Answer to Referee #1

- **General comment**

This manuscript presents an interesting study on evaluating the usage of Meteosat Second Generation satellite derived solar and longwave radiation products in coarse-scale models. For this the radiation products were first compared to ground measurements in the French Alps and the Pyrenees as well as to forecast fields from the AROME model and analyses fields from the SAFRAN system. While the shortwave satellite radiation products showed lower errors with in situ measurements than modelled fields, a clear conclusion for the longwave satellite radiation products could not be drawn (differences around ground measurement uncertainties). Together with forecasts from AROME the satellite radiation products were then used to drive snowpack simulations using the snowpack model Crocus in the French Alps and the Pyrenees. An evaluation with measured snow depth revealed increased biases when using satellite-derived products.
Irradiances derived from the Satellite Application Facility on Land Surface Analysis (LSA SAF) are thoroughly evaluated over mountainous, snow-covered regions, at single points as well as analyzed spatially. The manuscript therefore presents a step towards assimilating high-resolution LSA SAF satellite irradiance products (3 km) over mountainous regions into models with grid cell sizes of only a few kilometers. Overall, the manuscript is well written and I suggest this manuscript to be published once the major comments and corrections listed below were addressed.

We thank the referee for the time dedicated to this review and his insightful comments. We answered below to all his points. His comments are in bold while our answers appear in blue. Changes in the manuscript appear in red.

- **Major comments**

My major comment or concern is about the used satellite-derived products which has an impact on the evaluation applied over mountainous regions. I might be wrong, but I could not find out if the satellite-derived products were corrected for topographical influences on radiation using digital elevation models which is however important.
For instance, spatial de-biasing in the shortwave radiation products needs to be conducted to reduce errors when applied over mountainous terrain. Meteosat Second Generation satellite-derived solar radiation was corrected before, see e.g. the HelioMont method (Stoeckli (2013), Castelli et al. (2014)).
Also, did you consider any topographic corrections for the downward longwave radiation product; I believe the algorithm suggested by Trigo et al. (2010) does not introduce limited sky view.
Please clarify and discuss both.

The reviewer is right to underline this concern about topographical influence on radiation in mountains. We will try to answer this concern in three points: 1) limitations of LSA SAF irradiance products in mountains; 2) limitations of the use of LSA SAF irradiance products to provide the radiative forcing for distributed snowpack simulations in mountains; 3) influence of the surrounding topography on the evaluation at stations. These elements are developed in a new section added to the discussion.

1) The Heliomont solar irradiance product (Stöckli, 2013; Castelli et al., 2014) is calculated using the MSG SEVIRI High Resolution Visible (HRV; 0.45-1.1 µm) channel and five other near-infrared and infrared channels (0.6, 0.8, 1.6, 10.8, 12.0 µm). In this method, the satellite data depending on the HRV channel (at 1 km resolution) requires an orthorectification to avoid artificial geometric shifts in terrain due to its high resolution compared to the terrain elevation (Stöckli, 2013), while the satellite data from the other channels (at lower spatial resolution) are not orthorectified. The DSSF and the DSLF only use data from the 0.6, 0.8 and 1.6 µm channels, which do not require orthorectification, similarly to the Heliomont method. These details have been added in the manuscript.
At a 2.5 km scale, the LSA SAF irradiance products have limitations specific to their use in mountains. Although the satellite observations used (cloud mask, top-of-atmosphere reflectance) have a similar spatial resolution, the products rely on meteorological inputs at a coarser resolution (ECMWF forecasts). To represent the topographical influence on these variables, a downscaling is required. As DSLF highly relies on its meteorological inputs, we used finer resolution meteorological inputs for DSLFnew. DSSF relies more on satellite observations, but still includes in the calculations the total column water vapour (TCWV) forecast by ECMWF. It is the main limitation of DSSF in mountains at kilometric scale, since the atmospheric vapour content depends on the elevation. The next step to improve solar irradiance products would thus be to develop a "DSSFnew" using TCWV forecasts at finer scale. It has been mentioned as a limitation and perspective in the new discussion section.

2) The context of this study is to assess different radiative forcings, including satellite-derived irradiance products, for distributed snowpack simulations at a 2.5 km grid spacing. The local topography (aspect, slope, surrounding terrain) is not taken into account in the snowpack simulations made on a flat terrain, similarly to Vionnet et al. (2016) and Quéno et al. (2016). The aim of these simulations is indeed to represent the mean state of the snowpack over the considered pixel (at a given altitude and a given location in the mountain range), thus discarding subgrid topographical specificities. Indeed, the 2.5 km resolution does not enable to reproduce the distribution of slopes and aspects found in a given mountainous area. Consequently, topographical influences on the incoming radiation such as slope/aspect effects, terrain shadowing, limited sky view factor and terrain reflections for SW↓, terrain thermal radiation and limited sky view factor for LW↓, are not taken into account for these ideal "flat pixel" simulations. Besides, they are not taken into account in the existing radiative forcings (SAFRAN and AROME). The aim of our study is to assess the practical benefits of LSA SAF products to provide a more accurate estimate of solar and longwave irradiances over a 2.5 km wide pixel for snowpack simulations. The main limitation of these simulations is that they do not capture a large part of the irradiance variability, strongly determined by the local topography, because these processes occur at a finer scale, requiring sub-grid radiation parameterisations, as developed by Helbig and Löwe (2012) or in the Heliomont method (Stöckli, 2013; Castelli et al., 2014).

3) The local topography affects in situ measurements used for the evaluation of the products. The station locations are generally set up in flat and open fields, which only partly reduces this influence. An effort was made to mitigate the impact of the surrounding terrain on the observations used for comparison to the irradiance products: for SW↓, the effect of terrain shadowing on direct radiation has been taken into account by discarding periods when the sun was masked by the topography.

However, the effects of limited sky view and reflections on diffuse radiation were too difficult to take into account. The terrain effects on measured LW↓ (limited sky view and terrain thermal radiation) were not computed either. Concerning the influence of different elevations in the comparisons, we preferred to indicate the elevation differences (Table 1) than apply a correction.

--- CHANGES IN MANUSCRIPT (lines 490-531) ---

5.2 Limitations due to the topographical influence on radiation

Limitations to the use of kilometric-resolution irradiance products in complex terrain arise from the high topographical influence on incoming radiation. These limitations are tackled here following three axes: (i) limitations of satellite-derived irradiance products in mountainous terrain, (ii) local topographical effects on radiation in the radiative forcing of snowpack simulations, and (iii) influence of local topography on the evaluation of the irradiance products and snowpack simulations.

First, satellite data sometimes require corrections when applied over mountains. For instance, the HelioMont solar irradiance product (Stöckli, 2013; Castelli et al., 2014) is calculated using the MSG SEVIRI High Resolution Visible (HRV; 0.45-1.1 µm) channel and five other near-infrared and infrared channels (0.6, 0.8, 1.6, 10.8, 12.0 µm). In this method, the satellite data depending on the HRV channel (at 1 km resolution) requires an orthorectification to avoid artificial geometric shifts in terrain due to its high resolution compared to the terrain elevation (Stöckli, 2013), while the satellite data from the other channels (at MSG pixel resolution, i.e. more than 3 km) are not orthorectified. The DSSF and the DSLF only use data from the 0.6, 0.8 and 1.6 µm channels, which do not require orthorectification, similarly to the HelioMont method. Corrections may also be applied to the meteorological inputs. The DSSF does not rely as much as the DSLF on meteorological forecasts but it still uses the total column water vapour content (TCWV) forecast from ECMWF IFS at 16 km resolution. Since the TCWV is dependent on the elevation, the DSSF could be improved with AROME forecasts of TCWV at kilometric resolution, similarly to DSLFnew. Despite that, the DSSF still exhibits a better performance than AROME and SAFRAN in mountains.

At sub-kilometric scale, the local topography strongly influences the solar and longwave irradiance variability. Oliphant et al. (2003) identified the following surface characteristics as causes of radiative flux variability, by order of importance: slope aspect, slope angle, elevation, albedo, shading, sky view factor, and leaf area index. These local factors are not taken into account in AROME, SAFRAN and LSA SAF irradiance products. This study aims at assessing the practical benefits of different irradiance datasets to be used as radiative forcing for distributed snowpack simulations at 2.5 km resolution in mountains. In the context of representing the mean state of the snowpack over a considered flat pixel, at a given altitude and a given location in the mountain range, the terrain influence on the radiation does not need to be taken into account in the radiative forcing. However, to capture the sub-kilometric variability of the snowpack, it will be necessary to consider sub-grid effects of the surrounding terrain on the radiation, and thus a topographical correction of irradiance products (e.g. Helbig and Löwe, 2012) as done for MSG satellite-derived solar fluxes by the HelioMont method (Stöckli, 2013; Castelli et al., 2014).

The main limitation implied by local topography effects regards the evaluation of the irradiance products and the snowpack simulations through in situ comparisons. Indeed, in situ irradiance and snow depth measurements are affected by these effects. The location of stations in flat and open fields reduces the impacts of slope, aspect and vegetation. The evaluation of solar irradiances at periods when the sun is not masked by the surrounding topography enables to discard the terrain shadowing effect on direct solar radiation. However, this effect is not considered for snow depth comparisons. Additionally, the limited sky view and the reflection effects on diffuse solar radiation are not taken into account, as well as the limited sky view and terrain thermal radiation effects on longwave irradiance.

- **Specific comments**

Line 102-104: What are you using the AROME forcing for? Please explain why you are using air temperature and relative humidity in 2m but wind and precipitation in 10m.

The word “forcing” was indeed used with no further explanation. We have explained in the new version of the manuscript the use of the forcing to drive snowpack simulations.
The air temperature and humidity are taken at 2 m above the ground and the wind at 10 m above the ground because it corresponds to the heights of the diagnostic variables provided by AROME. The detailed snowpack model Crocus uses forcings taken at these heights when driven by AROME. The precipitation is taken at ground level. It has been precised in the new version of the manuscript.

--- CHANGES IN MANUSCRIPT (lines 102-108) ---
In this study, we built a continuous atmospheric forcing dataset to drive snowpack simulations using hourly AROME forecasts issued from the 0 UTC analysis time, from + 6 h to + 29 h, extracted on a regular latitude/longitude grid with a 0.025resolution over the period and domains of study (Sect. 2.1, Fig. 1), similarly to Quéno et al. (2016) and Vionnet et al. (2016). Besides incoming shortwave and longwave irradiances, 2 m temperature and humidity, as well as 10 m wind speed and ground-level precipitation (amount of rainfall and snowfall) are part of the AROME forcing. The variable heights correspond to the heights of the diagnostic variables provided by AROME.

Line 120-122: How are the reanalyses interpolated at the exact station locations? Please clarify.

The reanalyses are interpolated at the station locations through a weighted mean of SAFRAN reanalyses at the two closest elevation levels in the considered massif. It has been mentioned in the new manuscript.

--- CHANGES IN MANUSCRIPT (lines 124-127) ---
For comparisons to in situ irradiance observations, the reanalyses were interpolated at the exact elevation of the stations, through a weighted mean of SAFRAN reanalyses at the two closest elevation levels in the considered massif.

Line 151-152 and Line 160-161 : Please clarify the given target accuracy, i.e. citation. How was it derived ?

The target accuracy is derived from a quality control of both products through comparisons with radiation measurements of the BSRN (Baseline Surface Radiation Network; Ohmura et al., 1998). A citation of the Product Requirement Document (Trigo and Viterbo, 2009) has been added.

--- CHANGES IN MANUSCRIPT (lines 157-159) ---
The target accuracy of the DSSF is 10% or 20 W m^{-2} for values lower than 200 W m^{-2} (Trigo and Viterbo, 2009).

--- CHANGES IN MANUSCRIPT (lines 171-172) ---
The target accuracy of the DSLF is 10% (Trigo and Viterbo, 2009).

Line 170 : If possible, can you add the approximate or range of height of the "first operational atmospheric level" of AROME?

More exactly, air temperature and dew point are taken at 20 m above ground in the archive of the AROME operational forecast. This height corresponds approximately to the height of the first prognostic level in the operation version of AROME over the period 2010-2014 (Seity et al., 2011). It has been added in the new manuscript.

--- CHANGES IN MANUSCRIPT (lines 180-183) ---
Air temperature and dew point were taken at 20 m above ground in the archive of the AROME operational forecast. This height corresponds approximately to the height of the first prognostic level in the operation version of AROME over the period 2010–2014 (Seity et al., 2011).

Section 2.2.4 "New DSLF product using AROME forecasts" : What is the reason that you first interpolate the AROME forecasts to the LSA SAF grid? Would it be possible to apply the algorithm directly on the AROME grid assuming the same cloud fraction in all AROME grid cells covered by the coarser LSA SAF grid cell? Maybe this way you could profit from the higher resolution temperature fields as the improvement between DSLF and DSLFnew is not that obvious based on Figure 3.

We chose to interpolate AROME forecasts to the LSA SAF grid because we wanted to generate a new product (DSLFnew) on the exact same grid as DSLF, in order to enable direct comparisons (e.g. in Fig. 5d and Fig. 6d). This product is thus on the same grid as the observed cloud mask, which provides the most important input to derive the LW irradiance.

--- CHANGES IN MANUSCRIPT (lines 186-188) ---
The new product was generated on the exact same grid as DSLF, in order to enable direct comparisons, so AROME forecasts were interpolated over the LSA SAF grid through a closest-neighbour method (similar grid spacing).

Line 182-183: I believe the statement that elevation is one of the most significant factors of surface radiation needs clarification. I guess this depends on scale, i.e. represented topographic complexity. There are also differences for shortwave and longwave radiation (as you also found (Figure 7)). Please discuss.

We agree with the reviewer that stating "elevation is one of the most significant factor of surface radiation variability" would require an additional discussion. However, this discussion would have no place in this section of dataset description. So the sentence was reformulated and a discussion about factors of surface radiation variability is made in the new discussion paragraph.

--- CHANGES IN MANUSCRIPT (lines 196-198) ---
As elevation influences incoming radiation (Oliphant et al., 2003), stations were not used for evaluation if the difference between the station elevation and the elevation of the four closest AROME and LSA SAF grid points was higher than 300 m.

Line 230-233 : Did you evaluate the scenario : shortwave from DSSF and longwave from DSLF ? How do the results compare to those from your scenario c) ?

This scenario was not evaluated, because of the similarity of DSLF and DSLFnew, compared to the other LW irradiance products. DSLFnew was chosen because of it mitigates the negative bias of DSLF up to 2200 m.

Line 242-243 : Why did you select a maximum elevation difference of 150 m between AROME grid cell elevation and station elevation for compiling a set of suitable snow depth measurements? In Line 185 you selected a maximum elevation difference of 300 m

between AROME grid cell elevation and LSA SAF grid points. What are the reasons for the differing values? Please discuss.

For snow depth measurements, a maximum elevation difference of 150 m between model grid point and station was chosen in order to keep the same observation dataset as in Quéno et al. (2016) for the Pyrenees and Vionnet et al. (2016) for the Alps. This value enabled to mitigate the differences of snow depth between simulations and observations arising from elevation differences, and keep at the same time a significant and representative ensemble of stations (172).
For irradiance measurements, the tolerance of elevation difference had to be higher due to the scarcity of stations in the Alps and the Pyrenees, in order to keep a representative dataset. As elevation differences up to 300 m can influence the comparisons of irradiances, the altitude of grid points and stations is indicated in Table 1.
A short discussion has been added in the manuscript.

--- CHANGES IN MANUSCRIPT (lines 199-201) ---
As elevation differences up to 300 m may have an influence on the comparisons, the altitudes of the grid points associated with each station are listed in Table 1 and should be kept in mind when analyzing the evaluation statistics.

--- CHANGES IN MANUSCRIPT (lines 261-263) ---
Only stations with less than 150 m elevation difference to the model topography were selected, in order to use the same dataset as Quéno et al. (2016) and Vionnet et al. (2016).

Line 253-255: Please mention briefly which method you used to derive the terrain horizon, e.g. interval size or add a reference.

The terrain horizon is calculated at an interval size of 5°, from a 25 m resolution DEM. Figure R1 provides an example of the terrain horizon at Bassies station.

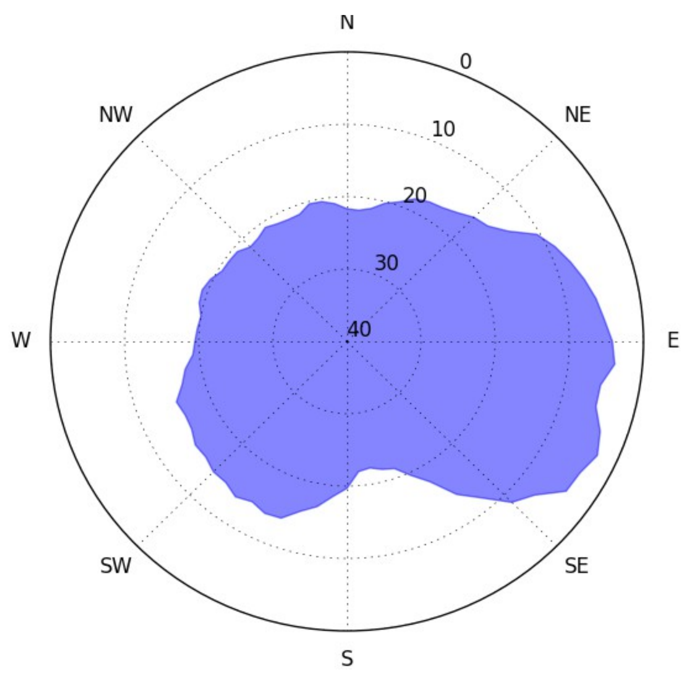

Figure R1. Terrain horizon (in degrees) at Bassies station (1650 m, Pyrenees), calculated at 5° intervals.

--- CHANGES IN MANUSCRIPT (lines 273-277) ---

To account for topographic shading on irradiance in situ measurements, a topographic mask was computed with a 5° interval size after a 25 m resolution digital elevation model (DEM) of IGN (French National Institute of Geographical and Forest Information), and applied to the SW↓ irradiance products at all stations except Andorre and Envalira, because the DEM of IGN was only available on the French territory.

Line 255: Please specify why the horizon was not computed for Andorre and Envalira. Are those stations without topography in the surroundings?

The horizon was not computed for Andorre and Envalira because the DEM of IGN (French National Institute of Geographical and Forest Information) was not available for these stations located outside of the French territory. These stations are also located in mountains, that is why a threshold SW value was used to discard periods when the sun was masked by the topography. It has been specified in the new manuscript.

--- CHANGES IN MANUSCRIPT (lines 273-277) ---

To account for topographic shading on irradiance in situ measurements, a topographic mask was computed with a 5° interval size after a 25 m resolution digital elevation model (DEM) of IGN (French National Institute of Geographical and Forest Information), and applied to the SW↓ irradiance products at all stations except Andorre and Envalira, because the DEM of IGN was only available on the French territory. *The SW↓ irradiance products were only evaluated when the sun was above the horizon, or when the observed value was higher than 20 W m^{-2} at Andorre and Envalira stations (to discard periods when the sun is masked by the terrain).*

Line 273: Can you add a similar table for the longwave measurements as Table 1 for the shortwave measurements? The table would give additional inside to the performance with regards to measurement uncertainties and altitude differences between model grid cell and station.

Table 1 actually lists measurement uncertainties and altitude differences between model grid cell and station for LW (columns 3 to 6). Contrary to Fig. 2 where the SW metrics are aggregated by domain and range of altitude, Fig. 3 shows bias and RMSE at each station measuring incoming LW fluxes. We think it would be redundant to add a table with these metrics.

- **Technical comments**

Line 40: Consider removing "were".
Done.

Figure 4: Please increase all labels and legend.
Done.

Line 295: Please rephrase : "Whatever the hour, AROME overestimates SW."
Done.

Figure 9: Please rephrase the last sentence in the caption.
Done.

Line 442 : Consider refering to Figure 2.
Done.

Line 446 : Consider refering to Figure 3.
Done.

Line 534-536 : Consider adding to "…due to a too strong altitudinal gradient. " that the gradient arises from the cold bias in AROME air temperatures.
Done.

References

Castelli, M., Stöckli, R., Zardi, D., Tetzlaff, A.,Wagner, J., Belluardo, G., Zebisch, M., and Petitta, M.: The HelioMont method for assessing solar irradiance over complex terrain: Validation and improvements, Remote Sens. Environ., 152, 603–613, doi:10.1016/j.rse.2014.07.018, 2014.

Helbig, N. and Löwe, H.: Shortwave radiation parameterization scheme for subgrid topography, J. Geophys. Res. Atmos., 117, doi:10.1029/2011JD016465, 2012.

Ohmura, A., Gilgen, H., Hegner, H., Müller, G., Wild, M., Dutton, E. G., Forgan, B., Fröhlich, C., Philipona, R., Heimo, A., König-Langlo, G., McArthur, B., Pinker, R., Whitlock, C. H., and Dehne, K.: Baseline Surface Radiation Network (BSRN/WCRP): New Precision Radiometry for Climate Research, Bull. Amer. Meteor. Soc., 79, 2115–2136, doi:10.1175/1520-0477(1998)079<2115:BSRNBW>2.0.CO;2, 1998.

Quéno, L., Vionnet, V., Dombrowski-Etchevers, I., Lafaysse, M., Dumont, M., and Karbou, F.: Snowpack modelling in the Pyrenees driven by kilometric-resolution meteorological forecasts, The Cryosphere, 10, 1571–1589, doi:10.5194/tc-10-1571-2016, 2016.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France convective scale operational model, Mon. Weather Rev., 129, 976–991, doi:10.1175/2010MWR3425.1, 2011.

Stöckli, R.: The HelioMont Surface Solar Radiation Processing, Tech. Rep. 93, MeteoSwiss, https://www.meteoswiss.admin.ch/content/dam/meteoswiss/de/service-und-publikationen/Publikationen/doc/sr93stoeckli.pdf, 2013.

Trigo, I. and Viterbo, P.: Product Requirement Document, Tech. Rep. 1.11, The EUMETSAT Satellite Application Facility on Land Surface Analysis (LSA SAF), https://landsaf.ipma.pt/GetDocument.do?id=281, 2009.

Vionnet, V., Dombrowski-Etchevers, I., Lafaysse, M., Quéno, L., Seity, Y., and Bazile, E.: Numerical weather forecasts at kilometer scale in the French Alps: evaluation and applications for snowpack modelling, J. Hydrometeor., 17, 2591–2614, doi:10.1175/JHM-D-15-0241.1, 2016.