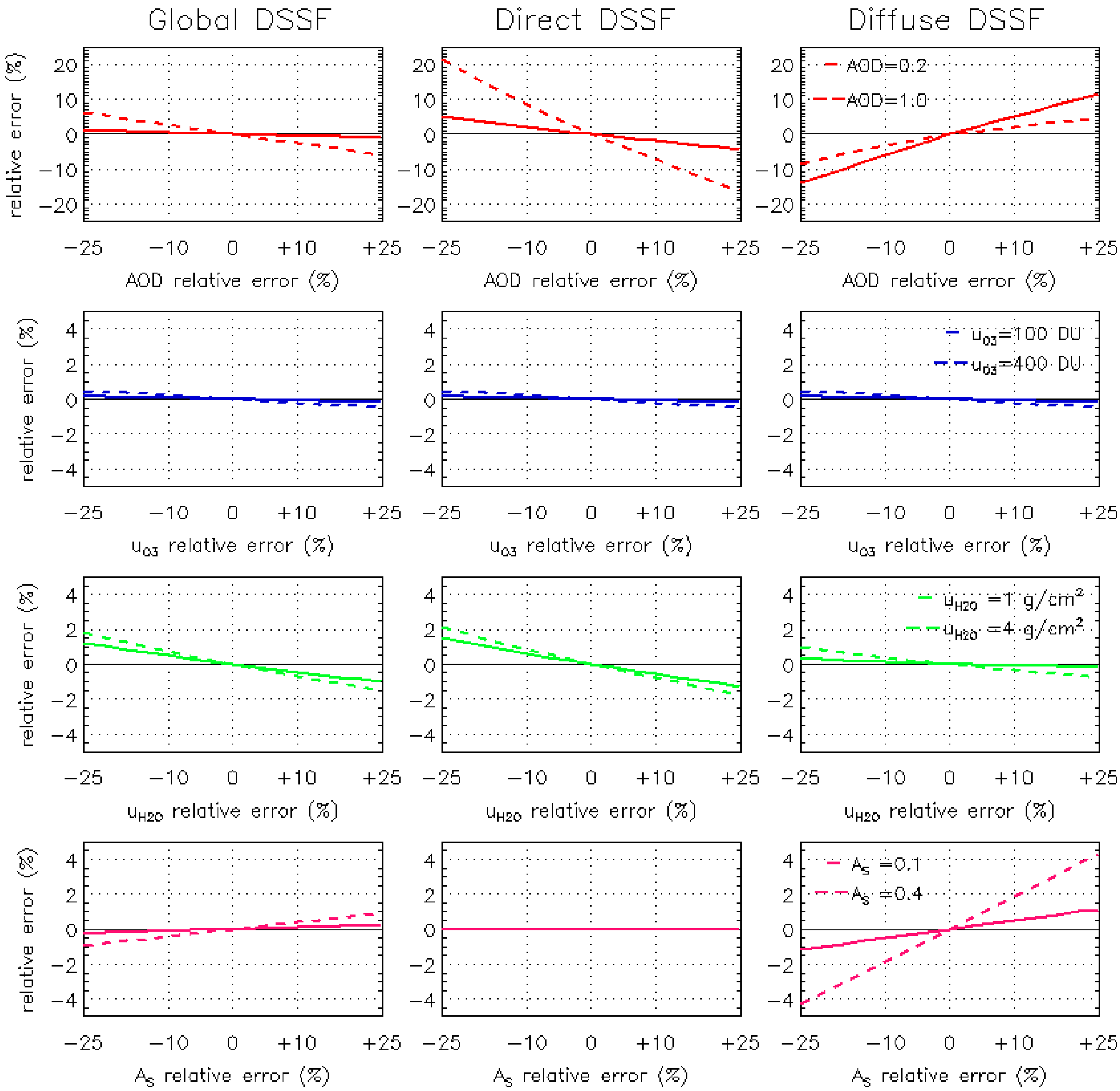

Figure B – *Improved version of Figure 5 after including the surface albedo (As) as a free parameter* - Relative error on global, direct, and diffuse DSSF calculated with SIRAMix caused by uncertainties in terms of relative error affecting AOD (1st line figures. Note different vertical scale), ozone 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 prefered not to include an additional study based on KT based on the following reasons:

1. 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.
2. 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).
3. 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.
4. 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.
5. 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 asses 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. A comment on this matter has been included at the beginning of Section of the revised manuscript.

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 now explained in detail in the new Section 2.3.2 of the revised manuscript. Likewise, the information on the clear sky strategy used by McClear is explained in Section 3.1.2.

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) has been discussed in Section 5 (see 4th paragraph) of the new version of the manuscript.

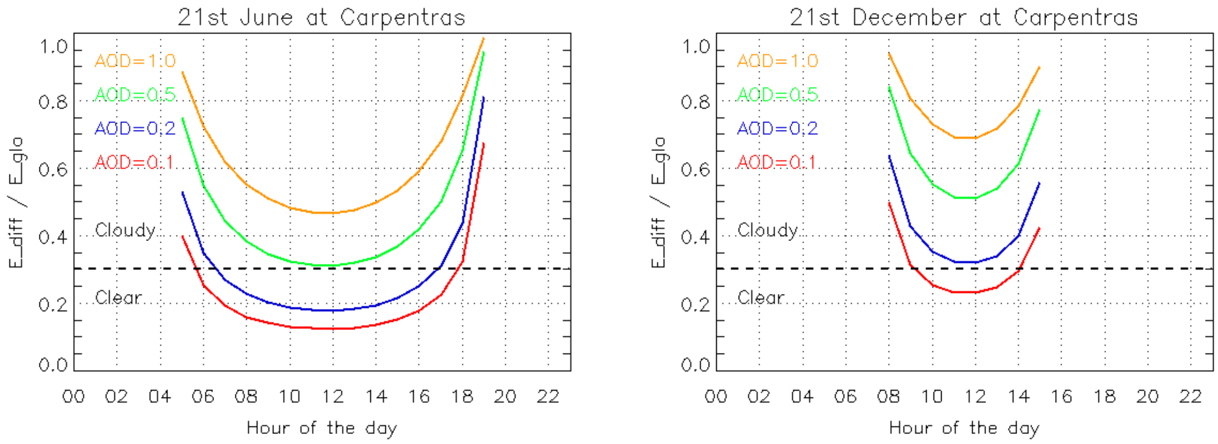

Figure C – Daily evolution of ratio *E_diff* / *E_glo* simulated by libRadtran over Carpentras for (right) the summer solstice and (left) the winter solstice. A U.S. Standard cloudless atmosphere was chosen with four different values of AOD: 0.1 in red, 0.2 in blue, 0.5 in green, and 1.0 in orange. The 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 have been clarified and can be found in Section 3.2 of 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 has been included in Section 5 (see 4th paragraph) of the new 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 has been included in Section 4.2.3 of the revised article.

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 has beeb included in Section 4.2.2 of the new version of the manuscript.

Best regards,

The authors