Interactive comment on "Determining the infrared radiative effects of Saharan dust: a radiative transfer modelling study based on vertically resolved measurements at Lampedusa" by Daniela Meloni et al.
Anonymous Referee #2
Received and published: 28 October 2017

Review of paper: acp-2017-591 "Determining the infrared radiative effects of Saharan dust: a radiative transfer modelling study based on vertically resolved measurements at Lampedusa" by D. Meloni et al.

General comments
In this paper radiation closure experiments are made in order to determine the infrared radiative effects of dust and to assess the role of dust size distribution (SD) and refractive index (RI). To this aim, in situ data from aircraft (ATR-42 and Falcon), surface (AERONET, radiometer, pyranometers, pyrgeometers and pyrometer) radiosonde and satellite (IASI) measurements are utilized for the closure. The measurements come from the ADRIMED/ChArMEx campaign in 2013. The vertically resolved simulations are performed with the MODTRAN radiative transfer model (RTM) initialized by insitu vertical and remotely sensed columnar SD and RI along with data for a series of surface and atmospheric parameters relevant to LW radiation transfer, coming from radiosoundings, spectrophotometer measurements, ECMWF reanalysis and MODIS satellite products. The assessment lies in comparing simulated and measured LW irradiances and brightness temperatures (BTs), while the dust LW radiative forcing (ARF) and atmospheric heating/cooling rates (AHR) is estimated with the RTM. Three cases (summer days) during a period of dust intrusions (late June and early July) are examined, and the study is performed for Lampedusa in central Meditteranean, in proximity to northern Africa and Sahara.
The study is detailed and makes synergistic use of a variety of data. Some interesting findings are reported, from which some are not always new, e.g. that the dust LW radiative effects are non negligible or that the heating rate profile of dust depends on its vertical distribution as well as on SD and RI. Yet, some others provide new information and give insight regarding the role of dust SD and RI for their LW radiative and thermal effects and for BT, e.g. that using dust RI from local dust sources (Algeria and Morrocco, DB2017) produces best agreement with observations or that the use of inaccurate, although optically equivalent SD and RI has a large impact on the dust ARF. The paper is well organized and nicely written although it sometimes lacks clarity in the discussion of its results.
The main issue is that the paper seems to fail to convince about the best performance and appropriateness, and to provide a clear message on what is the optimal combination of dust properties for achieving the radiation closure. The relevant messages drawn from the simulations-measurements comparisons of LW and WINDOW fluxes, and of BTs, are not consistent and appear to be somewhat contradictory, as it is for ex- ample the case in Table 3. Even the authors state (page 15, lines 20-21) that "the MODTRAN spectral resolution impacts the standard deviation of the model-measurements differences, making the results obtained with different AOPs equivalent". More specifically:

Main Comments
1) In general, quite small differences between the 7 examined configurations, consisting in different model setups (Table 2), are found between results obtained without aerosols and with aerosols, as well as between the 6 configurations with aerosols (3 columnar and 3 in-situ). This does not help to draw a clear conclusion on which one configuration and aerosol properties combination is the best, although this is expected to be the main finding of such a radiation closure study.

The study is aimed at investigating how dust particles affect various radiation quantities (irradiance and BT at the surface and in the atmosphere, BT at the top of the atmosphere) compared to the aerosol-free case and how the magnitude of the dust radiative effect depends on different aerosol optical properties.
This study has been carried out with information on the atmospheric vertical structure and surface characteristics, as well as on the aerosol burden and physical properties, derived from observations. Specific values of the dust complex refractive index, including some recently determined region-dependent values, have been used. The model outputs are then compared with measurements from various instruments installed on different platforms (from the surface radiometers and pyrometer to the airborne radiometers, to the satellite IASI interferometer).
The aim is not the determination of the best combination of aerosol size distribution and refractive index. In fact, we have used the vertically resolved in situ measured size distribution as a reference, since it is directly measured and, in our opinion, best represents the occurring aerosol properties. Under this assumption, we show that the more recent determination of regionally dependent refractive indices perform better than the frequently used literature values.
The use of vertically averaged SDs derived from AERONET with the most commonly used values of RIs (as it is quite frequently done in similar studies) is finalized at assessing the influence they have on the radiation field.
It must be pointed out that the closure is done on a quite large number of radiation measurements, made at the surface, airborne (and at different altitudes), and on satellite. In our opinion this is a quite unique analysis, and the comparison with different types of radiation measurements gives strong constraints and robustness on the results.
It must be also said that unfortunately the atmospheric conditions occurring during the ChArMEx/ADRIMED campaign did not bring large AOD and this aspect, combined with the model and measurements uncertainties, causes the modelled LW irradiances to be equivalent. Nonetheless, we show that this is not true for the WINDOW and the IR zenith BT, for which significant differences are obtained for the *in situ* SD when using OPAC and DB2017 RIs.
For the day with the largest AOD, i.e. 22 June, we also assessed which combination of SD and RI gives the best model-measurement match for the overall set of LW irradiances (downward at the surface, upward and downward components on the ATR-42 and Falcon 20) by calculating the RMSD of all model-measurement absolute differences, and selecting only those AOPs for which the RMSD is below the ±5 W m^{-2} threshold value. For the AERONET SD all the three RIs meet the requirement (RMSD between 3.2 and 3.3 W m^{-2}), while for the *in situ* SD only the DB2017 (RMSD 4.7 W m^{-2}). This conclusion has been added at the end of Section 4.2. If we assume that the vertically-resolved *in situ* SD is the best representation of the effective SD, the DB2017 RIs provide the best AOPs.

2) The ascertained/computed differences of each one of 7 configurations with respect to measurements (LW, WINDOW, BTs) mostly fall within the range of uncertainty of measurements, making difficult to decide on which one is really the best configuration.
This is true for the LW and WINDOW irradiances, but not for the zenith BT. We added a sentence at the end of Section 4.1 summarizing the best combination of SD and RI that provide the best model-measurement match: “The final results of the analysis of the surface measurements show that irradiances, either broadband and in the 8-14 µm spectral interval, are not useful to reduce the uncertainty on the dust RI, since the impact of different RIs is below the measurement and model uncertainty. On the contrary, narrowband zenith BT seems to be suitable to constrain the dust RI which better represents the dust optical properties either in moderate and in low dust loading conditions. Under the assumption that the in situ SD is the most representative of the real aerosol dimensions, the DB2017 RI provides the best agreement between model and measurements, either LW and WINDOW irradiances, and sky BT in the two cases where the atmospheric meteorological profiles and the in situ SD are measured down to surface level (22 and 28 June).”

3) A main conclusion drawn from the analysis is that there is a systematic model overestimation of upward LW fluxes within the peak of dust layers, in all 3 days. In other words, there seems to be an inherent problem with the modelling tool, which needs to be assessed.
We believe that a modeling problem is difficult to expect, either because the CLIMAT BT, obtained with the same input parameters as the irradiance components, are very well reproduced, and because for the 22 June case the model succeeds in resolving both irradiances and BTs. We have further investigated this aspect and have tentatively attributed the observed bias to the CGR4 slow time response; in fact, significant model-measurements differences are found when the aircraft velocities during the descents are too high. See details of the answer to point 26 of the specific comments.

4) The estimated small differences between the no-aerosol and aerosol configurations, indicate that the RTM LW computations are relatively insensitive to dust.
That is true but depends mainly on the AOD value. For the 22 June case, (AOD at 500 nm of 0.36) the increase in LW irradiance at the surface due to dust compared to aerosol-free conditions (NOAER) is +4.8, 4.7, 3.3, 10.9, 10.6, and 8.1 Wm^{-2} with COL1, COL2, COL3, INSU1, INSU2, and INSU3, respectively. With the *in situ* SD the dust effect is larger than the uncertainty of LW irradiance measurement (5 Wm^{-2}) and of the model (4.2 Wm^{-2}).
The values decrease with AOD: indeed, the values for the 28 June case (AOD at 500 nm of 0.21) are +3.2, 1.7, 6.3, 5.1 Wm^{-2} with COL1, COL3, INSU1, and INSU3. This explains why the simulations in aerosol-free conditions agree with measurements within their uncertainties.
For larger AOD, like the case presented in the paper by Meloni et al., 2015, Altitude-resolved shortwave and longwave radiative effects of desert dust in the Mediterranean during the GAMARF campaign: Indications of a net daily cooling in the dust layer, J. Geophys. Res. Atmos., 120, 3386–3407, the perturbation induced by the dust with AOD at 500 nm of 0.59 was 16.2 and 16.1 Wm^{-2} with AOPs analogous to COL1 and INSU1, respectively. In that case the modelled LW irradiance in aerosol-free conditions is outside the expanded uncertainty of the measurements.

5) The reported conclusions are sometimes contradictory. For example in page 17, lines 14-15 it is stated that dust RI from DB2017 produces the best agreement with observations, but this is not supported by and it is not in line with the results of Table 3 where NOAER and COL1 also provide good results, even better than INSU3, if all three parameters, i.e. LW, WINDOW, BT, and three days are considered.
The sentence refers to the results obtained with the *in situ* size distribution. One point that has been clarified in Sections 4.1 and 4.2 is that we can assume that the *in situ* SD is more realistic than the AERONET one, since the first is derived from vertically resolved optical counter measurements covering the diameter range 0.032-32 µm, while the second is derived from surface visible and near-infrared radiance measurements and is representative of the whole atmospheric column. Under this assumption the best model-measurement agreement is obtained with INSU3 AOPs, as stated in page 15, line 32.

6) It is not clear why BTs were computed and are reported only at 3 levels, which sometimes are not collocated with the peaks of dust layers; why similar BT computations were not made at more levels.
The selected altitudes correspond to those witha horizontal flying attitude of the ATR-42. Moreover, while all the other simulated profiles (upward and downward LW irradiances) where obtained with a single model run, the CLIMAT BTs require a run for each altitude. The vertical profiles of the modelled irradiances show a smooth change with altitude, so we believe that BT at few altitudes are sufficient to describe the vertical variations and adding simulations at other altitudes would not provide additional information.

7) The conclusions drawn from the BT analysis are different from those obtained from the analysis of LW fluxes. This is for example the case of the results of profile 42, in Figures 10 and 11. May this point to a possible modelling problem/inconsistency?
The model has proven to perform well either for the LW fluxes and for the CLIMAT BTs on 22 June, so we cannot justify the results for flight F42 with model inconsistency. We also exclude as a cause an incorrect choice of the model input parameters, like the sea surface temperature, emissivity, and/or the atmospheric temperature/humidity profiles, because the CLIMAT BTs are well reproduced at all altitudes. Following a further check on the data, We found that the bias in the LW irradiances occurs where the ATR-42 descent velocity is high. We believe that the associated fast change of ambient air temperature, associated with the relatively long response time of the CGR4 pyrgeometer, may produce the bias. See details of the answer to point 26 of the specific comments.

8) The role of clouds is not reported. Were all the tree days/cases cloud-free? If so, how is this confirmed/ensured? A relevant discussion should be made since the effect of clouds on LW is significant and interplay or even dominate the effect of dust (e.g. possible implications for Fig. 2).
The three flights were carried out under cloud-free conditions.
The sky conditions at Lampedusa were monitored by the TSI-440 sky imager, collecting hemispheric pictures every minute. The absence of clouds during the flights is also confirmed by the time series of downward SW, LW and WINDOW irradiances, and by the information provided by the zenith-looking pyrometer and lidar at the surface. The profiles of downward LW irradiance from airborne radiometers also show that clouds were not detected: indeed the signal due to cloud emission would have been evident in the measurements with positive spikes, like those in Figure 2. A sentence has been added to the description of Figure 2, to highlight the large increase in irradiance/BT due to cloud presence.

Specific Comments
1. Page 1, line 28: define IASI acronym.
The acronym has been defined.

2. Page 4, Figure 1: the AERONET AOD may also be overplotted.
We received a comment from Reviewer #1 saying that we presented many instruments but some of them were not used in the analysis, and he/she suggested to describe only instruments and measurements that were actually used. We used MFRSR AOD measurements because of their higher temporal resolution (about 1 minute) compared to that of the Cimel sunphotometer (about 15 minutes). Moreover, some Cimel data are missing because of some malfunctioning of the instrument.
However, di Sarra et al., Empirical correction of MFRSR aerosol optical depths for the aerosol forward scattering and development of a long-term integrated MFRSR-Cimel dataset at Lampedusa, Appl. Optics, 54, 2725-2737, 2015, compare MFRSR and AERONET AOD measurements, deriving a mean bias in AOD not larger than 0.004 and a root mean square difference ≤0.031 at all wavelengths. Plotting the available AERONET AOD with the MFRSR AOD would require some discussion. Since the paper is already long we prefer not to add the AERONET AOD measurements.

3. Page 4, line 18: the reported angstrom exponent is high, it is about the maximum one; give a more realistic value (range).
During the first days of the campaign (17-20 June) the Ångström exponent is between 0.3 and 1.75. The text has been modified accordingly.

4. Page 5, lines 32-35: why? Please explain.
Figure 1 of the paper by Gröbner, J., Wacker, S., Vuilleumier, L., and Kämpfer, N.: Effective atmospheric boundary layer temperature from longwave radiation measurements, J. Geophys. Res., 114, D19116, doi:10.1029/2009JD012274, 2009 is very explicative. 99% of the downward LW irradiance comes from the lowest atmospheric layers, those where most of the water vapour is concentrated. The atmosphere is nearly transparent in the 8-14 µm spectral range, and the radiation in the interval is emitted from the upper layers.

5. Page 5, lines 35-37, "The pyrometer . . . for the IRP BT": this sentence is oversimplified. A quick look at the 3 figures reveals significant differences between BT and irradiances. For example, what happens in June 24 and 25 (when LW-WINDOW curves do not have peaks, opposite to IRP BT)? What about the role of temperature and clouds?
The pyrometer has a narrow field of view (2.6°), the broadband CGR4 has 180° and the CGR3 has 150° FOV. So the pyrometer is able to detect each single cloud entering its FOV, while the same single cloud can have a minor impact on the LW-WINDOW irradiance measured by radiometers with broad FOVs if the rest of the sky is cloud-free. That is the reason why the IRP BT time series has more peaks than irradiances time series.

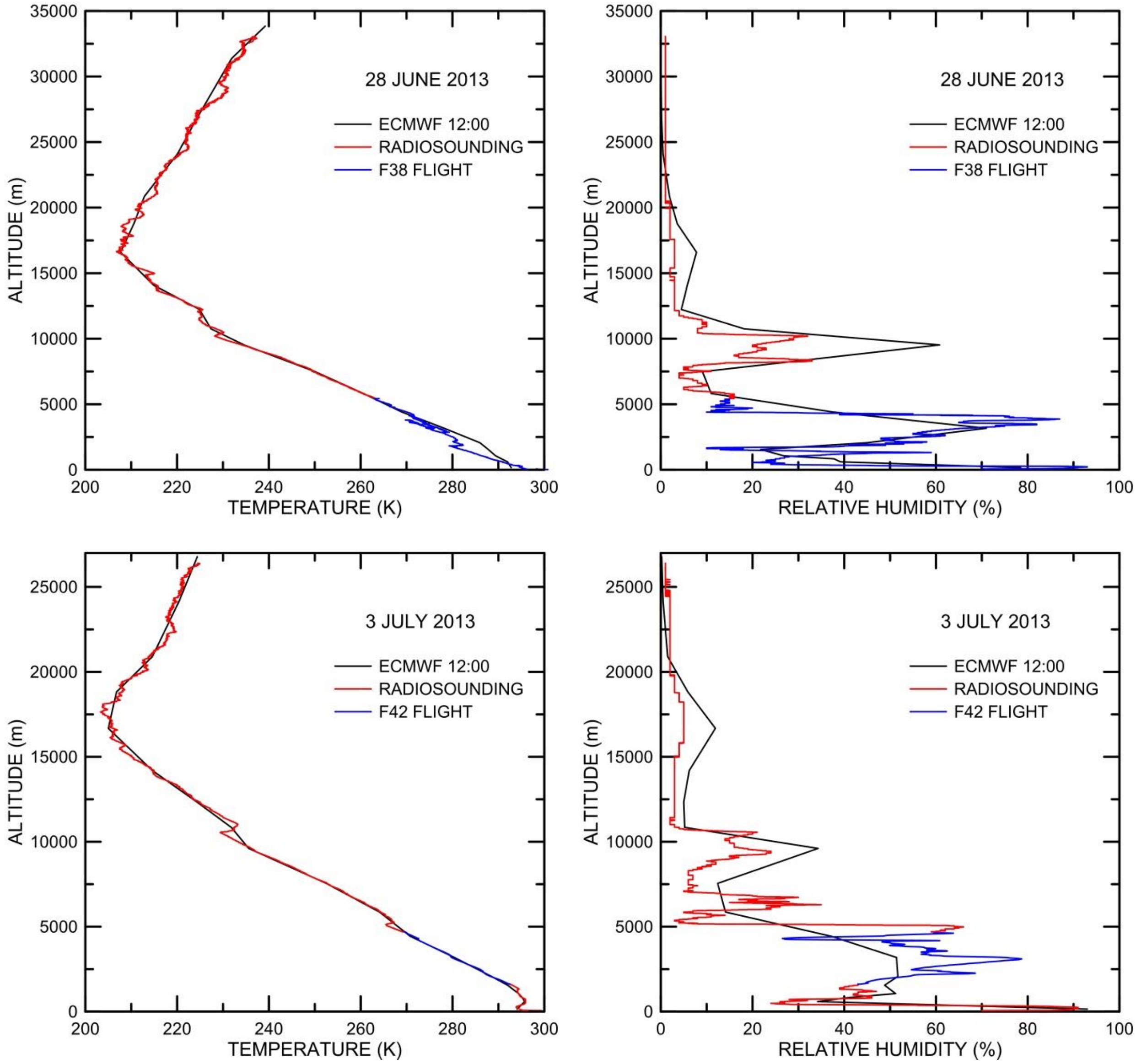
6. Page 7, line 7: define FWHM acronym.
The acronym has been defined in the revised text.

7. Page 7, line 30: up to which altitudes? How much the use of standard profiles can affect the radiative fluxes? Was any sensitivity study performed to assess this? Especially the LW fluxes should be sensitive.
The standard mid-latitude profiles have been used above the maximum altitude of the radiosounding on 28 June and 3 July, i.e. above 32 km and 26 km, respectively. The surface, as well as the profiles, irradiances and BT are not sensitive to variations of upper level atmospheric profiles.
In the revised manuscript a sensitivity study of the modelled quantities has been performed with respect to the main parameters affecting infrared radiation, either in aerosol-free conditions (i.e. IWV, temperature profile, SST, and surface emission) and with aerosol (we have tested the sensitivity to AOD and to the imaginary part of the dust refractive index).
An increase of 0.3 K in the temperature profile causes a 2.2 Wm^{-2} increase in downward LW irradiance at the surface, decreasing for increasing altitudes (becoming 1.3 Wm^{-2} at the ATR-42 top altitude of 5.7 km, and 0.6 Wm^{-2} at the Falcon 20 altitude of 10 km). The upward LW irradiance increases by 1.8 Wm^{-2} at 5.7 km, by 1.5 Wm^{-2} at 10 km, and by 1.4 Wm^{-2} at the TOA.

8. Page 7, line 33, regarding the absorbing gases: similarly, it would be worth to discuss/assess the sensitivity of fluxes to these parameters, especially given the scaling applied to their vertically distributed values.
The sensitivity of the irradiance to changes in the integrated water vapour has been tested by increasing the average measured value by its uncertainty (+0.2 mm), which is of the same order of magnitude of the IWV standard deviation within the considered time interval.
The downward LW irradiance increases by 0.4 Wm^{-2} at the surface, by 0.1 Wm^{-2} at 5.7 km, and by 0.06 Wm^{-2} at 10 km, while the upward LW irradiance at the TOA decreases by 0.1 Wm^{-2}.
The sensitivity to other absorbing aerosols has not been tested.

9. Page 8, about ECMWF: The use of reanalysis data is inevitable in this case. However, an assessment of the induced uncertainties associated with their coarser resolution could be made by comparing similar ECMWF data with available measurements for the other two days. This could provide an estimation of induced uncertainties in June 22.

As suggested by the reviewer, the ECMWF profiles at 12:00 have been compared with the temperature and relative humidity profiles measured by the radiosonde and by the airborne meteorological instruments. The T profile is well reproduced on both days, although the ECMWF profiles does not capture the fine vertical structures due its vertical resolution. An exception is the atmosphere below 4 km on 28 June, where the profile sounded by the ATR-42 presents lower temperatures compared to that of the ECMWF profile. The RH profile of the ECMWF operational model generally follows the measurements, but differences can be large at certain levels, like around 3 km on 3 July.



10. Page 8, line 10: a few words about the measured aerosol properties and the identified aerosol layers can be added. For example, apart from the layers and their extension neither information is given nor reference is made to the type of aerosols in each layer, with reference to corresponding measurements that cloud provide this kind of information.

A short discussion on the aerosol stratification identified by the airborne measurements and by the airmass back trajectory analyses has been included in Sections 3.1 and 3.1.1. A detailed analysis is presented in the paper by Denjean et al. (2016).

11. Page 8, line 14: so, what values of emissivity were assumed in the study? Do they differ and how much from day to day.
The values are the same for all the three cases because surface wind speeds are similar. A sentence has been added in the text to better explain, and some spectral values are included.

12. Page 8, Figure 4: the quality should be improved, e.g. by thicknenning the curves, so that the coloured curves can be more easily discerned.
The quality of the figure has been improved.

13. Page 8, line 29: As mentioned, different factors influence and differentiate the AERONET and in-situ SDs, one important being their different value, .e. columnar versus vertically resolved. The value of detailed measurements is that they provide vertically resolved SDs. Therefore, emphasis should be given to them. Discuss a bit more how the measured SDs differ to AERONET ones, referring to their agreement and disagreement. For example, larger differences appear in June 22 than in July 03. Refer to this difference referring to the nature of vertical profiles of Fig. 3 and the type of aerosols that are present in the different layers of every daily profile.
Table S1 has been produced as Supplement Material with the median radius, standard deviation, and normalized number concentration for each mode of either the AERONET and the *in situ* log-normal size distributions for the layers identified in the three cases. The differences in SD among the various layers have been discussed further, as suggested by the reviewer, and related to the transport pathways and to the mixing of dust with pollution particles.

14. Page 9, line 19: explain why this choice of water soluble RI was made and not any other.
Polluted maritime aerosol is the most probable aerosol type characterizing the lowest atmospheric layers over Lampedusa. This is supported by the analysis of the airmass back trajectories in the boundary layer, showing airmasses originating from Europe or recirculating within the Mediterranean basis, and from the chemical analyses carried out on the aerosol samples (Denjean et al., 2016). Water soluble is one of the components of the maritime aerosol, either clean and polluted, according to the OPAC definitions (see Hess et al., 1998). The other components are sea-salt and soot (on for the polluted maritime). Among the components, water soluble and sea salt have similar values of the imaginary part of the RIs below 8 µm and above 11 µm, while the water soluble is more absorbing than sea salt in the 8-11 µm interval. The RI of soot is too absorbing to be representative of the average aerosol. For these reasons the water soluble RI has been chosen.

15. Page 9, lines 20-26: Table 2 is not discussed enough. It should be said more clearly what exactly has been done and how the Mie-based computations of AOD compare to AERONET ones, whenever applicable, i.e. in visible wavelengths.
The combination of the SDs and RIs for the three flights has been better clarified in a new version of Table 2, which includes three tables (one for each flight), describing the SD and RI used in each aerosol layer identified by the ATR-42 and lidar vertical profiles. Moreover, the choice of the different RIs has been better explained in Section 3.1.1.
The spectral AOPs (extinction coefficient, absorption coefficient) accepted by MODTRAN are all referred to the extinction coefficient at 550 nm, for which the vertical distribution, $ext_{550}(z)$, is derived from the lidar backscatter profile and AOD measurements.
So the AOD at 550 nm is fixed, and corresponds to the value obtained from the MFRSR measurements at 500 and 868 nm using the Ångström law. AOPs are allowed to differ in no more than four aerosol layers in the troposphere.
We compute the spectral AOPs from Mie theory for a single particle in a wavelength grid from 2 to 100 µm and including 550 nm, and then divide them by the calculated extinction coefficient at 550 nm. So we have:

$$ext_{\lambda}(z) = \frac{ext_{\lambda}(z)}{ext_{550}(z)}$$
$$abs_{\lambda}(z) = \frac{abs_{\lambda}(z)}{ext_{550}(z)}$$

This ensures that the values of the AOD at 550 nm remain constant, whatever the AOPs.

16. Page 10, lines 16-17: does this refer to July 03? In Table 3 no results for INSU2 are displayed.
The sentence originally referred to 22 June. Section 4.1 has been modified in order to be clearer for the reader.

17. Page 10, lines 19-20: why the stronger infrared emission? Is it a matter of larger mass? Please explain.
For a particle with diameter D the absorption (and emission according to Kirchhoff law) and scattering properties are calculated with the Mie theory if D is of the same order of magnitude of the wavelength, as is the case for dust particles in the infrared spectral region. According to Mie theory, the absorption and scattering efficiency ($Q_{abs,ext}$) depend on the complex refractive index and on D and the absorption and scattering coefficient ($\sigma_{abs,ext}$) are proportional to Q and to the particle's cross section πD^2. Thus increasing the particles' dimension increases the absorption and emission coefficient.

18. Page 10, line 24: clarify that "all cases" refer to LW, WINDOW and IR BT.
"all cases" refers to the three days.

19. Page 10, line 34: here it should be clarified what is exactly the spectral interval/coverage of the measurements (IRP). This not clear based on what is said in page 7, line 6, about the IRP centered at 3 wavelengths etc. It is essential to clarify what is exactly the spectral coverage of measurements since they are used as the reference to which the simulations are compared, and given the significant sensitivity of theoretical computations to the spectral interval. Also explain why the reduction in WINDOW irradiances has different magnitude despite the same spectral reduction (0.4 microns) in different spectral parts.
The infrared pyrometer (IRP) measures the zenith BT in the 9.6-11.5 μm interval, while the CLIMAT measures the nadir BT in three infrared channels centred at 8.7, 10.6, and 12 μm with about 1 μm full width at half maximum. The sentence on page 10, line 34, refers to the WINDOW irradiance, measured by the PMOD/WRC CGR3 modified pyrgeometer, which is sensitive to the radiation in the 8-14 μm spectral interval.
The sensitivity analysis carried out on the WINDOW irradiance shows that reducing the spectral interval by a small amount (0.4 μm) significantly reduces the downward irradiance, and the reduction depends on where in the spectrum the reduction is operated: this is caused by the asymmetry of the irradiance spectrum with respect to the centre of the interval.

20. Section 4.1: what is missing is a critical approach providing insight into possible physical reasons for better agreement between the 7 examined cases. A quite exhaustive and very detailed description of results is made, referring to various numbers (Table 3). This is not enough while it turns to be confusing to the reader. What is more important is to determine which set of AOPs is more efficient and compared better to the measurements for the 3 cases. The discussion should conclude on this, stating at least if there is a "best" choice or if there is not and why. Moreover, in both cases, the discussion should provide a physical basis for the outcome of the analysis and the closure of Table 3. For example, a summary of the results of Table 3 should point to NOAER being the most efficient simulation, providing better results than the other 6 sets of AOPs in 4 cases (out of totally 9, i.e. 3 days by 3 parameters). NOAER is followed by COL1 (3 cases with best performance) and INSU3 (2 cases). So, questions may arise, like why simulations without aerosols

should be more appropriate/realistic, or why INSU3, which may be expected to be the most realistic, is finally not.
Section 4.1 has been modified to better present the results and the conclusions. We state that for the conditions met during the campaign, with moderate and not really large AOD, the surface irradiances are not useful to reduce the uncertainty on the dust RI, since the impact of different RIs is below the measurement and model uncertainty. This also explains why NOAER simulations agree with measurements within their respective uncertainties. On the contrary, narrowband zenith BT seems to be suitable to constrain the dust RI which better represents the dust optical properties either in moderate and in low dust loading conditions, assuming that the *in situ* airborne measurements better describes the local aerosol distribution. Indeed, in the two cases where the atmospheric meteorological profiles and the *in situ* SD are measured down to surface level (22 and 28 June) the zenith BT is well reproduced with the DB2017 RIs.

21. Page 12, line 3: as to upward LW, authors may want to comment on why the smallest differences are for COL1 in Table 5, while the smallest RMSDs in Table 4 are for INSU1.
Tables 4 and 5 show that the upward LW irradiance at Falcon 20 and ATR-42 altitudes is reproduced within model and measurement uncertainties with all AOPs, including the NOAER case. The fact that COL1 AOPs give the best match with Falcon 20 measurements while INSU1 AOPs provide the best agreement with the ATR-42 ones may be attributed to different reasons: among them, the Falcon 20 passage is not exactly simultaneous with the ATR-42; the Falcon 20 simulations may be affected by the ECMWF temperature/humidity profile above 6 km, while the ATR-42 upward irradiance rely on the *in situ* measurements of the meteorological vertical profiles.

22. Figure 7: why only points for NOAER, COL1 and INSU3 are given and not for the other cases? All these appear in Table 6.
Overlapping the points corresponding to all AOPs would have made the Figure 7 very difficult to read. After the reviewer's suggestion, we have considered that presenting the results of COL3, instead of COL1, may be better to show the dust perturbation compared to NOAER and the effect of changing the SD, but not the RI, compared to INSU3. The estimated uncertainties on the calculated BTs are also shown.

23. Page 12, lines 15-16: add "in-situ" before SD. This sentence needs to be re-written, since it is introduced all suddenly without being given evidence and discussed based on the results of Fig. 7.
The sentence has been modified as follows: "These results show the better sensitivity of BT to dust optical properties than broadband irradiance. When considering that *in situ* SD better represents the local aerosol distribution, the DB2017 RI from Algeria and Morocco provide the best model-measurement agreement".

24. Page 12, lines 17-19: while discussion is made no results are shown/given.
This comment is not clear. Figure 6 for the upward LW irradiance and Figure 7 for the CLIMAT BTs show that model simulations with and without aerosols are overlapped below 4 km, and differ at higher altitudes, where the aerosol effect is discernible. Moreover, the differences in BT due to dust compared to the aerosol-free simulations are provided on page 12, lines 19-20.

25. Section 4.2.1: A quite exhaustive discussion is made making frequent reference to numbers that differ a while between the 6 examined cases. Also the question arises why NOAER sometimes performs equally or better than dust-including cases. It could point to potential artifacts due to counteracting effects of other parameters than aerosol, which affect the LW radiation transfer and BT.

The limited aerosol effect is due to the moderate AOD measured during the flights and not to modelling problems, and this is the reason why in some cases (more often for the irradiance but not for the BT) the NOAER simulations agree with measurements. This aspect has been remarked in the revised manuscript (Sections 4.1, 4.2.1, and in the conclusions).

26. Page 12, lines 35-36, “These differences . . . airborne instrumentations”: so is there an inherent problem with the model?
We cannot state that there is problem with the model, because the same input parameters allow to fairly reproduce the CLIMAT BTs. Moreover, the LW irradiance is well reproduced below 2000 m and above 4000 m and at the Falcon20 altitude. As discussed, a possible bias in the CGR4 measurements may be found when the air temperature is fastly changing, due to inhomogeneities in the instrument temperature. For example the ATR-42 path (Figure 4b) shows a steep aircraft fall from 3.5 km to 1.2 km. A rapid decrease in altitude, with a consequent increase in temperature, may be not registered by the pyrgeometer, which has 6 seconds response time (1/e), but needs a longer time to establish equilibrium with the ambient air temperature (as also found by Albrecht et al., Pyrgeometer measurements from aircraft, Rev. Sci. Instrum., 45, 33–38, 1974). The same problem may have affected the irradiance measurements on 3 July. This may also explain the slight overestimation of the downward LW irradiance by the model in the same altitude ranges where the upward component is overestimated. The CLIMAT BT measurements is not expected to suffer from the same problem because the response time of the instrument is much faster (about 160 ms).

27. Page 14, line 20” add “was” before “evaluated”.
Done.

28. Page 14, line 21, “resampled”: how it was done?
IASI spectra have been averaged in 15 cm^{-1} intervals. The first interval is centred at 652.5 cm^{-1} and includes the BT values between 645.0 and 660.0, the second interval is centred at 667.5 cm^{-1}, and so on.

29. Page 14, lines 32-37: why there is difference on what provides the best match with reference to best match with the measured spectra and BTs?
The IASI spectra used in the comparison with the model are averaged over a region, instead of instantaneous measurements like those from CLIMAT. Moreover, IASI and CLIMAT measurements are not simultaneous. Finally, measurements are made over different spectral intervals, and the AOPs producing the best agreement with the mode result may somewhat differ. This has been stated in the text.

30. Page 15, lines 2-3: this is not applicable to 780-980/cm for June 22 and 28.
The sentence states that when the AOD is sufficiently large the aerosol perturbation to the BT simulated in aerosol-free conditions is significant, so this does not apply for the 28 June case. About 22 June, this applies clearly to the 1070-1200 cm^{-1} interval for different AOPs, and only for the COL1 combination in the 780-980 cm^{-1} interval.

31. Page 15, lines 20-22, “In our case, . . . AOPs equivalent.”: what exactly is it meant by this? By which means. Please explain. Is it implied that this (having very high resolution) is preferable? If so, why? If valid, it would mean that AOPs are not important for accurately computing LW radiation and dust LW radiative effects. Is this the meaning?
The sentence refers to the fact that MODTRAN resolution (0.1 cm^{-1}) is lower than that of the IASI measurements and of the σ-IASI-as radiative transfer model (0.01 cm^{-1}), so we do not expect model-measurements differences as low as those found by Liuzzi et al. (2017). The sentence has been eliminated and text has been modified, also to answer to Reviewer #1, as follows: “The

limitation in the MODTRAN5 resolution does not allow to reproduce the high resolution IASI spectra: however, the scope of simulating the IASI measurements is to show that TOA BTs are sensitive to the dust presence and to the various AOPs, and that they can be reproduced with the same input parameters that allow to simulate irradiance and BT at the surface and in the atmosphere.".

32. Page 15, line 28, "The combination . . . downward": this is not clearly evidenced in the discussion of sections 4.1 and 4.2.
The sentence has been moved at the end of Section 4.2.

33. Page 26, Table 2: the Table needs further/better explanation, it is not very easy for the reader to understand what exactly is the information given in this Table.
We agree with the reviewer. A new Table 2 has been prepared, which includes three tables, one for each day. The tables present the combination of SD and RI used in each layer identified by the lidar and ATR-42 measurements, and the AOD value at 8.6 µm.

34. Page 37, Figure 5: what have been the criteria for the design of flight paths.? Nothing is said about this and deserves to be mentioned in the text.
This aspect has been treated in Section 2.1(Aircraft strategy) of the paper by Denjean et al. (2016). In particular the following text explains the flight strategy: "The general flight strategy consisted of two main parts. First, profiles from 300m up to 6 km above sea level (a.s.l.) were conducted by performing a spiral trajectory 10–20 km wide to sound the vertical structure of the atmosphere and identify interesting dust layers. Afterwards, the identified dust layers were probed by straight levelled runs (SLRs), where the aircraft flew at fixed altitudes, to provide information on dust spatial variability and properties. Horizontal flight legs in the dust layers lasted 20–40 min to allow aerosol collection on filters for chemical analyses in the laboratory".
The sentence "The ATR-42 sounded the atmosphere during profile descents and ascents to infer the vertical structure and composition and identify layers with different properties" in Section 2.2 clarifies the criteria beyond the flight paths.

35. Page 44, Figure 12: wavelengths could be added, e.g. on the top x-axis.
The top x-axis has been expressed in wavelength units, as suggest by the reviewer.