

We would like to thank both referees for their time and effort to evaluate the manuscript and for their insightful comments, which surely help make the manuscript even better.

We have gathered in a single document the full comments of both reviewers, and our answers. The former are reproduced in black plain text, while the latter are in blue italic police. Except when stated otherwise, the line numbers in our answers are those from the revised manuscript (without tracked changes). We have not included here the modified text from the manuscript as it would make this long document even much longer.

Referee 1

This manuscript presents a comprehensive and valuable update of the MAPIR algorithm (v5.1) for dust aerosol retrieval from IASI observations. The work is technically solid, the dataset is highly relevant for the community, and the effort to build a long-term, consistent TIR dust product is particularly commendable. The paper also demonstrates a strong level of expertise and provides a thorough description of the methodology and evaluation.

However, despite these clear strengths, I have several major concerns that should be addressed before publication. In my view, some of these issues are substantial and currently limit the robustness and interpretability of the results. In particular, questions related to the treatment of spectral noise, the handling of known issues in the IASI-C dataset, and the validation strategy need to be clarified and/or strengthened.

For these reasons, I recommend **major revisions**. I believe that addressing the comments below would significantly strengthen the manuscript and make it a strong and impactful contribution to the field.

Major comments:

- **IASI-C - different pre-processing**
 - This appears to be a blocking issue. Publishing a dataset affected by such a bug raises concerns. I would strongly recommend either resolving this issue before publication or restricting the study to IASI-A.

We agree with the reviewer that any processing difference hinders the merging possibilities to create long data sets. However, there is no bug here. The IASI available data (spectra / level 1 and atmospheric data / level 2) is different in its format, by essence, because EUMETSAT has not reprocessed the IASI-C data. However, as detailed in section 2.1.1, from these different data formats we extract the same information, which aims towards best possible consistency, at least until a new version of the PWLR3 or PCS is implemented. There was a new version of the PCS on 30 March 2023, which represents a “non-solvable” (currently) discontinuity in the time series, however with no detectable impact as was discussed in the manuscript section 3.8. The preprocessing detailed in section 2.2.1 is also different because different quality flags exist in the different IASI data sets, and therefore the filtering had to be adapted although aiming at maximum similarity.

It is rather possible that the reviewer meant the issue about the missing land surface altitude in some of the IASI-C spectra, discussed in section 2.3. This is indeed an unfortunate and

unexpected issue in the input IASI data from EUMETSAT, and is not linked to the MAPIR pre-processing. As stated in the manuscript, the correction is indeed easy (use an external digital elevation model) and will be implemented for any future full processing while it is already considered in the NRT algorithm. However, this can not be considered a bug (not in MAPIR anyway), since it affects the number of spectra that are successfully processed, but it does not hinder the quality of the obtained results. A full reprocessing of the IASI-C time series, although technically simple, represents a non-negligible financial cost which is currently not available. Furthermore, the data has been publicly available for more than 2 years now (both from our website and from Copernicus Climate Data Store, as stated in the Data Availability section), and therefore its full description and evaluation is needed for both IASI-A and IASI-C even though the second misses some useful data points. The description in the form of an ATBD (algorithm theoretical basis document) is already available with the Copernicus data set, and some level 3 validation was done under Copernicus, both available in the Copernicus data set documentation among the documentation of all the aerosol data sets (current direct links to the documents are: http://dast.copernicus-climate.eu/documents/satellite-aerosol-properties/C3S2_312a_Lot2_FDDP-AER/C3S2_312a_Lot2_D-WP2-FDDP-AER_202311_ATBD_AER_Annex_F_LMD_v3.0_final2.pdf and http://dast.copernicus-climate.eu/documents/satellite-aerosol-properties/C3S2_312a_Lot2_FDDP-AER/C3S2_312a_Lot2_D-WP2-FDDP-AER_202402_PQAR_AER_v3.2.pdf - should these not work any more, then it is best to use the data DOI referenced in the manuscript, and go to the section documentation then search the desired files).

With all that in mind, we hope that the reviewer can meet us in the thought that the current full data set (IASI-A and IASI-C) needs proper description and evaluation, including currently known issues, to inform users in the best possible way.

- **Length and readability**

- The paper is very long and highly information-dense. Several sections could be streamlined to improve readability without losing scientific content. For example, L305–315 could be simplified, and L317–319 may not be necessary. Overall, reducing overly technical detail in some parts would make the manuscript more accessible.

We are aware that this is a very long manuscript with some very technical parts, and made further efforts to remove / rephrase where relevant, aiming at being more concise and clear. The technical parts are necessary to allow reproducing the retrieval, which is a requirement in such a publication.

Lines 305-315 AMTD version, now lines 309-319: answered in the comment on validation and in specific comments; the contents can not be removed, especially since the reviewer him/herself asks for more information and clarification

Lines 317-319 AMTD version, now lines 321-322 have slightly been shorted; however, the contents are useful and would probably have been requested if we had not included it.

Sections 2.2.3 and 2.2.4 were improved using bold and italic to highlight some words, separating different concepts treated in these relatively long and technical sections.

Introduction to section 3 was reduced to its first paragraph; the rest is discussed in the further sections. Section 3.2 has been shortened and improved.

And many small changes were done across the manuscript.

- **Treatment of spectral noise (L345–350)**

- The artificial reduction of spectral noise does not increase the information content of the retrieval. The information content is governed by the Jacobians (i.e. sensitivity of the signal to the state variables), not by the noise itself. What is actually modified here is the balance between observation and prior in the cost function. While this may improve convergence, it is not physically justified. The spectral noise should reflect the true instrument characteristics, not be tuned to fit the retrieval. Please clarify how the values used in v5.1 were determined.

There is no artificial reduction in the spectral noise: even in MAPIR v5.1 the values used are larger than the “true” expected spectral noise (for which 2 references are given lines 216-17, with different reported values). In OEM retrievals, the spectral noise is often (if not always) “inflated” to larger values than the true spectral noise. In this way, “errors/uncertainties” that can not be properly modelled (e.g. in the radiative transfer model, in surface emissivity, ...) can be taken into account by the minimisation process. As those uncertainties are not known, the spectral noise to be used in the OEM is determined empirically as the lowest value (minimal value being the “true” spectral noise) allowing convergence for a sufficient fraction of retrievals. We are not aware of a more deterministic way to define the OEM spectral noise.

The difference between MAPIR v4.1 and MAPIR v5.1 is that the noise inflation was reduced: the OEM spectral noise is still larger than the “true” noise, but is lower in MAPIR v5.1 than it was in MAPIR v4.1.

Reducing the spectral noise in the OEM does increase the information content of the retrieval (in our work, we use the number of degrees of freedom to quantify this) because it modifies the constraint applied during the retrieval. The number of degrees of freedom (DOFs) is the trace of the averaging kernel matrix, which is itself the product of the Jacobians and the Gain matrix. The latter contains the constraints due to a priori information and spectral noise (see e.g. Rodgers 2000, section 2.4.2). Therefore, modifying the constraints in one of these affects the number of degrees of freedom of the retrieval. Reducing the OEM spectral noise does indeed also give more weight to observations with respect to a priori, in the cost function. In the retrieval results this effect was observed with a reduction of the “background AOD” (the retrieved low AOD values where it should be almost 0) together with the increase in DOFs.

Sections 2.2.2 and 2.2.4 have been modified to better explain this and avoid any possible confusion about the use of the spectral noise in MAPIR.

- **Vertical profile validation**

- It is stated that another manuscript is under preparation. I would then suggest publishing these 2 papers together, as companion papers. Section 3.6

(profiles) would be better placed in that companion paper rather than included here.

Section 2.6 and the quantitative profile validation will be moved to the second manuscript. The latter is almost ready, however we would like to finalise the current publication even if the profile validation manuscript is not yet submitted, avoiding a long time before the revised version becomes available. We propose to keep a very short qualitative profile demonstration section in the current manuscript, with the CALIOP comparisons that were already shown, and some test-case comparisons with EARLINET. This is not a validation, but a short demonstration of the profiling capability of MAPIR, while the validation of those profiles will be in the second manuscript.

- **Dust AOD Validation**

- The validation effort is appreciated, especially given the known challenges of TIR aerosol validation. It should be stressed here that these results should indeed be considered qualitatively, as a proper comparison is not directly possible. However, it is clear that the a priori of 2 particles cm⁻³ leads indeed to an overestimation for low AOD. How do these results compare to the previous version of MAPIR and to other dust retrievals from IASI (e.g. DLR IMARS, LMD, ULB)? Given that this paper focuses on algorithm improvements, such benchmarking is essential.

We thank the reviewer for her/his appreciation of the challenges in this exercise. We stress that we do not call it “validation” but “evaluation” due to this qualitative nature of the comparisons, even when statistics are provided. We indeed mention a number of times in the manuscript that the evaluation results are to be taken with caution, e.g. beginning of section 3, end of section 3.2, lines 567-569, and in the conclusion.

The lowest bound of the a priori does indeed lead to overestimation for low AOD. It was however necessary to ensure that, in the case of events where usually there is no dust (e.g. central or North Europe), the retrieval is not overconstrained to very low dust AOD and fails or modifies mostly the surface temperature (therefore missing the dust). This is explained in detail in lines 309 to 319.

The retrieval method itself is clearly compared to its previous version in the manuscript. Comparing the data to the previous data set in detail would be quite strenuous. Almost all input data was updated to reprocessed or more recent data sets, therefore the comparison of the previous MAPIR v4.1 data set with the current MAPIR v5.1 data set would include differences both linked to retrieval changes and input data improvement. In addition, the time and geographic coverage are different between versions. That would mean that the manuscript would need to contain two evaluation sections, one with data that match the MAPIR v4.1 coverage, and one for the full-time coverage and global data. For the sake of readability, we had intentionally left this comparison between versions out.

Comparing the MAPIR data to the other IASI algorithms would surely provide a lot of useful information on strengths and weaknesses of each algorithm. This is however a huge task and out of the scope of the current manuscript. Large intercomparison exercises typically represent a paper by themselves, with an external expert as leading author. Some information/comparisons are available in the Copernicus dataset documentation / quality

assessment report (link shared above in the response to the IASI-C comment). These, however, relate to level 3 data.

Minor comments:

- Ensure all acronyms are consistently defined (e.g. OEM, AOD appear multiple times). Consider using a glossary package (e.g. glossaries in LaTeX).

We carefully re-checked all abbreviations and removed the full definition where it was repeated.

- Add DOIs and clickable links for all references where possible.

Indeed, there were 2 articles and 2 reports for which the DOI was missing. For most technical documents there is no DOI but a reference that is specific to the provider; this reference has been added with the document title. Clickable links other than DOIs often expire after some time and are less useful. All links were also rechecked to ensure they are clickable.

- Figures are often too small and difficult to read:
 - Increase font sizes (comparable to main text)
 - Harmonise figure styles
 - Use more descriptive, self-contained titles (e.g. include instrument names directly in titles)

The size of figures 1 to 9, 11, 17 to 20 (revised manuscript numbers) was increased (with different factors depending on the figure). The captions of figures 2 and 3 were shortened. Figures 4 and 5 received a left title. Figure 6 received a new colour scale following a specific comment. Figures 7 and 8 received different markers for day and night following a specific comment. Figure 10 received an additional “super title” and colour bar legend. Figure 18 received an addition “super title”. Figures 12 to 14 text was uniformised in size.

Specific comments

- L26 remove extra punctuation → “wind speed (e.g. Choobari et al., 2014)” -> *Done*
- L41 best -> highest -> *Done*
- L113 Why mention the OEM here? -> *Removed; the precision was indeed not needed (the EUMETSAT retrieval for Ts, T and H2 profiles contains 2 steps, the first being the PWLR3 and the second an OEM, for clear-sky observations only)*
- L126 I would clarify early that the setup described is for the climate setup -> *This was already mentioned in the introduction, lines 61-63*
- L126-132 Clarify the reasons behind the different setups. Also, what does it mean “when the OE doesn’t perform well”? How is this evaluated? -> *The reason is simply the data availability: in NRT, only one set of L2 data is distributed by EUMETSAT, which is either the OEM (when performed and passed the QC) or the PWLR3. The sentence was clarified in the manuscript.*
- L139 spherical assumption - please mention already here that this assumption is valid in the TIR but less in the VIS (and keep reference to section 3.2) -> *Done*

- L143 high-resolution, what does it mean exactly? -> *60m vertical resolution, same as CALIOP original data; added in the manuscript*
- L146 The resolution of MAPIR retrieval vertical grid should be specified here to ease readability -> *done*
- L163 Does it make sense to use a fixed concentration for all other gases? How much gas absorption is there in the TIR? What are the default concentrations in RTTOV? What is the impact of assuming the surface Lambertian in the TIR? -> *The MAPIR dust retrieval spectral windows lie in the so-called atmospheric window, where there is very little gas absorption; the dominant is water vapour which is accounted for since we are using the IASI level 2 data. Other absorbing gases in those retrieval windows would be CO₂ (only small lines), O₃ (but not the main TIR absorption band), SO₂, HNO₃, CFC-11 and CFC-12. These are all very small absorption contributions in the used spectral windows (see for example in Clerbaux et al. 2009 - reference added in the manuscript). Therefore, their variability is not critical to reproduce, while to include those gases in the retrieval would be computationally costly for some, impossible with RTTOV for others. Using external data sets for all these gases would require finding consistent data throughout the time series, including for NRT exploitation, which would represent a huge work for very limited expected benefit. The default RTTOV concentrations are defined in the RTTOV science and validation report for version 12 (and unchanged in version 13); reference has been added in the manuscript. The Lambertian assumption is common in the TIR spectral range, stating that both emission and reflection are isotropic, and is mandatory in RTTOV when using the DOM for aerosol scattering modelling. We apologize for not making this clear in the original manuscript. This is now corrected. The impact of this assumption has not been analysed, as it would require using a different RT model.*
- L173 The “negligible effect” should be demonstrated -> *In the test, the difference in AOD at 10µm was below 0.01 for u and v components of 20m/s (wind speed of 28m/s or about 100km/h) -> this has been added in a short note in the manuscript.*
- L174 snow fraction set to 0 -> Is it a valid assumption? In which cases? -> *The CAMEL emissivity (monthly) atlas already contains snow where its presence is usual. When the snow fraction is set to 0, the grid pixel is treated as snow-free. If the atlas emissivity includes snow, the algorithm searches for an alternative snow-free emissivity value for that pixel. When no snow-free value is available, the original atlas emissivity is retained, as occurs in regions where snow persists throughout the month. For non-zero snow fractions, the treatment depends on whether the atlas emissivity already contains a snow component. When such a component is present, the atlas emissivity is retained. Otherwise, the atlas emissivity is combined with a fixed snow emissivity spectrum, using the prescribed snow fraction as the weighting factor. Sensitivity tests showed that using a snow fraction of 50 substantially increases the surface emissivity, relative to the snow-free case, in regions and periods where snow would not normally be expected. In contrast, the emissivity remains largely unchanged in regions where snow is expected. Therefore, this assumption (snow = 0) would be invalid only for a dust cloud over an occasional snow cover. A short explanation has been added in the manuscript, avoiding too much technical detail.*
- L210 no need to repeat the OEM acronym -> *done as part as all acronyms checks*

- L226 - 232 this paragraph might be technical and not necessarily needed in an already quite long scientific paper -> *Indeed this is very technical, however it is an important part of the retrieval setup, needed to allow reproducing the retrieval; therefore it is necessary to include it in the paper*
- L246 PSD not defined -> *Done (added at first use, in section 2.1.2)*
- L255 how is this uncertainty reported then? -> *It is not reported in the files, only evaluated with the methodology described in section 3.7.*
- L272: clarify → “not prefiltered in MAPIR v5.1” -> *the clouds are not filtered during the pre-processing. This part of the sentence has been removed, and we repeated the information that the cloud flag is not used for pre-processing, lines 262-263*
- L278 “decent” quality is vague -> *good performances; we agree that it remains vague, however we do not wish to reproduce here the PWLR3 validation; reference has been added.*
- L282 is there some validation of this filter? -> *We agree that an independent validation of the proposed filter would be desirable. However, such a validation is currently not straightforward, as there is no existing dust/cloud reference filter or dataset that is sufficiently reliable for direct use in the context of our analysis. This is precisely why a dedicated filter was introduced. The filter was developed empirically, through visual inspection of a limited set of representative cases and expert judgement regarding the separation between dust-affected scenes and thin-cloud contamination. We have therefore clarified in the manuscript that the filter should not be regarded as a fully validated classification product, but rather as a pragmatic screening approach designed to reduce evident misclassifications. While imperfect, it represents the most appropriate solution currently available for the specific needs of this study.*
- L284 Provide a reference or an expanded explanation -> *To our knowledge, this specific explanation has not been formally reported in literature. Nevertheless, it is readily apparent in the analysis of intense dust events, particularly when the IASI cloud fraction or cloud flag is compared with true-colour imagery from visible sensors. We have added a brief explanation of this to the manuscript.*
- L289 was shown - where? -> *wrong choice of word, it is indeed not “shown” but comes from analysis of the MAPIR v4.1 data during the set-up of MAPIR v5.1. Shown was replaced by “discovered”.*
- L292 wrongly categorizing intense dust events as clouds - provide reference -> *see answer to the comment on L284*
- L296-298 The very small impact of reducing streams seems surprising → needs justification -> *The reason is that scattering is a minor contribution in the dust radiative impact in the TIR (dominated by absorption/emission). A sentence was added in the manuscript.*
- L305-315 unclear → needs clarification or simplification -> *This paragraph is indeed very technical. It is however important because it explains the “background” dust AOD retrieved in absence of dust. We have clarified the paragraph as much as we could without removing necessary content.*
- L335 -337 If not quantified, consider removing -> *removed*
- Figure 2 is too small, hard to read -> *size increased in the pdf, hopefully this is acceptable for the type setting*
- L362 - 368 Only ~35% of retrievals pass final QC → is processing cloudy scenes justified? *We agree that the final number of valid retrievals is relatively low. However,*

processing scenes before applying the final cloud-related quality control is necessary for the objectives of this product. As discussed in Section 2.2.4, a cloud pre-filter would remove a substantial fraction of the strongest dust events, as well as weaker events occurring in regions where dust is less expected. Excluding these scenes before the retrieval would bias long-term analyses, affect source identification, and limit the usefulness of near-real-time data in assimilation, forecast, and alert systems. This approach is computationally costly, and we agree that a reliable pre-retrieval cloud filter would be highly beneficial. However, such a filter would need to distinguish clouds from dust without systematically rejecting important dust cases. We currently do not have a robust solution that meets this requirement.

- L371 This shall be explained before the publication of this paper. At least, it would be nice to see that some investigation has been performed and possible factors excluded -> *The EUMETSAT user notification service has been investigated again and a plausible explanation has been published. This has been added in the manuscript, alongside with the explanation for the discontinuity in IASI-C in 2022.*
- L372 Again, this is a blocking issue for publication -> *Please see our detailed response to Major Comment 1. The IASI-C differences are not due to a MAPIR bug, but to differences in the available EUMETSAT input data formats and quality flags. The processing was adapted to extract the same physical information as consistently as possible for IASI-A and IASI-C. The missing land-surface altitude in some IASI-C spectra is an input-data issue and affects data coverage, not the quality of the retained retrievals. These points have been clarified in the manuscript.*
- Figure 3 is too small -> *size increased in the pdf, hopefully this is acceptable for the type setting*
- L533 If it really only slightly changes the statistics, why do you remove the worst 3%? Especially if this 3% includes high values of AOD, the statistics (both correlation and RMSE) might change significantly. I would refrain from such a filter, especially given that the data entering the comparison is already pre-filtered. -> *As mentioned in the document the elimination of the 3% outliers is only altering slightly the statistics. We provide hereunder, for comparison, the same as Figures 4a and 6 but without outlier removal, and Table 5 including values with and without outlier removal. Only the correlation coefficient differs, which is expected due to the different noise level in the comparisons. It has to be noted that the 3% outliers do not correspond to the highest AOD values, but to the highest AOD differences. The decision for the exclusion of these 3% has to do with the spatiotemporal comparison limitations, not only with the data quality itself (which is indeed already pre-filtered, although no filter is ever perfect). Satellite-based and point measurement comparison of any variable are directly linked with spatiotemporal comparison thresholds versus the variability of the parameter under investigation inside the threshold limits. Here we assume that aerosols are homogeneous within 25km and 15 minutes from the AERONET measurements. This may however not be strict enough and still cause outliers in comparisons due to (1) a high spatio-temporal inhomogeneity in dust distributions, especially for severe dust cases; (2) the different optical path for sun-looking instruments and satellite nadir instruments even though both provide vertical columns at the end; (3) the cloud structures within 25km and 15 minutes, which may be captured differently by these different optical paths. Therefore, based on the analysis we decided to exclude 3% of data as outliers, justified by the above reasoning. A*

similar approach was done for MAPIR v4.1 (Callewaert et al, 2019), and also by Capelle (2018), but indeed not in Clarisse (2019).

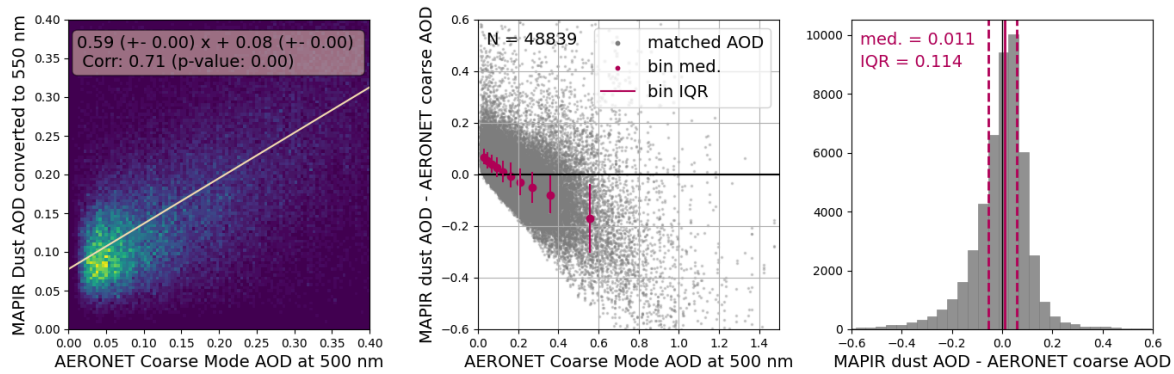


Figure 1: Same as the manuscript paper Figure 4a, but without the 3% outlier removal: MAPIR v5.1 IASI-A 550nm converted AOD against AERONET SDA CAOD 500nm.

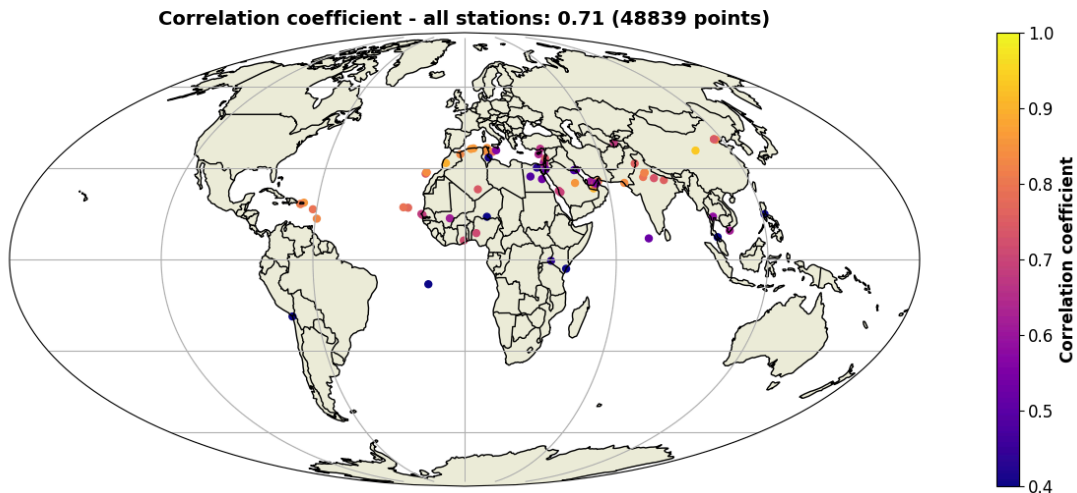


Figure 2: Same as the manuscript paper Figure 6, but without the 3% outlier removal: Correlation coefficient between MAPIR v5.1 IASI-A 550nm converted AOD and AERONET SDA CAOD 500nm, for all stations with at least 100 collocations

Table 1: Same as the manuscript Table 5: statistics of MAPIR v5.1 AOD converted to 550nm versus AERONET SDA coarse mode CAOD with 3% worse data as outlier removal (values in the manuscript) / without outlier removal. Note that the difference in the number of points is not exactly 3% because the removal is done by station, and 3% is not always precisely an integer number of points.

<i>param</i>	<i>A - solar</i>	<i>C - solar</i>	<i>A - lunar</i>	<i>C - lunar</i>
<i>N points</i>	<i>47 310 / 48 839</i>	<i>1 024 / 1 053</i>	<i>12 167 / 12 579</i>	<i>959 / 995</i>
<i>Corr coeff</i>	<i>0.78 / 0.71</i>	<i>0.67 / 0.56</i>	<i>0.79 / 0.72</i>	<i>0.7 / 0.45</i>
<i>Mean abs bias (std dev)</i>	<i>-0.002 (0.11) / -0.005 (0.14)</i>	<i>-0.011 (0.18) / -0.02 (0.23)</i>	<i>0.002 (0.09) / 0.000 (0.11)</i>	<i>-0.001 (0.09) / 0.001 (0.16)</i>
<i>Median ab bias</i>	<i>0.012 / 0.011</i>	<i>-0.01 / -0.013</i>	<i>0.011 / 0.01</i>	<i>0.008 / 0.007</i>
<i>Median rel. bias</i>	<i>9% / 8.4%</i>	<i>-5% / -5.6%</i>	<i>10% / 9%</i>	<i>6.2% / 5.8%</i>
<i>Comp with bias lower than 0.1</i>	<i>76% / 73%</i>	<i>58% / 56%</i>	<i>81% / 78%</i>	<i>78% / 75%</i>

- Figure 7 day/night distinction unclear → use different markers -> *Done, also for Figure 8 to keep consistency*
- Redundancy between Figures 7 and 8. Also, from Figure 7, it looks like there should be some colocation for IASI-A. Clarify why IASI-A colocations seem to be missing. -> *Lines 607-608 state that “For Ragged Point, there are not enough lunar colocations for IASI-A, therefore only IASI-C is shown”; there are more solar co-locations but the goal here was to show also day/night consistency. Figure 7 shows all data points, while Figure 8 shows only co-locations. Showing both allows the reader to examine the sampling of the full data that is used for the comparisons.*
- L741 Could Jacobians be computed numerically (perturbation approach)? Shouldn't sensitivity also depend on other parameters (e.g. gas concentrations, AOD magnitude)? -> *The perturbation approach is indeed what was done with using a surface emissivity of 0.95 times the original value. The impact of the surface emissivity on the retrieved AOD does depend on the AOD, which was already mentioned in the manuscript in the next sentence, together with other parameters. The impact of the gas concentration on the dust retrievals with different surface emissivity was not analysed but is not expected to be significant, as the spectral windows used bear only very little gas absorption. Using a numerical approach to generate emissivity Jacobians and then use them in the OEM would be an option to investigate; however for computational time sakes it would be necessary to avoid calculating those Jacobians for each retrieval. This is therefore huge work. The better evaluation of the surface emissivity uncertainty and its impact on the dust retrieval is a good topic for a future study, and a sentence was added in the conclusions.*
- L790: Could reduced IASI-C coverage be explained by erroneous cloud filtering? -> *No cloud filtering is done before the retrieval; this reduced IASI-C coverage is due to*

missing the surface altitude in some pixels, for which there is no explanation. Please refer to the answer to the first major comment for relevant details.

- L793: If discussed later, avoid mentioning prematurely -> *sentence removed*
- L794–797: Clarify reference to Figure 18 -> *reference to Figure 18 (old manuscript, now Figure 17) was a reference to the way the data was handled, to avoid repetition. The reference to the figure was removed and the sentence was modified to contain necessary information.*

Referee 2

This study describes version 5.1 of the MAPIR algorithm for retrieving dust concentration profiles and column-integrated AOD at 10 μm from IASI observations onboard the Metop-A and Metop-C satellites. Key updates from version 4.1 include extending the vertical retrieval range to 10 km, removing pre-filters to maximize data coverage, adopting updated surface emissivity and radiative transfer models, reducing spectral noise inflation, and propagating temperature and humidity uncertainties. The algorithm is evaluated against AERONET, CALIOP, and EARLINET. Long-term stability and IASI-A versus IASI-C consistency are also assessed. Overall, the paper is comprehensive, transparent about limitations, and delivers a scientifically valuable long-term data product. However, there are still some major concerns needed to be addressed before considering acceptance.

Separation of retrieval quality from the TIR-to-VIS conversion uncertainty. In sect. 3.3, the authors compare MAPIR converted dust AOD at 550 nm against AERONET SDA coarse-mode AOD at 500 nm, using a fixed conversion factor of 1.78. The authors mentioned that the AOD-dependent bias could be a constant bias if a different conversion factor (2.85) were used. This means that the central validation metric conflates two distinct sources of uncertainty: the quality of the 10 μm retrieval itself and the accuracy of the wavelength conversion, which depends on assumed particle size and refractive index (Capelle et al., 2018; Clarisse et al., 2019; Zheng et al., 2023, 2026).

As the authors acknowledge that direct reference to AOD at 10 μm is currently unavailable, it is suggested to provide a sensitivity analysis in which the comparison statistics (correlation, slope, bias, same as Figures 4 and 5) for a range of plausible conversion factors (e.g., 0.9 to 3.5, as quoted from Clarisse et al., 2019), showing how the statistics change. In addition, it is suggested to acknowledge the necessity of constraining dust microphysical properties (e.g., size, shape, refractive index) to reduce the uncertainty contributed by the conversion factor, as previous studies mentioned (Capelle et al., 2018; Clarisse et al., 2019; Zheng et al., 2023, 2026).

This is indeed a central issue in general for the validation of TIR-retrieved AOD, and why we here use the word “evaluation” and not “validation”. The conversion is simply a multiplication by the factor, therefore performing the validation with different conversion factors leads to exactly the same correlation coefficient, while the slope and single point bias will simply be multiplied by the ratio between the different conversion factors. We consider it interesting to discuss which conversion factor would make the dependency with the AOD disappear, but showing additional plots would not provide useful additional information.

The dust microphysical properties are constrained, as described in section 2.1.2 and the same particle size distribution and refractive index are used to compute the conversion factor as in the retrieval. Using different sizes / refractive indices would affect both the retrieval and the conversion factor, and MAPIR does not currently contain any sort of size / composition retrieval, only fixed properties. This issue is under investigation for a future version of the algorithm, and represents very significant work.

The whole discussion in section 3.2 is aimed at showing that there are so large differences in the physical processes at play in TIR versus VIS, that obtaining a theoretically correct conversion factor does not currently seem feasible. Therefore, without reference data in TIR, one can only perform qualitative evaluations such as presented in this manuscript.

We thank the referee for pointing us to the work of Zheng et al, especially the most recent paper, which brings additional light to this whole conversion discussion. A reference to it was added in the manuscript in section 3.2 and also when discussing the AERONET validation in section 3.3.2.

It would also be informative to note which end of the conversion factor range corresponds to which particle sizes, such as "smaller particles produce higher conversion factors," which would help the reader build intuition.

This was done by Zheng et al, 2026, although with different sets of RI. We now refer to that paper to interpret the conversion factor, in section 3.3.2.

Quantitative profile validation needs strengthening. The profile comparisons with CALIOP (Sect. 3.6.1, Figs. 12–14) are presented as qualitative, side-by-side visualizations without summary statistics. The EARLINET validation (Sect. 3.6.2) is limited to 12 dust events at one station (Limassol, Cyprus). Given that MAPIR's unique selling point relative to column-only TIR products is precisely the vertical profiling capability, a more rigorous profile validation would significantly strengthen the paper.

Respectfully, we somewhat disagree with the reviewer that the unique selling point of MAPIR is the profiling capability. The MAPIR AOD product itself is also a worthy product, and the validation run under the Copernicus program for the climate data store compares the different IASI dust algorithms, finding that each has strengths and weaknesses. The profiling capability is an additional feature of MAPIR, like the mean altitude and selection of aerosol model (among two) are the additional features of Capelle et al.

Therefore, for the CALIOP comparisons, it is suggested to add layer-by-layer statistics (mean bias and correlation) for the matched orbit segments, even if the comparison is approximate due to different retrieved quantities (extinction vs. concentration). If possible, for the EARLINET comparisons, the authors could also consider expanding the analysis beyond Limassol to other EARLINET stations in the dust belt. Otherwise, it would be better to provide the mean and standard deviation of the MAPIR–lidar difference as a function of altitude for all 12 cases, with and without the averaging kernel correction, to provide a quantitative sense of the systematic vertical bias.

Although the authors acknowledge that a dedicated profile validation paper is in preparation, a brief quantitative summary here would improve the current manuscript substantially.

The whole quantitative profile validation will be moved to the second manuscript, following a recommendation from the other reviewer. The paper will contain much more comparison cases, from different EARLINET stations. We propose to keep a very short qualitative profile demonstration section in the current manuscript, with the CALIOP comparisons that were already shown, and some test-case comparisons with EARLINET. This is not a validation, but a short demonstration of the profiling capability of MAPIR, while the validation of those profiles will be in the second manuscript.

The a priori positive bias floor and its implications for climatological averages. The increased minimum a priori concentration of 2 particles cm^{-3} across the entire profile (Sect. 2.1.3) creates a retrieval floor of approximately 0.06 at 10 μm AOD (~ 0.14 at 550 nm), which the authors acknowledge (lines 558–565 and 835–836). While this is understandable from a retrieval stability perspective, the consequences for users constructing climatologies, computing regional means, or performing trend analyses are significant. In dust-free regions, the product would always report non-zero dust AOD, inflating global or regional averages. I suggest the authors consider whether a bias correction or an additional quality flag (e.g., a flag indicating that the retrieval has not moved meaningfully from the a priori) could be included in future data versions.

We do understand the issue caused for regional means or long-term averages. However, the previous setup used to miss events where the a priori was very low, because it could not deviate from it. Our assessment is that it is better to have a background overestimation (that is a known feature and could be removed by users) than to miss important events.

The current recommendation to users is to consider AOD up to 0.06 as non-significant and a sentence was added in the conclusion to emphasize that point and how users should handle it. We are not convinced that bias correction is the correct option: this is not a bias on all data, as it applies only for the low AOD cases with low sensitivity. We acknowledge the suggestion of including a specific flag for cases where the retrieved profile is close to the a priori. However, it would mostly flag low dust AOD cases. Users might get confused by the flag and simply remove all these data points from analysis, which is also not necessarily the correct interpretation as it would also remove the information that there is no or very little dust.

The MAPIR retrieval uses a single set of dust optical properties: a log-normal size distribution with 0.6 μm mean radius, a refractive index from GEISA–HITRAN, and spherical particles (Sect. 2.1.2). However, previous studies have shown that there are significant global variations of dust size distribution (Formenti and Di Biagio, 2024) and dust TIR refractive index (Di Biagio et al., 2017). It is suggested to add discussions on whether the use of regionally varying refractive indices (as done by the Di Biagio et al. dust CRI database) could improve the retrieval, at least as a future development. If such sensitivity tests have already been performed for earlier MAPIR versions, referencing those results here would suffice.

We fully agree with the reviewer that this is important future development to be considered, both for the retrieval and for the conversion to VIS. Some work was done while designing this new version of MAPIR, using the refractive indices from Di Biagio et al. (2017) and different particle sizes. However, the work could not be finalised for use in an “operational”

production line with the intention to build long consistent time series (and also operate in NRT). The reasons were the computational time constraints (can't again increase significantly), the necessity to design a retrieval of the best RI / size (because running multiple retrievals and selecting the best result is not an option, due to computing time), the need to deal with horizontal continuity in the optical properties (to avoid discontinuities in the retrieved concentration / AOD). It was decided that this improvement would take too long to implement and that the already made improvements justified a new algorithm version bringing better data to users even if there is certainly still room for improvement.

Although assuming dust to be spherical has been demonstrated to have similar optical properties in TIR as in non-spherical shapes, the authors also calculate the conversion ratio to visible based on the spherical assumption. The impact of dust non-sphericity is not negligible in visible (Huang et al., 2020; Saito et al., 2021). A discussion of whether the spherical assumption introduces systematic biases in the retrieval and in the conversion factor calculation is needed.

The spherical assumption impact is already shortly discussed line 480-484 (with complementary references). No significant bias is expected in the retrieval itself, and the non-sphericity "uncertainty/bias" is most certainly lower than uncertainty/bias arising from uncertainty in the other particle optical properties.

Providing a more quantitative assessment of the potential bias the spherical assumption brings in the conversion would require to define a non-spherical dust model and compute 550nm extinction from that model. This is significant work, and would indeed probably improve the trust in the conversion factor. However, the conversion factor contains many more uncertainties than the spherical assumption (starting with the PSD and RI) and those can not be solved at the moment. Therefore it does not seem relevant to yet quantify precisely the impact of the spherical assumption.

Minor comments:

Abstract, line 5: The abstract mentions retrieval of "10 μm AOD and mean altitude" but does not mention that the evaluation is performed using the 550 nm converted AOD. Since most readers will use the 550 nm product, briefly noting the conversion and its associated uncertainty in the abstract would set appropriate expectations.

Done

Sect. 2.1.5, lines 170–178: The surface emissivity bias in CAMEL v2 over the Sahara during summer is noted but not quantified in terms of its impact on AOD retrieval. Since the surface emissivity uncertainty analysis in Sect. 3.7 shows up to 0.05 AOD uncertainty; the authors should discuss whether this uncertainty is dominated by the CAMEL bias or by intrinsic emissivity variability. If the former, would using an alternative emissivity product (e.g., MODIS-based monthly emissivity) reduce this uncertainty?

We are not aware of a CAMEL v2 bias over the Sahara during the summer, and it is not discussed in our manuscript. Lines 326-328 mentions such a bias but it refers to the

previously used surface emissivity from Zhou (2011). We realize that the sentence could lead to confusion and modified it.

Sect. 2.2.4, lines 285–300: The reduction from 8 to 4 DOM streams is validated with a maximum AOD difference of 0.015. It would be helpful to state the conditions under which this test was performed (e.g., dust AOD range, viewing geometry). At large viewing angles or high optical depths, the sensitivity to stream number may increase.

This reduction is not validated; the introduced bias was deemed acceptable with respect to other errors/uncertainties, considering the huge gain in computation time. The test was performed on 20 June 2020 (Godzilla extreme event) for global data. The provided maximum AOD difference is for the large AOD and viewing angles during that day.

Sect. 3.3.3, Figure 6: The colorbar range (0.1 to 1.0) is too smooth to tell the variation. Since most stations have $R > 0.6$, consider using a narrower range (e.g., 0.4 to 1.0) or a colormap with limited intervals to better distinguish good from mediocre stations.

Done

Sect. 3.9, IASI-A vs. IASI-C: The discussion of the Metop-A orbital drift (lines 793–809) is important, and the recommendation to switch to IASI-C in October 2019 is well justified. Consider providing a concrete prescription in the data usage guidelines, e.g., “For long-term analyses, use IASI-A from July 2007 to September 2019 and IASI-C from October 2019 onward.” Users would benefit from this explicit guidance rather than having to piece it together from the discussion.

We have added the specific month of change (October 2019) in the corresponding section of the conclusion. The time range of both instruments was already mentioned in the second paragraph of the conclusion. We keep this in mind for any other user documentation than this paper, together with the recommendation regarding AOD below 0.06.

References:

Capelle, V., Chédin, A., Pondrom, M., Crevoisier, C., Armante, R., Crepeau, L., and Scott, N. A.: Infrared dust aerosol optical depth retrieved daily from IASI and comparison with AERONET over the period 2007–2016, *Remote Sens. Environ.*, 206, 15–32, <https://doi.org/10.1016/j.rse.2017.12.008>, 2018.

Clarisse, L., Clerbaux, C., Franco, B., Hadji-Lazaro, J., Whitburn, S., Kopp, A. K., Hurtmans, D., and Coheur, P.-F.: A decadal data set of global atmospheric dust retrieved from IASI satellite measurements, *J. Geophys. Res. Atmos.*, 124, 1618–1647, <https://doi.org/10.1029/2018JD029701>, 2019.

Di Biagio, C., Formenti, P., Balkanski, Y., et al.: Global scale variability of the mineral dust long-wave refractive index: a new dataset of in situ measurements for climate modeling and remote sensing, *Atmos. Chem. Phys.*, 17, 1901–1929, <https://doi.org/10.5194/acp-17-1901-2017>, 2017.

Huang, Y., Kok, J. F., Kandler, K., Lindqvist, H., Nousiainen, T., Sakai, T., Adebisi, A., and Jokinen, O.: Climate Models and Remote Sensing Retrievals Neglect Substantial Desert Dust Asphericity, *Geophys Res Lett*, 47, e2019GL086592, <https://doi.org/10.1029/2019GL086592>, 2020.

Saito, M., Yang, P., Ding, J., and Liu, X.: A Comprehensive Database of the Optical Properties of Irregular Aerosol Particles for Radiative Transfer Simulations, *J Atmos Sci*, 78, 2089-2111, [10.1175/jas-d-20-0338.1](https://doi.org/10.1175/jas-d-20-0338.1), 2021.

Zheng, J., Zhang, Z., Yu, H., Garnier, A., Song, Q., Wang, C., Di Biagio, C., Kok, J. F., Derimian, Y., and Ryder, C.: Thermal infrared dust optical depth and coarse-mode effective diameter over oceans retrieved from collocated MODIS and CALIOP observations, *Atmos. Chem. Phys.*, 23, 8271–8304, <https://doi.org/10.5194/acp-23-8271-2023>, 2023.

Zheng, J., Yu, H., Zhou, Y., Shi, Y., Zhang, Z., Di Biagio, C., Formenti, P., and Smirnov, A.: A novel retrieval of global dust optical depth and effective diameter based on MODIS thermal infrared observations, *Remote Sens. Environ.*, 332, 115083, <https://doi.org/10.1016/j.rse.2025.115083>, 2026.