Here we first report the referee comments (in black), and then we provide our responses (in blue). In our replies, pages and lines (p. xx, l. xx) refer to the updated manuscript.

Anonymous Referee #1

This study evaluates the AMATERASS product of the Himawari-8 satellite based on the EXAM algorithm. Ground-based measurements from SKYNET, JMA and BSRN were used for this evaluation under all-sky and clear sky conditions. It is a well written paper which with some minor revisions it could be published in the AMT journal.

R -> We are grateful to Referee #1 for carefully reading our manuscript and for the helpful suggestions that allowed improving the quality of the study.

My most serious comments have to do (i) with the lack of comparable results from the bibliography and (ii) with the non-inclusion of an aerosol input to the model. This fact makes the evaluated product almost blind to the aerosol effects. However, the impact of this additional uncertainty was assessed, but there is still a reliability gap. I recommend the authors at least to provide some information about potential sources of operational aerosol optical properties and ways of integrating them into the EXAM algorithm (additional extension or a new model approach?). On the other hand, the discussion about RE effects is valuable for the improvement of satellite algorithms as to take them into account, but the overall impact in terms of estimated solar energy potential for the natural energy resource exploitation is meaningless compared to the aerosol effect.

Concerning point (i):
Due to limitations in the temporal resolution of the satellite products on which the algorithms are based, there is a lack of concurrent operational products of surface solar radiation with a resolution comparable to AMATERASS (i.e. at 2.5 min). Therefore, in the introduction of the revised manuscript (p. 3, l. 5-16), we reviewed additional references of validations performed at 15 min (Kosmopoulos et al., 2018; Qu et al., 2017; Ruf et al., 2016; Zo et al., 2016). Then, following your comment below (cf. page 7, lines 4-5), in the revised discussion section (p. 17, l. 1-27), we compared our findings with such previous results (both at 15 min and 1 h).

Concerning point (ii):
Although assimilation/forecast datasets from the European MACC project or the Japanese MRI/JMA are suitable sources of operational aerosol optical properties, aerosol estimations provided by Himawari-8 itself would be the best products to account for the aerosols effects in a future version of AMATERASS. Nevertheless, in order to implement such correction, two steps should be necessary. Firstly, the retrieval of aerosol optical properties (i.e. aerosol optical thickness and Angstrom exponent) must be accomplished by exploiting Himawari-8 observations. This activity has been initially performed by using the algorithms of Higurashi and Nakajima (1999) for ocean and Fukuda et al. (2013) for land and recently further improved in the retrieval of urban aerosols (Hashimoto and Nakajima, 2017). Such preliminary results are still under evaluation and an initial validation with observations recorded at the Japanese SKYNET stations showed encouraging results. Then, the following step will be the inclusion of Himawari-8 aerosol parameters into EXAM and the creation of an updated version which will include the aerosol effects on the solar radiation. As shown in Takenaka et al. (2011), EXAM was designed to account for aerosols in a neural network for clear sky conditions. In the original scheme, three aerosol optical properties (i.e. AOD, the imaginary part of the refractive index and the size distribution) and five additional parameters (i.e. solar zenith angle,
surface albedo, surface pressure, ozone, water vapor) were included in the neural network and the achieved results were satisficing. A similar approach will be used in the next version of EXAM.

In the revised manuscript, we included the above discussion at p. 5, l. 15-28.

In the Introduction section, the authors provide good bibliography about hourly-based validations but there is no reference for higher temporal resolution. There is a need here for additional references using satellite data in finer time steps.

In the revised introduction (p. 3, l. 5-16) we included further references of validations performed at high temporal resolution and described their main results (e.g. Ruf et al., 2016; Zo et al., 2016; Qu et al., 2017; Kosmopoulos et al., 2018). Then, following your comment below (cf. page 7, lines 4-5), in the updated discussion section we discussed our findings in the light of these previous results (p. 17, l. 1-27).

On page 2, lines 3-4, rephrase as "However, to implement such an EMS system, surface solar irradiance data must be supplied as accurately as possible."

R -> Thank you!

On page 2, lines 29-30, the percentages seem to refer to relative RMSE.

R -> Yes, thank you!

On page 3, line 7, provide reference.

R -> We included a further reference (Perez et al., 2016)

Page 3, lines 20-21: Indeed the RE can introduce bias into the evaluation of satellite estimates. Mention other uncertainties (altitude corrections, aerosol optical properties (aod, angstrom, ssa), sza and shading from adjacent mountains) with relevant references.

R -> In the revised manuscript (p. 4, l. 14-13), we included an additional discussion as follows: "Many additional factors can introduce a bias in the results of a validation exercise and make satellite-based estimates more uncertain (Polo et al., 2016). Among others, we remind the important role played by aerosols in reducing the surface solar irradiance, thus this is an important parameter to be accounted in the algorithms especially for polluted regions or deserts under clear sky conditions (e.g. Qu et al., 2017). Then, it is worth to mention the negative effect of the complex morphology of mountainous regions which cannot be easily accounted due to the limited satellite spatial resolution and the local fast-changing weather conditions (Dürr et al., 2010; Urraca et al., 2017; Federico et al., 2017). Further, the negative impact of the high solar zenith angle (SZA) and satellite viewing zenith angle on the quality of the estimates must be also mentioned: the former can cause low clouds to be overshadowed by higher clouds while the latter produces a parallax effect in the clouds position (Polo et al., 2016; Qu et al., 2017). Finally, uncertainties in the surface albedo especially over bright surfaces (e.g. snow or desert) would make hard cloud identification (e.g. Tanskanen et al., 2007)."

Page 4, line 10: The EXAM algorithm will include the aerosol effect in the same neural network or in a separate one only for cloudless conditions? The combination of clouds and aerosols in the same NN could result further uncertainties with aerosol mixtures, aerosol and cloud mixtures, multiple aerosol and cloud layers etc. Explain in brief how this inclusion will be addressed.

R -> In the revised version of the manuscript, here (p. 5, l. 15-28) we included the discussion at point (ii) described above. In particular, EXAM was originally designed to account for aerosols in a neural network for clear sky conditions (Takenaka et al., 2011). In the original scheme, three aerosol optical properties (i.e. AOD, the imaginary part of the refractive index and the size distribution) and five
additional parameters (i.e. solar zenith angle, surface albedo, surface pressure, ozone, water vapor) were included in the neural network and the achieved results were satisfying. A similar approach will be used in the next version of EXAM.

Page 4, line 11: The mentioned bias from the absence of aerosol inputs may be confused with the aforementioned RE and this is an issue for the reliability and accuracy of results. A clarification is needed here (e.g. underestimation for RE conditions and overestimation because of aerosols).

R -> In the revised version of the manuscript (p. 5, l. 11-13), we further discussed and made clear this potential issue. The overestimation of the satellite-based estimates caused by the absence of the aerosol correction has been evaluated only under clear sky conditions. On the other hand, the investigation concerning the RE effects, which contribute to decrease the positive bias of the satellite-based estimates, has been conducted under all sky conditions.

Provide also reference for the aerosol impact on radiation and relevant sensitivity analysis in order to quantify the overall bias by the absence of aerosol input.

R -> In the revised version of the manuscript here (p. 5, l. 3-10) we included further references for the aerosol impact on radiation (e.g. Xia et al., 2007; Cachorro et al., 2008; Di Biagio et al., 2009, 2010; Papadimas et al., 2012; Huttunen et al., 2014) and we reported that, under clear sky conditions, for East Asia previous studies found a daily mean direct aerosol forcing on surface global radiation ranging from -8 to -64 W/m² while for Japan (in the Kanto region) roughly in the range -8 to -23 W/m² (Kudo et al., 2010).

Moreover, we stated that although aerosols modulate the amount of solar radiation reaching the ground under all sky conditions, it is generally thought that, depending on relative position/altitude of clouds and aerosol layer, usually aerosol effects are small compared with cloud effects for the solar global radiation while their impact is more important for the direct solar radiation (e.g. Qu et al., 2017; Kosmopoulos et al., 2018).

Page 4, line 20: The applicability of EXAM with the Himawari-8 input, needs to be tested under various climatological conditions, so the selection of ground-based measurements may include stations in high altitude, mainland, near the sea, near to urban sites like Chiba etc. Here, one station is near urban site, and three affected by desert and continental regions. Discuss the representativeness of the stations selected as well as further necessary test cases for future similar evaluations.

R -> Following your comment, in the revised manuscript (p. 6, l. 10-18) we further extended the validation results and discussed the necessity of an additional analysis for specific climatological conditions. Here we included:

“Possibly, the reliability of the AMATERASS dataset needs to be tested under various climatological conditions. Therefore, ground-based measurements should include stations at high altitude, in the mainland, near the sea, near to urban sites etc. The SKYNET Chiba station can be considered representative of the urban conditions of the mainland region, while the other SKYNET stations can be roughly grouped as representative of a subtropical region possibly affected by desert and continental aerosols. Because of the necessity of examining other climates and different conditions and to extend our validation to the whole Japan, we accomplished further comparisons with respect to 47 stations of the Japanese Meteorological Agency (JMA) surface network of pyranometers, some of them also belong to the Baseline Surface Radiation Network (BSRN), relying on the rigorous quality control performed by JMA. This allowed to distinguish three main climatological regions and additional smaller areas presenting a distinct response.”
Then, we further faced this issue in Section 3.2 when discussing the new right panel of the updated Fig. 3 as follows (p. 9, l. 14 to p. 10, l. 11):

“The right panel of Fig. 3 shows the correlation between monthly RMSE and cloud fraction at each JMA station for January to December 2016 plotted over the Japan Digital Elevation Model derived from GTOPO-30 (https://lta.cr.usgs.gov/GTOPO30). Overall, the correlation was always positive but it ranged from more than 0.9 to less than 0.2 for the different stations. As Japan extends from north to south for about 3000 km, it is characterized by a variety of climatic regions which affected the pattern of the correlation. Indeed, according to previous studies (e.g. Ohtake et al., 2015), we can distinguish at least three main climatological regions and additional smaller areas. The first region (enclosed by the blue dashed line) is the norther part of Japan, mostly characterized by a subarctic climate, which provided a uniform response and small differences between Hokkaido and the north of the mainland (i.e. Tohoku). Here, except for two stations located in the Pacific sector of Japan, usually the correlation was lower than 0.4. Then, the large central region (within the red dashed line), which is characterized by humid and temperate climate, presents a more articulated pattern. The flat and strongly urbanized Kanto region (i.e. around Tokyo) showed the highest correlation values (r > 0.9). The SKYNET station of Chiba University is located here and it is supposed to be representative of this area. Then, correlations became slightly lower toward south with an evident distinction between the east coast, characterized by higher correlations, and the west coast, which presents lower values. It is worth noting that in winter the west coast is usually affected by elevate snowfall levels, while the Pacific coast usually shows frequent clear sky conditions. Fukue SKYNET station is located in the south-west sector of this region.

Stations in mountain regions usually present peculiar features that distinguish them from the stations located near the sea. A recent validation of three satellite-based radiation products over an extensive network of 313 pyranometers across Europe (Urraca et al., 2017) showed that stations sited in the Alps and Pyrenees have errors (i.e. RMSE and MB) two or three times larger than the ones of other locations. In mountainous regions, the altitude varies sharply and affects not only surface related parameters, but also the state of the atmosphere. Therefore, satellite models can fail in such areas because the spatial and temporal resolutions are not high enough to account for the sharp terrain and changing weather conditions (Dürr et al., 2010; Castelli et al., 2014). Accordingly, although not shown here, in the central and norther Japan we found larger RMSE values for stations in mountainous regions compared with seaside stations. Finally, we note a somewhat uniform correlation (r > 0.6-0.8) in the subtropical region (enclosed by the green dashed line) along the Pacific Ocean where the largest precipitation usually occurs. The SKYNET stations of Cape Hedo and Miyako are located in this latter region.

The peculiarity of the different regions is manifest when focusing on the annual cycle of RMSE (see inset in the right panel of Fig. 3). Generally, higher RMSE values were found in summer being the values highest in the subtropical region, followed by the central and the norther region. On the other hand, in winter RMSE values in the north were found similar to the ones of the subtropical region. Future analyses, based on longer time series, are expected to further highlight the satellite uncertainties at the different locations and, potentially, improve the accuracy of the satellite-derived dataset by site-adaptation methods (Polo et al., 2016).”
The overall evaluation was performed with Skynet, BSRN and other stations from the JMA, so the Title of this paper could be optionally renamed as ”... by ground-based measurements”.

R -> We changed the title as suggested.

On page 5, line 25 and page 9, line 30, is there a classification of sky conditions in the bibliography based on the CSI? This will be helpful for the readers in order to have a sense of quantification for the various sky conditions.

R -> although to our knowledge there is not such specific classification, using the clear sky index or the clearness index for defining the three main categories of sky conditions (i.e. clear, broken and overcast) is widely accepted. For example, some authors have performed analyses based on the following values: [0.00-0.34] overcast, [0.34, 0.65] broken, [0.65, 1.00] clear (e.g. Serrano et al., 2006; Bech et al., 2015). We included such considerations in the revised manuscript (p. 7, 1.16-18).

Page 6, line 4: Mention additionally the minimum solar energy potential in such large SZAs for strengthening this consideration.
We mentioned that for Chiba station the GHI is around 300 and 150 W/m² at SZA of 70 and 80, respectively.

On page 7, lines 4-5, there is a need for discussion of these results (MB and RMSE) with the mentioned (line 3) or additional references, as to provide direct comparison with similar approaches and satellites.

Following your advice, we extensively compared those as well as other results of the present validation exercise with previous references at both 15 min and 1 h resolution. Note that, in the revised manuscript, we faced this issue in the updated discussion section (p. 17, l. 1-27).

Page 7, line 17: This is an important aspect and it needs a short description of the magnitude of this effect (with reference).

A recent validation of three satellite-based radiation products over an extensive network of 313 pyranometers across Europe (Urraca et al., 2017) showed that stations located in the Alps and the Pyrenees have errors (i.e. RMSE and MB) two or three times larger than the ones obtained in most of the other locations. In such locations, the altitude varies sharply and affects not only surface related parameters, but also the state of the atmosphere. Satellite models fail on mountainous regions because the spatial and temporal resolutions are not high enough to account for the sharp terrain and changing weather conditions (Dürr et al., 2010; Castelli et al., 2014).

Note that we included this paragraph earlier in the new manuscript (p. 9, l. 30 to p. 10, l. 2).

Page 7, line 22: Is there a need for an altitude correction for the satellite estimations?

The negative bias occurs only in winter so it can hardly be connected with some specific altitude-related problem. However, only Nagano station is sited at a slightly relevant altitude (about 400 m a.s.l.) while the other stations are well below, so we cannot check altitudinal issues. On the other hand, issues mentioned in the previous comment likely affected the validation. For example, usually such locations also present a higher RMSE than other locations close to ocean. Potentially, for these locations, site-adaptation methods (Polo et al., 2016) could be important in improving the accuracy of the satellite-derived dataset.

Page 10, line 24: Provide comparable results from the bibliography.

We referred the reader to Fig. 4 of Nottrott and Kleissl (2010) where the authors showed a larger bias (annual mean 18 % and up to about 50 % in summer) when comparing the SUNY modeled dataset with weather stations in California.

Page 11, line 25: What about high aerosol loads without the impact of PW? Huttunen et al. (2014) explains only this (effect of water vapor on the determination of aerosol direct radiative effect). A more focused reference is needed.

We included additional more focused references (p.15, l. 18-21). For example, we reported the results of the study of Papadimas et al. (2012) where the direct radiative forcing efficiency was estimated for the whole Mediterranean basin. They obtained a direct radiative forcing efficiency at the surface of about 100 W/m² per AOD unit at the 96 % of the locations. Moreover, we also discussed the results of additional studies focused on the fact that the instantaneous forcing efficiency also depends on the different aerosol type as well as on the SZA (Cachorro et al., 2008; Di Biagio et al., 2009, 2010). Overall, these studies reported values roughly in the range 100-150 W/m² (Cachorro et al., 2008) and 100-200 W/m² (Di Biagio et al., 2009, 2010).
Percentages of energy attenuation are also welcome (lines 22-23) in order to be useful for pv reduced production because of high aerosol loads.

R -> In the revised version of the manuscript, in the original Fig. 9 (now Fig. 10, see below), we included two additional panels showing the attenuation in percentages. There, we tried to remove the SZA effect on AOD by computing the results in function of the aerosol optical depth slant (AODS i.e., AODS = AOD/cos(SZA), see Garcia et al., 2006) and we obtained -14.3 and -14.5 % per AODS unit at Chiba and Fukue, respectively.

Fig. 9 [updated, now Fig. 10] – (top panels) Scatter plots of the difference in clear-sky surface global radiation (SKYNET observations minus Himawari-8 estimates) and measured SKYNET AOD at 500 nm under clear-sky conditions between January and December 2016 at the Chiba (left panel) and Fukue (right panel) stations. The regression line, its slope, and the correlation coefficient are also shown. (bottom panels) Scatter plots as above but for percentage difference in global radiation and aerosol optical depth slant (AODS).

Finally, on page 12, lines 4-5, provide comparable results from the bibliography.

R -> these results are roughly consistent with a previous investigation (Kudo et al., 2010) based on ground-based observations and performed in a close location (Tsukuba, distance about 50 km). In this previous study the authors reported a direct aerosol forcing larger for spring to summer (-22 W/m²) and small in winter (-8.3 W/m²).

In the end of this paragraph the author may provide a short description of the potential EXAM upgrade with the inclusion of the aerosol impact.

R -> In the revised manuscript (p.16, l.29-32), the following short discussion was included: “A recent study showed that the diurnal variation patterns in Himawari-8 AOD data are consistent with those seen in SKYNET observations (Irie et al., 2017). Thus, Himawari-8 aerosol products would provide a unique spatial and diurnal variation information which, once included in the next version of the EXAM algorithm, is expected to reduce this positive bias under clear sky conditions.”

The high spatial and temporal resolution of the Himawari-8 satellite in conjunction with
near real-time algorithms like the EXAM, will improve the precision of the solar farms planning and production control with clear benefits for the local energy transmission and distribution system operators. This paper after the above corrections could be a step forward to the efficient and full integration of the natural energy resource to the electricity grid and will contribute to the development of upgraded energy management systems.

R -> thank you!

New references


