Response to interactive comments from Referee #1

The referee is thanked for the careful reading of and constructive comments to the manuscript. The referee's comments are repeated below in italic font. The responses to the comments are shown in roman font.

Comments

- *Abstract: the quantitative discussion here only really covers the IASI data sets. A sentence or two about GOME-2 and SCIAMACHY performance could be added as well.*

  We have added the following text to the Abstract to include GOME-2 and SCIAMACHY performance. "For the solar sensors it is found that SCIAMACHY data on the average are lower by -1.097 km (-0.961 km) compared to the CALIOP geometric mean (cumulative extinction) height and GOME-2 lower by -1.393 km (-0.818 km)."

- *Table 1: I suggest changing the institute column to cover both the institute and the algorithm name, as it is algorithm names (e.g. MAPIR, IMARS) which are often referred to in other publications and may be most memorable.*

  The algorithm names have been included in tables 1, 2 and 3 as suggested.

- *P5L9: the MAPIR data presented are stated to be from two different versions, v3.2 and 3.4. However, it is hard to tell how significant the differences are. From the text it seems to only affect the a priori temperatures used in the retrieval, but it is not clear if this is the only difference. This should be clarified in the text. If the difference in retrievals from the change to the prior is negligible, then this should be stated and the sentence reworded to avoid potential reader confusion/concern. My suggestion, both for simplicity and consistency (both internal and with any future available data set) would be to just reprocess all the data with the latest version of the algorithm (v3.4).*

  Indeed, the only difference between the two versions is the a priori surface temperature. The reason for this change is that for periods in time when the IASI level 2 data from EUMETSAT is in version 4, the retrieved surface temperature is often completely off, especially over deserts, therefore being a bad a priori for our retrievals. In those cases, the ECMWF ERA-interim reanalysis makes a better a priori. For later versions of the EUMETSAT IASI level 2 product, the retrieved surface temperature makes a better a priori for MAPIR than ERA-interim (closer to truth, and also easier to use). The change between EUMETSAT IASI level 2 version 4 to 5 occurred on 14 september 2010, so it concerns the 4th period in the current analysis. We have tested that using the a priori surface temperature from ECMWF ERA-interim or from EUMETSAT IASI level 2 version 5 is indeed negligible.

  As there is now a publication available that contains the full algorithm description, where we have used the MAPIR version name 3.5 to join together versions 3.2

and 3.4 for simplicity, we have changed this here too. The discussion of the Ts a priori has been removed from here, as it is not required for this manuscript, has little impact on the results, and may now be found in the other manuscript which reference is provided.

- *P7, general: some of this information can be removed for readability, since there is already a reference for the algorithm in table 1 and some of this information is probably not directly necessary for the interpretation of the results here. For example I dont think the section from L7-L18 about the a priori can be shortened or removed, unless the authors feel this is significantly different from what is described in the prior algorithm paper or is somehow crucial for the understanding.*

  Indeed we agreee that there was too much information in that section. It was due to the fact that the algorithm used here was very significantly different from the previously published one and not yet described elsewhere. We have now simplified the MAPIR description, only maintaining a general description and the parts which are useful for the discussion.

- *Section 2.2.4: this algorithm description can also be shortened, as Table 1 provides a reference and some details for the interested reader, and much of the information provided is not necessary for the interpretation of the results. For example P10L2-7, P10L18-22.*

  The section has been shortened as suggested by the referee.

- *P13L10-24: from this it seems like the movement did not make a large systematic change to the dust height (0.02 km or less). Since the standard deviation of the height change was 0.25 km, this also implies that the error which would result from not attempting to account for the time difference is small, and that the altitude of dust layers is fairly stable over the time period of (as I understand) up to 5 hours. So that in itself is an interesting result as it suggests temporal sampling errors are not a big problem for this type of comparison, at least when looking at the big picture (since 0.25 km is somewhat smaller than the other retrieval uncertainties). The other side of the coin, which I think should be discussed more, is whether FLEXTRAs assessment of the transport during this time period is accurate. I would like to see some quantification on the reliability of FLEXTRA here as I assume it depends in part on the resolution and quality of the meteorology data ingested, since this is what will determine the horizontal and vertical motion from the trajectory. So for example which ECMWF data version was used? Has FLEXTRA been quantitatively evaluated? If you use another meteorological data set or change some other input parameter is the solution stable? I know that e.g. HYSPLIT lets you run an ensemble of trajectories as one way to assess stability; I dont know if FLEXTRA does this too. The bottom line is that the comparability between CALIPSO and other sensors rests on the use of FLEXTRA to assess and correct for the change in aerosol location between observations, and this is not covered in much detail in the present version of the manuscript, so more discussion is needed.*

To address the referee's comments we have added the following to the FLEXTRA description: "Quantification of trajectory errors is always difficult, due to a general lack of ground-truth data. However, FLEXTRA has been quantitatively evaluated in the past. Comparisons of FLEXTRA trajectories driven with ECMWF data with balloon trajectories have revealed typical horizontal transport errors of about 20% of the travel distance, but with large variability from case to case (Baumann and Stohl, 1997; Stohl and Koffi, 1998; Riddle et al., 2006). Evaluation against meteorological tracers such as potential vorticity suggests errors of a similar magnitude (Stohl and Seibert, 1998). Thanks to improvements in the meteorological analysis data, slightly smaller errors may be assumed for more recent years, but the order of magnitude of the errors is likely still similar.".

Information about ECMWF data version and temporal and spatial resolution has been added.

- *Figures 2-7: I understand the intention here, but find these figures hard to interpret. In the top panel the grey is hard to see unless you zoom a lot, and the blue lines are somewhat similar to tones in the colour table. For the bottom panels, the background lidar curtain is the same between the IASI figures but there are too many different coloured symbols overlaid to quickly and easily compare and get the picture of what is going on. I wonder if, for the IASI plots at least, these could be redrawn. For example, make the top panels of Figures 2-5 their own figure, since these are all the same date and same map bounds, so we can directly compare the maps of coverage and heights between the different algorithms for IASI. Then also make the bottom panels their own figure and simplify in some way for clarity. For example we probably dont need to see the lidar curtain (although this could be added as a panel by itself at the top, or as a separate figure) since what is being compared is the effective heights, and the curtain just adds noise when the eye is trying to see coloured symbols. The deep pink pluses and red crosses appear similar in tone, again taking the lidar curtain off would allow one to use a more contrasting color for one such as blue or black. The points for CALIPSO AOD could also be removed since this is in black and naturally the most eye-catching, yet it is the least relevant since it is one set of points which does not represent a height. (In my mind the think you want the reader to focus on should probably be black or red since these stick out.) Alternatively the CALIPSO AOD could be just overlaid on the lidar curtain. The point here is that when looking at figures like this I want to answer two questions: (1) how similar are the spatial coverage and distribution of heights from the data sets and (2) how do they compare to CALIPSO? The present Figures do not allow me to do this effectively.*

  To make the figures more readable, we have split the figures as suggested and changed the colours of the markers and included legends. The text as been updated accordingly.

- *I realise that Figures 6 and 7 are a different date so cant be combined with the IASI figures for my suggested redrawing. Is there no date with all instruments providing data? If so that would be better to show. Since both have much sparser coverage than the IASI plots, though (both these Figures are essentially a lot of white space*

and then overlapping symbols which are hard to distinguish without zooming), unless a better common case can be found, perhaps these two examples could be removed. I understand that from parity you probably want to show one case from each algorithm, but surely there must be a more instructive case than this; there is so little data that its hard to assess looking at it whether these data sets are reasonable or not. And the summary statistics are in Table 3, so if the data are always sparse I dont know that we need to see GOME/SCIAMACHY case studies.

We have omitted these figures as suggested. The text have been updated accordingly.

- *Table 3: as with Table 1, I'd add algorithm name here in the column headers. I'd also add instrument name, for ease of reference. I also have some formatting questions/comments here. In general this table is not well organised because the values in the headers only seem to refer to some of the statistics (I guess the top four of the eight boxes in each subset?) and there are too many colours. I think this could be much simplified, and the clarity improved, by condensing things and labelling individual rows rather than relying on colour-coding and large column headers. If I interpret the legend right, the top left and top right numbers in each column are the mean and standard deviation of height difference. So this could be represented as one entry, e.g. 0.590 +/- 1.213 km for BIRA-IASB. And the row legend would just read "height difference, km". Then youd have other rows for number of points and "inlay". Labelling rows (which would then repeat for the all points, day/night, land/ocean splits) would remove the need for the complicated colour coding, and the header would only say instrument and institute/algorithm name. The statistics for cumulative vs. geometric heights could be split as left-right subcolumns for each algorithm instead. Does this make sense? Changing to that layout would make each row/columns content unambiguous, and dis- pense with the need to have 10 different colours (plus white) to code the table. As it stands, again, it is very difficult to pick out the key numbers from the table.*

  We agree with the referee that the table was unnecessarily complicated to read and appreciate the suggestions for changes. They have been fully adopted in the revised manuscript.

- *Figure 8 (and later): if these are scatter density plots on the first two rows then shouldnt they be shown as filled rectangles rather than clouds of points? There is also no colour scale to indicate what is shown (absolute or relative frequency, how many points are we looking at)? I suggest redrawing. For the bottom panels, it would be clearer if a vertical line for a height difference of 0 km is added. This will again aid in direct comparisons of the data sets. I would take off the Normal distribution fits; they dont add anything that we cant already see from the bins, and dont appear to be great fits in some cases anyway (the actual distributions seem to have higher kurtosis than a Normal distribution, at least for IASI).*

  The two first rows in figure 8 (and later) show the probability density, using kernel density estimation. Two-dimensional histogram, which gives filled rectangles, does not reveal the features seen with the probability density. The caption of the figure has been changed to clarify what is presented. Also color scales have been added

in the two first rows, and in the bottom row the normal distribution fit has been removed and a vertical line added for the zero height difference, as suggested.

- *Figures 11, 12: I dont think these add anything to the discussion, and should be deleted. This is essentially another visualisation of the data already presented in Fig- ures 8-10 and Table 3, albeit also sliced by time. The lines are also too faint to see. The authors can just modify the discussion on P24, L12-16 to note that optical charac- teristics may have been different but no clear temporal variation was found when the data were examined. We dont need to see the plots.*

  The figures have been removed and the text modified accordingly as suggested.

- *P24L19-23: From the inlay columns it seems overall that this has values between 5.3% (LISA, ocean, night, cumulative) and 68.9% (LMD, land, night, cumulative), with typical values being something of order 30%. That means that typically two thirds or so of retrieved heights are entirely outside the dust layer (i.e. the retrievals put the dust somewhere totally without dust). Is that interpretation really correct? If so, that sounds pretty bad, and seems at odds with the other statistics presented, which show a standard deviation of around 1 km or so (and to me seems like a good result). I would double-check the calculation of this inlay statistic or reword the text if I have misunderstood what it means. The only way I can think of to reconcile this discrepancy is if the vertical extent of the CALIPSO dust layers is typically significantly smaller than the roughly 1 km IASI retrieval error. So perhaps some statistics about the CALIPSO dust layer geometric thickness should be presented here. Either way, there appears to be some unresolved issue in this statistic or its interpretation.*

  Unfortunately there was an error in the calculation of the "inlay" statistics. This error, which affected only the "inlay" statistics, has been corrected and revised results entered Table 3. The text has been updated accordingly. A table showing CALIOP cloud height and layer thickness statistics have been added together with an accompanying discussion in the text.

- *P27L14-15: Table 1 indicates that several of the algorithms use the Optimal Estimation Method (OEM) or similar techniques, which should be able to provide pixel-level uncertainty estimates on retrieved aerosol height. It would be instructive to compare these estimates in a statistical sense to the level of agreement with CALIPSO, to see whether these uncertainties are reasonable. This does not appear to have been done; I suggest adding it in the revised manuscript, since this analysis provides a useful way to validate the uncertainty estimates. I realise that this cant be done for all the data sets, but since this is a big advantage of OEM, it makes sense to use it! For the OEM retrievals, it should be there already. If it is not, why not? It is definitely naturally within the scope of the existing study.*

  Some of the algorithms use methods such as the OEM which in principle may be used to provide pixel-level uncertainties. However, these uncertainties include only those associated with the quantities included in the retrieval and are not fully representative because we are not able to compute the full uncertainty estimate.

To compute the full uncertainty estimate would require computing derivatives with respect to the parameters with significant impact on the retrieval. For the MAPIR algorithm for example, this would at least include derivatives of the temperature profile, surface emissivity, aerosols size distribution and refractive index. Within the current MAPIR framework this is not possible. For the other algorithms similar concerns are valid, and/or full error estimates have not been included at this stage of algorithm development. Uncertainty estimates of the retrievals are certainly of great interest, but also not trivial to implement. At the current algorithm development stage we are unfortunately not able to present this quantity.

Response to interactive comments from Referee #3

The referee is thanked for the careful reading of and constructive comments to the manuscript. The referee's comments are repeated below in italic font. The responses to the comments are shown in roman font.

Major comments

1. *I have the same feeling to the first reviewer that Figure 2-7 and their discussions are difficult to follow. I also suggest Figures 2-5 and 6-7 be merged, so that readers can easily compare the spatial distribution of heights between different algorithms. I would also recommend a legend be added to the curtain plot to indicate the meanings of each symbol.*

   Figures 2-7 have been changed as suggested by referee #1. Legends have been added to indicate the meanings of each symbol.

2. *Table 3 is particularly hard to follow. I recommend, instead of using table, use bar plots to compare those statistics to different algorithms.*

   We have clarified and cleaned up Table 3 as suggested by referee #1. Bar plots, as suggested, may be an alternative. However, we feel that the numbers themselves include more detailed information about the results and, as such, may have more value for possible future studies.

3. *Too much text is used to present Figures in the Appendix. Those figures should be briefly mentioned, so only the major findings from them be presented.*

   We have moved the mentioned text to the appendix and points the reader to the appendix for discussion and presentation of these results.

4. *This study found that solar algorithms yield larger bias (> 1 km) for the case of dust aerosol height than the IR algorithms. However, it should be noted that some studies have shown an accuracy of about 0.5 km of dust layer height from O2-A might (studies listed below). So authors may need to compare and justify the performance of the current study to those studies:*

   Kokhanovsky, A. A., and V. V. Rozanov (2010), The determination of dust cloud altitudes from a satellite using hyperspectral measurements in the gaseous absorption band, International Journal of Remote Sensing, 31(10), 2729-2744, doi:10.1080/01431160903085644.

   Dubuisson, P., R. Frouin, D. Dessailly, L. Dufor TM t, J.-F. . L on, K. Voss, and D. Antoine (2009), Estimating the altitude of aerosol plumes over the ocean from reflectance ratio measurements in the O2 A-band, Remote Sensing of Environment, 113(9), 1899-1911, doi:10.1016/j.rse.2009.04.018.

Xu, X., J. Wang, Y. Wang, J. Zeng, O. Torres, Y. Yang, A. Marshak, J. Reid, and S. Miller (2017), Passive remote sensing of altitude and optical depth of dust plumes using the oxygen A and B bands: First results from EPIC/DSCOVR at Lagrange-1 C2point, Geophysical Research Letters, 44(14), 7544-7554, doi:10.1002/2017GL073939.

We thank the referee for these references. In the Discussion section we have included these references and a discussion of the results, highlighting additional differences between the retrieval setups of this work and mentioned literature.

Specific comments:

- *P2, L11: I'd to bring an attention to a recent review article about passive remote sens- ing of aerosol height by Xu et al. 2018, which is worth to cite: Xiaoguang Xu, Jun Wang, Yi Wang and Alexander Kokhanovsky, Chapter 1 - Passive Remote Sensing of Aerosol Height, In Remote Sensing of Aerosols, Clouds, and Precipitation, Elsevier, 2018, Pages 1-22, ISBN 9780128104378, https://doi.org/10.1016/B978-0-12-810437- 8.00001-3*

  The paper has been cited in the introduction.

- *P3, L5-10: It mentioned here that these selected dust events are of Saharan origin, but the studied area are also frequently affected by dust emitted from Middle East, India, and Western China. Please be accurate.*

  We have rephrased the sentences on P3, L5-12, to also include dust emitted from Middle East, India, and Western China.

- *P14, Figure 2(bottom): Symbols are hard to follow. A legend may be added to indicate the meaning of each symbol.*

  This Figure and similar ones have been revised. Legends have been added.

- *P15, Table 2: I dont quite understand the bracketed numbers in the third and forth rows. Please clarify in the Table caption (or using table footnote, as the caption is already very long).*

  We have added footnotes explaining the bracketed numbers.

- *P21, Figure 8: A colorbar is needed for the density of the scatters (similarly in Figure 9-10). The definition of the density is also necessary in the figure caption.*

  Colorbars have been added to Figure 8 and similar Figures. The definition of the density have been added to the figure caption.

Response to interactive comments from Referee #3

The referee is thanked for the careful reading of and constructive comments to the manuscript. The referee's comments are repeated below in italic font. The responses to the comments are shown in roman font.

Major comments

1. *I have the same feeling to the first reviewer that Figure 2-7 and their discussions are difficult to follow. I also suggest Figures 2-5 and 6-7 be merged, so that readers can easily compare the spatial distribution of heights between different algorithms. I would also recommend a legend be added to the curtain plot to indicate the meanings of each symbol.*

   Figures 2-7 have been changed as suggested by referee #1. Legends have been added to indicate the meanings of each symbol.

2. *Table 3 is particularly hard to follow. I recommend, instead of using table, use bar plots to compare those statistics to different algorithms.*

   We have clarified and cleaned up Table 3 as suggested by referee #1. Bar plots, as suggested, may be an alternative. However, we feel that the numbers themselves include more detailed information about the results and, as such, may have more value for possible future studies.

3. *Too much text is used to present Figures in the Appendix. Those figures should be briefly mentioned, so only the major findings from them be presented.*

   We have moved the mentioned text to the appendix and points the reader to the appendix for discussion and presentation of these results.

4. *This study found that solar algorithms yield larger bias (> 1 km) for the case of dust aerosol height than the IR algorithms. However, it should be noted that some studies have shown an accuracy of about 0.5 km of dust layer height from O2-A might (studies listed below). So authors may need to compare and justify the performance of the current study to those studies:*

   Kokhanovsky, A. A., and V. V. Rozanov (2010), The determination of dust cloud altitudes from a satellite using hyperspectral measurements in the gaseous absorption band, International Journal of Remote Sensing, 31(10), 2729-2744, doi:10.1080/01431160903085644.

   Dubuisson, P., R. Frouin, D. Dessailly, L. Dufor TM t, J.-F. . L on, K. Voss, and D. Antoine (2009), Estimating the altitude of aerosol plumes over the ocean from reflectance ratio measurements in the O2 A-band, Remote Sensing of Environment, 113(9), 1899-1911, doi:10.1016/j.rse.2009.04.018.

Xu, X., J. Wang, Y. Wang, J. Zeng, O. Torres, Y. Yang, A. Marshak, J. Reid, and S. Miller (2017), Passive remote sensing of altitude and optical depth of dust plumes using the oxygen A and B bands: First results from EPIC/DSCOVR at Lagrange-1 C2point, Geophysical Research Letters, 44(14), 7544-7554, doi:10.1002/2017GL073939.

We thank the referee for these references. In the Discussion section we have included these references and a discussion of the results, highlighting additional differences between the retrieval setups of this work and mentioned literature.

Specific comments:

- *P2, L11: I'd to bring an attention to a recent review article about passive remote sens- ing of aerosol height by Xu et al. 2018, which is worth to cite: Xiaoguang Xu, Jun Wang, Yi Wang and Alexander Kokhanovsky, Chapter 1 - Passive Remote Sensing of Aerosol Height, In Remote Sensing of Aerosols, Clouds, and Precipitation, Elsevier, 2018, Pages 1-22, ISBN 9780128104378, https://doi.org/10.1016/B978-0-12-810437- 8.00001-3*

  The paper has been cited in the introduction.

- *P3, L5-10: It mentioned here that these selected dust events are of Saharan origin, but the studied area are also frequently affected by dust emitted from Middle East, India, and Western China. Please be accurate.*

  We have rephrased the sentences on P3, L5-12, to also include dust emitted from Middle East, India, and Western China.

- *P14, Figure 2(bottom): Symbols are hard to follow. A legend may be added to indicate the meaning of each symbol.*

  This Figure and similar ones have been revised. Legends have been added.

- *P15, Table 2: I dont quite understand the bracketed numbers in the third and forth rows. Please clarify in the Table caption (or using table footnote, as the caption is already very long).*

  We have added footnotes explaining the bracketed numbers.

- *P21, Figure 8: A colorbar is needed for the density of the scatters (similarly in Figure 9-10). The definition of the density is also necessary in the figure caption.*

  Colorbars have been added to Figure 8 and similar Figures. The definition of the density have been added to the figure caption.

Comparison of dust layer heights from active and passive satellite sensors

Arve Kylling[1], Sophie Vandenbussche[2], Virginie Capelle[3], Juan Cuesta[4], Lars Klüser[5], Luca Lelli[6], Thomas Popp[5], Kerstin Stebel[1], and Pepijn Veefkind[7,8]

[1]NILU - Norwegian Institute for Air Research, PO Box 100, 2027 Kjeller, Norway
[2]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
[3]Laboratoire de Météorologie Dynamique (LMD), UMR8539, CNRS/IPSL, Ecole Polytechnique, Palaiseau, France
[4]Laboratoire Interuniversitaire des Systémes Atmosphériques (LISA), CNRS UMR7583, Université Paris Est Créteil, Université Paris Diderot, Créteil, France
[5]Deutsches Zentrum für Luft-und Raumfahrt e. V. (DLR), Deutsches Fernerkundungsdatenzentrum (DFD), 82234 Oberpfaffenhofen, Germany
[6]Institute of Environmental Physics (IUP), University of Bremen, Bremen, Germany
[7]Royal Netherlands Meteorological Institute (KNMI), 3730 AE De Bilt, The Netherlands
[8]Geosciences and Remote Sensing, Delft University of Technology, 2628 AA Delft, The Netherlands

Correspondence to: Arve Kylling
(arve.kylling@nilu.no)

Abstract. Aerosol layer height is an essential parameter to understand the impact of aerosols on the climate system. As part of the European Space Agency Aerosol_cci project, aerosol layer height as derived from passive thermal and solar satellite sensors measurements, have been compared with aerosol layer heights estimated from CALIOP measurements. The Aerosol_cci project targeted dust type aerosol for this study. This ensures relatively unambiguous aerosol identification by the CALIOP processing chain. Dust layer height was estimated from thermal IASI measurements by four different algorithms (BIRA-IASB, DLR, LMD, LISA) and from solar GOME-2 (KNMI) and SCIAMACHY (IUP) measurements. Due to differences in overpass time of the various satellites, a trajectory model was used to move the CALIOP derived dust heights in space and time to the IASI, GOME-2 and SCIAMACHY dust height pixels. It is not possible to construct a unique dust layer height from the CALIOP data. Thus two CALIOP derived layer heights were used: the cumulative extinction height defined to be the height where the CALIOP extinction column is half of the total extinction column; and the geometric mean height which is defined as the geometrical mean of the top and bottom heights of the dust layer. In statistical average over all IASI data there is a general tendency to a positive bias of 0.5–0.8 km against CALIOP extinction-weighted height for three of the four algorithms assessed, while the fourth algorithm has almost no bias. When comparing to geometric mean height there is a shift by -0.5 km for all algorithms (getting close to zero for the three algorithms and turning negative for the fourth). The standard deviation of all algorithms is quite similar and ranges between 1.0 and ~~1.4~~1.3 km. When looking at different conditions (day, night, land, ocean) there is more detail in variabilities (e.g. all algorithms overestimate more at night than at day). For the solar sensors it is found that SCIAMACHY data on the average are lower by -1.097 km (-0.961 km) compared to the CALIOP geometric mean (cumulative extinction) height and GOME-2 lower by -1.393 km (-0.818 km).

1 Introduction

Aerosol is identified as an essential climate variable (ECV) by the Global Climate Observing System (GCOS, http://www.wmo.int/pages/prog/gcos/). The aerosol layer height (GCOS product A.10.3) is one of four aerosol parameters which is needed to enhance our understanding of the aerosols' role in the climate system. Furthermore a deeper insight is important for radiative budget analysis, study of chemical and physical interactions in the troposphere, weather forecast modelling, remote sensing and air quality initiatives. Ground-based methods (LIDARs) offer high accuracy and calibration benchmarks, however their geographical coverage is sparse. Hence, satellite observations of the aerosol layer height are warranted and the quality of such a product needs to be assessed. As part of the European Space Agency (ESA) Climate Change Initiative (CCI, Hollmann et al., 2013) the Aerosol_cci (Popp et al., 2016) project has conducted a comparison between dust type aerosol layer heights from passive and active sensors to identify strengths and possible weaknesses in the estimate of this parameter.

Both active and passive methods may be used to estimate the aerosol layer height. The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) onboard the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite (Winker et al., 2009, and https://www-calipso.larc.nasa.gov/) provides detailed vertical information with a vertical resolution of 30 m below 8.2 km and a horizontal footprint of 335 m. Passive solar and thermal infrared satellite instruments may provide global data on a daily basis with horizontal resolution on the order of tens of km. For example Vandenbussche et al. (2013) retrieved desert dust aerosol vertical profiles from Infrared Atmospheric Sounding Interferometer (IASI) measurements; Cuesta et al. (2015) described the three-dimensional distribution, including dust height, of a dust outbreak over East Asia also using IASI measurements; Sanders and de Haan (2013) used the O_2 A-band to retrieve aerosol layer height from the Global Ozone Monitoring Experiment-2A (GOME). Dust top height may also be estimated using stereo view techniques by either utilizing instruments with multi-angle capabilities (for example the Advanced Along Track Scanning Radiometer, AATSR, Virtanen et al., 2014) or by combining measurement from different sensors (see for example Merucci et al., 2016). For a comprehensive survey on the methodological approaches, technical and scientific challenges of the retrieval of aerosol height, we refer the reader to the recent review by Xu et al. (2018) .

The aim of this work is to assess the different aerosol layer height products from different algorithms for various solar (SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY-SCIAMACHY, GOME-2) and thermal (IASI) sensors by comparison with CALIOP. The Aerosol_cci project targeted dust type aerosol for this study. The relatively unambiguous classification of dust by CALIOP and the availability of large dust events possibly avoid any biases due to aerosol misclassfication in the aerosol height comparison.

Earlier studies (Capelle et al., 2014; Peyridieu et al., 2010, 2013) have compared monthly averaged and gridded data, or a specific episode (Vandenbussche et al., 2013; Cuesta et al., 2015). We perform a point by point comparison for selected episodes. Furthermore we account for differences in satellite overpass time by the use of trajectory model analysis. Finally, for the first time, utilizing data from GOME-2, SCIAMACHY and IASI with their respective spectral and spatial resolutions, results from the different passive infrared and solar algorithms are compared for the same dust episodes to identify strengths

and weaknesses. Note that this work focuses on the comparison of dust layer heights retrieved from active and passive sensors. Comparisons of aerosol dust amount are outside the scope of this work and is discussed elsewhere (Popp et al., 2016).

The remainder of the paper is organized as follows: In Sect. 2 the data and data analysis methods are presented. The results from the aerosol layer height comparison are given in Sect. 3. The results are discussed in Sect. 4 and followed by the conclusions.

2 Data and methodology

To allow inclusion of data from the SCIAMACHY instrument that ceased operation in 2012, four desert dust events ~~of Saharan origin~~ in 2010 were selected (total 40 days):

- 18-27 March (10 days)
- 22 May - 1 June (11 days)
- 1-12 July (12 days)
- 14-20 September (7 days)

~~Furthermore we choose to focus the comparison~~ The comparison focus on the region between 0-40°N and 80°W-120°E, see ~~also upper plots of Figs. ??-??~~Fig. 2, and is mainly affected by dust from Sahara, but is also influenced by dust from the Middle East, India and Western China.

2.1 Active instrument dust height retrievals - CALIOP

CALIPSO is the fourth of the six satellites in the A-train satellite constellation. All six of the A-Train satellites cross the equator within a few minutes of one another at around 13:30 local time. CALIOP is part of the payload of the CALIPSO platform. The CALIOP laser produces simultaneous coaligned pulses at 532 and 1064 nm that are used to measure the backscatter profile. The 532 nm pulse is linearly polarized. The return signal is polarized parallel and perpendicular to the outgoing plane and detected by two photomultiplier detectors. The CALIOP aerosol typing algorithm uses layer-averaged depolarization and the 532 nm attenuated backscatter to classify the aerosol into one of six types: clean marine, dust, polluted continental, clean continental, polluted dust, smoke (Omar et al., 2009). Of the six types the dust aerosol is largely nonspherical implying a relatively large depolarization ratio and hence relatively unambiguous classification. Numerous data products are available from CALIOP. We use the 5 km profile product from CALIOP data version V4-10. Only dust profiles with CALIOP cloud aerosol discrimination (CAD) values between -100 and -20 are included (Winker et al., 2013). Profiles containing polluted dust and water and ice clouds are excluded. Furthermore we only include dust layers that are continuous, ~~that is~~ thus, multi-layered dust clouds are excluded from the analysis. The dust layer height is estimated from the extinction coefficient at 532 nm (`Extinction_Coefficient_532`). The extinction coefficient is a retrieved quantity and we only include profiles for

which the quality control flag `Extinction_QC_Flag_532` equals 0 (unconstrained retrieval; initial lidar ratio unchanged during solution process) or 1 (constrained retrieval).

2.1.1 CALIOP dust layer height

Aerosol layers heights may be termed as either "effective" or "real" heights. The "effective" layer height represents the height where the total aerosol load should be placed in order to be representative for the radiative properties of this aerosol. The thickness of "effective" layers are typically assumed to be small, 500 m or 1 km. For climate impact studies "effective" layer height is an important parameter, as it, together with the aerosol optical depth, single scattering albedo and phase function, allows quantitative estimates of the aerosols' direct radiative forcing. The "real" aerosol height may be described in terms of layer boundaries or by the full vertical profile. It is required for example for the understanding and characterising of aerosols-clouds interactions, air quality and flight safety.

There is no unique way to calculate the height of a dust layer from CALIOP data. Possible methods include:

Threshold: Calculate the cumulative extinction and set the height to be where cumulative extinction is above a prescribed threshold.

Cumulative extinction: Calculate the cumulative extinction and set height to be where the extinction column is half of the total extinction column.

Geometric mean: Identify the top and bottom heights of the dust layer and set dust height to the mean of the two.

Extinction-weighted: weigh the dust layer height z_i for layer i with an appropriate parameter and calculate the weighted average, for example using the extinction coefficient β_i as in Koffi et al. (2012):

$$z_{\mathrm{CALIOP}} = \frac{\sum \beta_i z_i}{\sum \beta_i} \tag{1}$$

In Fig. 1 the extinction-weighted and geometric mean CALIOP dust heights are plotted versus the cumulative extinction CALIOP dust height for the days and region under study. The extinction-weighted and cumulative extinction methods (right plot) are fairly similar except for heights below about 2.5 km where the extinction-weighted method gives slightly larger heights. The geometric mean method (left plot) generally gives larger heights than the cumulative extinction method. The geometric mean method is purely geometrical. The cumulative extinction, threshold, and extinction-weighted methods use the profile extinction information from CALIOP. Below we present results for one CALIOP height method that includes extinction information and one that is purely geometrical. The extinction-weighted and cumulative extinction methods are nearly similar, R^2=0.9942, right plot Fig. 1. To avoid having to arbitrarily set a threshold for the threshold method, we thus below present results for the cumulative extinction and the geometric mean methods.

It is noted that ambiguities in dust heights derived from CALIOP are larger for thick and optically dense dust layers. In these cases, the inversion of lidar profiles is less accurate for the lower part of these layers (due to uncertainties in the lidar ratio

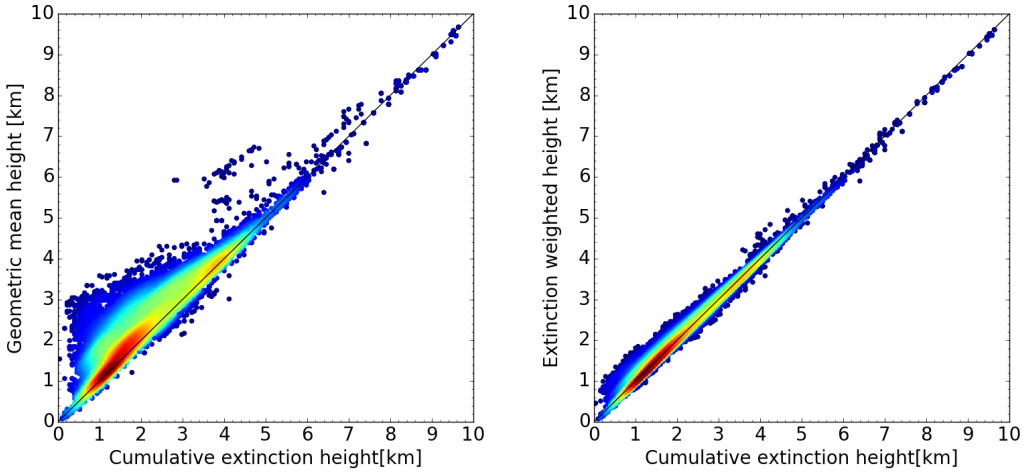

Figure 1. (Left) The CALIOP cumulative extinction height versus the CALIOP geometric mean height. Linear regression ($y = ax + b$) gives $a = 0.882$, $b = 0.6813$, R^2=0.941, and RMSE=0.652 km. (Right) The CALIOP cumulative extinction height versus the CALIOP extinction weighted height. Linear regression analysis gives, $a = 0.9619$, $b = 0.1597$, R^2=0.9942, and RMSE=0.182 km.

for dust ~~) and also due to~~ and multiple scattering effects (see e.g. Cuesta et al., 2009, 2015). Multiple scattering effects are neglected in the CALIOP operational products used here.

2.2 Passive instrument dust height retrievals

The dust layer height was estimated from measurements by IASI and GOME-2 onboard the MetOP-A satellite and SCIAMACHY onboard Envisat. MetOP-A orbits in a sun-synchronous mid-morning orbit, crossing the equator at 9:30, local solar time, in the descending node. Envisat was in a sun-synchronous polar orbit crossing the equator at 10:00 AM local solar time (MLST) in the descending node. The various dust height retrieval algorithms used in this study are summarized in Table 1 and described in more detail below.

2.2.1 The IASI algorithm at BIRA-IASB: MAPIR

The Mineral Aerosol Profiling from Thermal Infrared (MAPIR) retrieval algorithm is an extensive technical and scientific improvement of the algorithm first published by Vandenbussche et al. (2013). ~~Versions 3.2 and 3.4~~ Version 3.5 of the algorithm~~are used in this study. The improved algorithm will be~~ , fully described in ~~a separate publication, currently in writing~~Vandenbussche and De Maz is used in this study.

The MAPIR retrieval scheme is based on the Optimal Estimation Method (OEM, Rodgers, 2000), which iteratively adjusts a state vector, composed of 7 variables: the surface temperature (T_s) and the vertical profile of dust aerosols concentration, from

Table 1. Summary of dust height retrieval algorithms. See text for further details, including description of acronyms.

Institute (Algorithm name)	Radiative transfer	Satellite instrument	Algorithm specifics and aerosol type	Height	Reference(s)
BIRA-IASB (MAPIR)	Line-by-line Lidort, Mie	IASI	Optimal estimation Dust, ash	Vertical profiles and averaging kernels; 1 km resolution from 1 to 6 km; 1.5 to 2 degrees of freedom;	Vandenbussche et al. (2013) Vandenbussche and De Mazière (2017)
DLR (IMARS)	No direct forward modelling, optical properties from Mie	IASI	PCA, spectral matching, Bayesian probability Dust, ice clouds; possible application to ash	Effective layer height from emission temperature	Klüser et al. (2011), Klüser et al. (2012), Banks et al. (2013)
LMD	4A/OP-DISORT	IASI	Refractive indices from Volz (1972, 1973) and Balkanski et al. (2007)	Average weighted layer height	Peyridieu et al. (2010), Peyridieu et al. (2013), Capelle et al. (2014)
LISA (AEROIASI)	Line-by-line KOPRA, Mie	IASI	Tikhonov-Philips auto-adaptive regularisation, Dust	Vertical profiles and averaging kernels; 1 km resolution from 0 to 9 km; approx. 1.5 degrees of freedom;	Cuesta et al. (2015)
KNMI	Line-by-line DISAMAR Henyey-Greenstein phase function	GOME-2	Optimal Estimation All types	Effective layer height; accuracy of 0.5-1 km if AOD>0.3	Sanders and de Haan (2013)
IUP	Line-by-line SCIATRAN T-matrix	SCIAMACHY	Adjoint RTE Dust	Effective layer height	Lelli et al. (2017)

1 to 6 km height in steps of 1 km. ~~One of the strengths of the OEM is that it also produces averaging kernels, which enable the analysis of the sensitivity of the retrievals at the different heights, and the correlation between the retrieved parameters. The a priori T_s is , for retrievals using IASI data after 14 September 2010, taken from the IASI operational~~ The retrieval is performed only on cloud-free scenes (<10% cloud coverage in IASI level 2 ~~retrieval, version 5 and later (MAPIR version 3.2). For data prior to 14 September 2010, the T_s retrieved by the IASI operational level 2 algorithm version 4 or prior have significant issues, in particular over deserts (giving unrealistic values). Therefore, for dust retrievals at those dates we use the T_s from the European Centre Medium-Range Weather Forecast (ECMWF) era interim reanalysis as a priori (MAPIR version 3.4). The T_s a priori standard deviation is set to 5 in all retrievals. This might seem large but the diurnal variation of T_s in deserts reaches more than 20~~ cloud product). Unfortunately, the IASI cloud product seems to misflag some intense dust clouds as meteorological clouds, removing that data from our analysis.

~~As a priori for the~~ The a priori vertical profile of desert dust concentration ~~, we use~~ is obtained from the LIdar climatology of Vertical Aerosol Structure for space-based lidar simulation studies (LIVAS) monthly $1° \times 1°$climatology derived from CALIOP data (Amiridis et al., 2015). ~~That climatology contains mean vertical profiles of dust extinction at 1064 and 532 nm, with high vertical resolution. The extinctions in units of $^{-1}$ have to be converted to vertical profiles of dust particles number concentration in units of cm^{-3} at the vertical resolution of the retrieval. First, a linear interpolation is performed, from the CALIOP grid to the MAPIR grid. At altitudes higher than 7 , the aerosol concentration is set to zero. The extinction to concentration conversion~~ The conversion from 532 nm extinction to particle number concentration is done using the ~~visible (532 nm)~~ cross-section of the aerosol particles used in MAPIR~~(using a single particle size, the median radius of the particle size distribution, PSD). A second issue arises here due to~~ . To account for the fact that CALIOP measurements are sparse, ~~and therefore~~ plausibly impacting the continuity of the climatology amongst adjacent $1° \times 1°$~~cells is not ensured:~~ (the mean extinction in adjacent cells may come from measurements made for different days~~. To reduce this effect, we compute for the dust profile a priori~~), we use a horizontal running mean over 25 cells (5 in latitude, 5 in longitude). The standard deviation for the dust aerosol vertical profile is set to 100 % at all altitudes and all locations. ~~When the surface altitude in a scene is higher than one (or more) retrieval altitude(s), the dust concentration at that (those) altitude(s) is set to zero prior to the retrieval (and can not deviate from this value during the retrieval).~~

~~Three spectral windows in the thermal infrared are used for the retrievals, avoiding the huge ozone absorption band centered at 1040 $^{-1}$ (9.6): 905 to 927 $^{-1}$, 1098 to 1123 $^{-1}$ and 1202 to 1204 $^{-1}$. MAPIR relies on a very accurate radiative transfer code including multiple scattering: Lidort (Spurr et al., 2008) . Radiative transfer calculations are performed on line for each retrieval. The forward modelling is computed as to properly reproduce the sampling and resolution of level 1c IASI data. The Lidort radiative transfer code is linked to a Mie code preparing the spherical aerosol optical properties from their PSD (~~The dust aerosols in MAPIR are parameterized with a log-normal ~~,~~ particle size distribution with median radius 0.6 µm, geometric standard deviation of 2, corresponding to an effective size of 2 µm)~~and refractive index (GEISA-HITRAN dust-like, Massie, 1994; Massie and~~

~~Additional parameters are required for the radiative transfer calculations. The surface emissivity is taken from Newman et al. (2005) for oceans and from an updated version of the data reported in Zhou et al. (2011) for land. The two most important atmospheric~~

profiles (temperature and water vapour) are taken from IASI level 2 operational products from EUMETSAT (procession version number depends on the date of the data set, the quality described in August et al. (2012)). Other relevant atmospheric gas profiles (CO_2, O_3, N_2O, CH_4 and HNO_3 are included) are taken from the US Airforce Geophysics Laboratory tropical climatology (Anderson et al., 1986) . Line parameters of all gases come from the HITRAN 2012 database.

Only scenes with less than 10 cloud fraction from the IASI operational level 2 cloud product are retained for the dust retrievals. Unfortunately, the IASI cloud product seems to misflag some intense dust clouds as meteorological clouds, removing that data from our analysis. and the refractive index from Massie (GEISA-HITRAN dust-like, 1994); Massie and Goldman (GEISA-HITR

After the retrievals, quality filters are undertaken. The retrievals are marked to be of good quality when:

- the root mean square of spectral residuals is lower than 2 K over land, 1 K over oceans
- the final fitted AOD at 10 µm is lower than 8 (otherwise the probability is extremely high that the scene was cloudy and unflagged as such).

The dust detection is computed a posteriori, and based on the single criterion that the retrieved 10 µm AOD must be higher than 0.01. This is a very low threshold, and although ensuring that all dusty scenes are indeed flagged as such, it might flag scenes where the aerosol presence is questionable. The AOD is obtained by vertical integration of the concentration profile, and multiplication by the extinction coefficient at the desired wavelength. The mean height is obtained from the profile as linear interpolation of the height that would separate the aerosol column in two identical partial columns (in other words, half the aerosols are below the mean height and half are above it).

2.2.2 IASI DLR algorithm - IMARS

The retrieval method for IASI developed at DLR combines dust and ice cloud remote sensing using Principal Components Analysis (PCA) of the high resolution IASI spectra (version 4.2). Bayesian inference is used for differentiating between dust and ice clouds (Klüser et al., 2011, 2012; Banks et al., 2013; Klüser et al., 2015). The method may also be applied to volcanic ash retrievals (Klüser et al., 2015; Maes et al., 2016), with focus on the very variable composition of the ash of which there is lack of reliable reports about in the literature.

The retrieval uses spectral pattern matching between 8 and 12 µm in a subspace of the observation space, spanned by suited eigenvectors for inferring dust/ash properties from the observations. With this approach direct forward modelling of the infrared radiative transfer is avoided, as this would be strongly underdetermined due to the lack of information about surface emissivity (over deserts), atmospheric temperature, humidity profiles as well as about detailed information of dust/ash composition, particle size and sphericity. The composition of dust/ash is assumed to be represented by linear combinations of typical dust composition mixtures (Klüser et al., 2015). For extreme cases not represented by these mixtures (e.g. very high calcite of gypsum content in dust aerosols) the retrieval will not be able to correctly characterize the dust/ash load. Dust optical properties used here have been calculated with traditional Mie theory, thus ignoring particle non-sphericity (see Klüser et al., 2015, 2016; Maes et al., 2016). One of the outputs of the DLR algorithm is the dust/ash emission temperature. Using a vertical

temperature profile (standard atmosphere or model output) it is then converted to effective layer height (Klüser et al., 2015). De facto, the emission temperature is retrieved relative to the background, which implicitly delivers height information and not an absolute temperature.

The layer height is an effective emission height of a geometric thin (delta shape) dust layer. This makes its interpretation with regard to an averaged CALIPSO extinction profile non-intuitive. For optically thin dust layers the effective layer height is similar to the mean extinction of the profile; however, with growing dust AOD it moves further up in the profile (details depend on dust properties). Thus we expect a positive bias of the IMARS layer height to the cumulative extinction height from CALIPSO.

2.2.3 IASI LMD algorithm

The LMD method for the retrieval of dust characteristics from IASI observations was originally developed for application to the Atmospheric Infrared Sounder (AIRS) (Pierangelo et al., 2004, 2005), and then slightly modified as described in detail in Peyridieu et al. (2010, 2013) and in Capelle et al. (2014) for application to IASI. The method used to derive dust characteristics from IASI observations is a three-step physical algorithm based on a "Look-up Table" (LUT) approach. The first step constrains the atmospheric state (temperature and water profile) using 18 channels selected in the spectral range 4.5-14.5 $\mu\mathrm{m}$ and mostly sensitive to temperature and water profiles between 900 hPa and 200 hPa and not, or almost not, sensitive to surface characteristics (temperature, emissivity). The second step determines simultaneously the 10 $\mu\mathrm{m}$ AOD, the dust layer mean height and the surface temperature using 8 channels localized in three window regions: 8-9 $\mu\mathrm{m}$, 10-12 $\mu\mathrm{m}$ and 4.6-4.7 $\mu\mathrm{m}$. This selection of channels, both at short and long wavelengths, is aimed at decorrelating the contribution of AOD, height and surface temperature to the observed signal. The dust coarse mode particle effective radius can be determined in a third step.

For each step, LUTs of IASI simulated brightness temperatures are calculated using the forward coupled radiative transfer model 4A/OP-DISORT (available from http://4aop.noveltis.com, Scott and Chédin, 1981). Entries to the model include: AOD, height, surface pressure, surface temperature and emissivity, viewing angle, two refractive indices: Volz (1972, 1973) and Balkanski et al. (2007), and a set of 2311 atmospheric situations, selected by statistical methods from 80,000 radiosonde reports, and stored in the Thermodynamic Initial Guess Retrieval (TIGR) climatological database (Chédin et al., 1985; Chevallier et al., 1998). The PSD is modeled by a monomodal lognormal distribution described by the effective radius (R_{eff}) and the standard deviation of the distribution σ_g. Following results of previous sensitivity studies (Appendix A, Capelle et al., 2014; Pierangelo et al., 2005), fixed values are taken for the effective radius ($R_{\mathrm{eff}} = 2.3\ \mu\mathrm{m}$) and for the standard deviation of the size distribution ($\sigma_g = 0.65$).

The aerosol vertical distribution is supposed to be concentrated within a single homogeneous layer. While this assumption cannot describe correctly observations, in general more complex, the height retrieved here can be defined as an average weighted height for which half of the dust optical depth is below and half of the optical depth is above. This infrared optical equivalent to the real vertical profile is therefore appropriate for computing dust infrared forcing. It is worth noting that the resulting mean layer height corresponds to height above sea level. Several aspects of the retrieval algorithm: robustness to aerosol model (size distribution, shape, and refractive indices), possible contamination by other aerosol species, radiative

transfer model bias removal, or cloud mask including discrimination between clouds and aerosols, etc., were investigated and details may be found, for example, in Pierangelo et al. (2004) and Capelle et al. (2014). The surface emissivity spectrum is supposed to be known and is read from a 0.5° monthly grid retrieved from IASI (Capelle et al., 2012).

2.2.4 IASI LISA algorithm - AEROIASI

The IASI algorithm from LISA, called AEROIASI, has been conceived to observe the three-dimensional (3D) distribution of desert dust plumes for each overpass of IASI, both over land and ocean (Cuesta et al., 2015). For this, it derives vertical profiles of desert dust, in terms of the extinction coefficient at 10 μm, from individual thermal infrared spectra measured by IASI. It is a constrained-least-squares fit method, based on explicit radiative transfer calculations, where the vertical distribution and abundance of dust are iteratively adjusted in order to fit IASI observations. This approach uses auto-adaptive constraints for adjusting simultaneously the dust profile and surface temperature in order to offer particularly good adaptability for different atmospheric and surface conditions. This flexibility makes the aerosol retrieval possible for most cloud-free IASI pixels, both over ocean and land (even for bright surfaces and relatively low aerosol loads). The information on the vertical distribution of dust is mainly provided by their broadband radiative effect, which includes aerosol thermal emission depending at each height on the vertical profile of temperature (assuming local thermal equilibrium).

AEROIASI uses an a priori desert dust model (including dust microphysical properties) and meteorological profiles provided as inputs to the radiative transfer model. The line-by-line Karlsruhe Optimized and Precise Radiative transfer Algorithm (KOPRA, Stiller, 2000) is used to simulate thermal infrared radiance spectra and the inversion module KOPRAFIT to compare them to those measured by IASI, for 12 selected spectral micro-windows in the atmospheric window between 8 and 12 μm. KOPRA accounts for light absorption, emission, and single scattering by aerosols, using dust optical properties derived at each wavelength with a Mie code (Metzig, 1984) optimized as described in Deirmendjian et al. (1961) and Kerker (1969). The vertical grid of all profiles in the simulations is set between the surface and 9 km height asl (above mean sea level), with 1 km increments. For each pixel, we use atmospheric temperature profiles and first guesses of surface temperatures and water vapor profiles from ERA-Interim (ERAI) reanalysis (Dee et al., 2011) of the ECMWF. For all seasons and locations, AEROIASI uses a unique a priori vertical profile of dust derived from CALIOP average dust profiles over the Sahara on the summer of 2011.

~~In order to minimize the spectral residuals, the method adjusts iteratively the radiative transfer inputs (mainly the aerosol vertical profile and surface temperature) until reaching convergence (the maximum number of iterations is fixed to 10). The retrieved dust profile is obtained with an auto-adaptive Tikhonov-Phillips-type (Tikhonov, 2003) height-dependent regularization conceived to avoid unrealistic oscillations in the retrieved dust profiles while adapting the results to the amount of information provided by the IASI measurements. The overall results of AEROIASI show that the degrees of freedom or the number of independent pieces of information in the retrieval of dust profiles varies during the iterative procedure, a typical value of ~1.5 being used to determine the shape of the dust extinction profiles.~~

Once IASI spectra are fitted, a series of quality checks are performed in order to screen out cloudy measurements and aberrant retrievals. We exclude IASI pixels with derived surface temperatures below their ERAI reanalyses counterparts by more than 10 K and those pixels exhibiting too high root-mean-squared spectral residuals or horizontal variability with respect to their

closest pixels. For each quality-checked retrieval, we derive a vertical profile of dust extinction coefficient (α_{10} in km^{-1}) at 10 µm, the associated AOD (by vertical integration of the extinction profile), and mean and top heights of the observed dust layer. The mean height of the observed dust layer (the product used in the current paper) is defined as the height below which the integral of the extinction coefficient profile reaches 50 % of the AOD. Likewise, the height of the top of the dust layer is defined as the height below which the integral of the extinction coefficient profile reaches 95 % of the AOD. Indeed, AEROIASI provides valuable information not only on the mean height of the dust layers but also on their vertical extent (i.e., layer top heights and whether the layers reach the ground or are elevated.

In this study AEROIASI retrievals from version 2 of the algorithm are used. It mainly differs from the previous version described by Cuesta et al. (2015) in the a priori desert dust model and the surface emissivity database. Using these new databases, we obtain lower spectral residuals with respect to IASI measurements than with the previous version and higher adaptability for covering the large region analysed in this paper (i.e. the tropical dust belt). The climatological desert dust model consists of refractive indices, a single-mode lognormal particle size distribution and an a priori vertical profile of dust. Refractive indices are taken from field measurements of Saharan dust analysed by Di Biagio et al. (2014). The modal radius and width of the single-mode distribution are prescribed from average volume effective radius and width for the coarse mode ~~(for radii>0.6 µm of the AERONET size distributions, http://aeronet.gsfc.nasa.gov, Dubovik et al., 2002)~~ derived from Saharan ground-based stations in June ~~2011.~~ 2011 (for radii>0.6 µm of the AERONET size distributions, http://aeronet.gsfc.nasa.gov, Dubovik A unique first guess of dust vertical distribution (the same profile for all pixels and all seasons) is considered in the inversion, which is obtained from an average of CALIOP extinction vertical profiles for dust over the Sahara (during large dust outbreaks in late June 2011), scaled to particle concentration units (in order to set an a priori AOD at 10 µm of 0.03). Forward simulations include surface emissivity from a global monthly IASI-derived climatology over land (Paul et al., 2012) and a surface temperature dependent model over ocean (Newman et al., 2005).

2.2.5 GOME-2 KNMI algorithm

The deep oxygen lines (A band and/or B band) in the near infrared of the shortwave spectrum traditionally have been used for retrieval of the cloud height. In the absence of clouds, these bands contain information on the aerosol height (Wang et al., 2012). The algorithm developed at KNMI within the TROPOMI/Sentinel 5 Precursor programme (Veefkind et al., 2012) is based on the optimal estimation method and aims at deriving the aerosol layer height (Sanders and de Haan, 2013). This method has also been applied to Greenhouse gases Observing SATellite (GOSAT) and GOME-2 data within the on-going ESA AEROPRO study, in support of the Sentinel-4 development. The algorithm is sensitive to all aerosol types, including dust, biomass burning and industrial pollution plumes. Sensitivity analyses performed for the TROPOMI/Sentinel 5 Precursor ~~ATBD~~ Algorithm Theoretical Basis Document (ATBD) indicate that the aerosol layer height can be derived with an accuracy of 0.5-1 km, if the AOD is 0.3 or larger.

The algorithm uses the Determining Instrument Specifications and Analyzing Methods for Atmospheric Retrieval (DISAMAR) retrieval and simulation package developed at KNMI. In the setup that is used in this work, the aerosol is modelled as a 50 hPa thick layer, for which the height and the aerosol optical depth is fitted. The single scattering albedo of the aerosol particles

is assumed to 0.95 and we apply a Henyey-Greenstein phase function with an asymmetry parameter of 0.7. A climatological value for the surface reflectance is used. Pressure-temperature profiles are obtained from the operational ECMWF forecast. The algorithm uses a fit window between 758 and 762 nm.

The algorithm is only applied to cloud-cleared scenes for which the UV aerosol index has a value exceeding 1.0, indicating the presence of absorbing aerosol layers. For the GOME-2 data used in this work, the cloud clearing is done based on the GOME-2 data itself, which may result in undetected sub-pixel cloudiness. It is noted that the size of the GOME-2 ground pixels is much larger than for the TROPOMI instrument, for which the algorithm has been designed.

2.2.6 SCIAMACHY IUP algorithm

The height of an aerosol layer is determined using top-of-atmosphere (TOA) reflectances R (defined as the sun-normalized radiances, weighted by the cosine of the solar zenith angle) in the oxygen A-band, that is, in the range 758-772 nm, at the nominal spectral sampling 0.21 nm of SCIAMACHY, for a Gaussian instrument response function of 0.48 nm. The retrieval is based on the calculation of the weighting functions $W(h,b,\tau) = \partial R(h,b,\tau)/\partial(h,b,\tau)$, i.e., the Jacobians of R as function of the top and bottom altitude h, b and optical thickness τ of the aerosol layer. Upon linearization of the problem, the measured R in a gaseous absorption band can be written as a function of the desired h. Given that τ is inferred from an independent source, such as a non-absorbing channel outside the oxygen A-band, typically $\lambda = 758$ nm, the retrieval can be further simplified assuming either that the aerosol layer originates at the ground ($b = 0$ km) or is elevated ($b \neq 0$ km). The latter assumption implies that the prior geometrical thickness is preserved when retrieving h. Either way, the problem is reduced to the calculation of $W(h)$ (Rozanov, 2006; Rozanov et al., 2007) and the minimization of the difference between the forward-modelled and the measured reflectance, converging after $\sim$4 iterations on average, delivers the height of the layer.

Information on the local non-spherical dust optical properties, encoded in the spectral scattering T-matrix (Dubovik et al., 2006), as well as the single-scattering albedo and the aerosol extinction (box) profiles are embedded in $W(h)$. It has been assumed that these quantities are independent of height inside the aerosol layer. The HITRAN 2008 edition (Rothman et al., 2009) is used for the line intensities of the absorbing species (oxygen and water vapour) included in the forward problem. The full retrieval chain is implemented and carried out within the radiative transfer model SCIATRAN (Rozanov et al., 2014)

The selection of cloud-free SCIAMACHY pixels (60×40 km^2 of nominal footprint size) relies on the analysis of joint histograms of geometric cloud cover (CC < 0.1, from colocated 1×1 km^2 MERIS observations, Schlundt et al., 2011) and aerosol absorbing index (AAI > 1.0, de Graaf et al., 2005)) for the area of interest. Surface reflectivity is taken from the MERIS-derived black-sky dataset (Popp et al., 2011) which is the critical parameter for the accuracy of the retrieved h. Already an error of ±10 $\%$ in the a priori value of surface reflectivity can cause a bias up to ±1 km for τ = 0.25 and h> 3.0 km. More details about the IUP algorithm, as well as validation with independent measurements when applied to an elevated ash layer, are given in Lelli et al. (2017).

2.3 Data selection and comparison methodology

The selection of data and the comparison between CALIOP and the other satellite instrument estimates of dust heights proceed through the following steps for each date listed ~~above:~~ in the beginning of section 2:

1. Identify CALIOP swaths that are within the region of interest.

2. Identify closest CALIOP swath and IASI, GOME-2 and SCIAMACHY dust pixels in time and space. Due to the difference in the CALIPSO equator crossing time (13:30) and MetOP-A and Envisat equator crossing times (09:30 and 10:00), a maximum time difference of 5 hrs is allowed between CALIOP and IASI, GOME-2 and SCIAMACHY dust pixels in this step. To allow for possible movement of dust pixels between overpasses, pixels within 500 km were included for subsequent analysis. This allows for a maximum wind speed of 100 $\mathrm{km/hr}$.

3. For the CALIOP swaths from step 2 identify CALIOP dust profiles using the CALIOP dust flag and CAD score. Calculate CALIOP cumulative extinction and geometric mean dust layer heights.

4. Move CALIOP dust layer heights from the previous step backward in time to the Metop-A and Envisat overpass times using the FLEXTRA trajectory model.

5. After moving the CALIOP dust heights backward in time they may still be at locations different from the IASI, GOME-2 and SCIAMACHY dust heights. A second colocation is thus made to colocate the moved CALIOP dust heights with IASI, GOME-2 and SCIAMACHY dust heights. The maximum difference in distance is set to 20 km for IASI and 100 km for SCIAMACHY and GOME-2 reflecting the larger footprints of the latter two instruments.

6. Analysis of height differences including statistics.

In ~~the top panels of Figs. **??**-**??** are shown examples~~ Fig. 2 examples are shown of data from steps 1-5. ~~(Top) The IASI dust height from the BIRA-IASB analysis. The CALIOP profiles identified as dust, within 500 and 5 hrs time differences from nearest IASI pixel, are overlaid (grey dots). The location of the CALIOP height after shifting to IASI overpass time is shown by the blue dots. The time range (UTC) in the title gives the times of the first and last CALIOP points plotted. (Bottom) Curtain plot of the CALIOP extinction coefficent for heights identified as dust. IASI-BIRA-IASB dust layer heights are shown as purple dots (colocated with CALIOP cumulative extinction heights) and pink triangles (colocated with CALIOP geometric mean heights). Red crosses are CALIOP cumulative extinction heights that have been shifted to the IASI pixel location using FLEXTRA. The green crosses are the CALIOP cumulative extinction heights before the shift. Deep pink plusses are CALIOP geometric mean heights that have been shifted to the IASI pixel location using FLEXTRA. The lime plusses are the CALIOP geometric mean heights before the shift. The black dots are the column optical depth at 532 nm from CALIOP. The curtain is for the CALIOP data between 40-60°E in the top plot. The time range (UTC) in the title gives the times of the first and last CALIOP extinction coefficient profiles plotted. See text for further details.~~ ~~Similar to Fig. **??**, but IASI dust height from DLR analysis.~~ ~~Similar to Fig. **??**, but IASI dust height from LMD analysis.~~ The pixels identified as dust from IASI

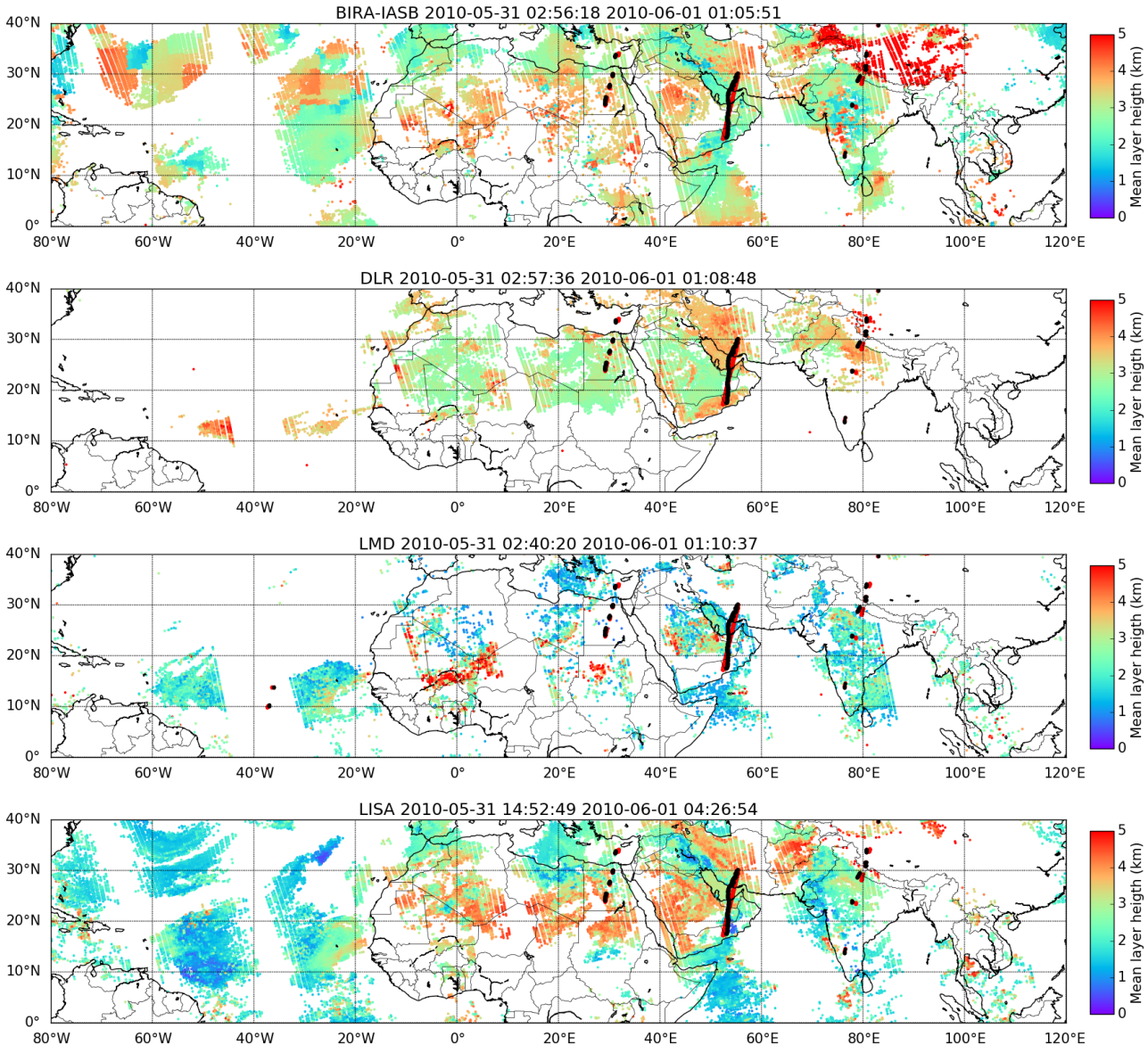

Figure 2. ~~Similar to Fig. ??, but~~ The IASI dust layer height from the BIRA-IASB (top plot), DLR (second plot), LMD (third plot) and LISA (bottom plot) analysis. The CALIOP profiles identified as dust, within 500 km and 5 hrs time differences from nearest IASI pixel, are overlaid (red dots). The location of the CALIOP height after shifting to IASI overpass time is shown by the black dots. The time range (UTC) in the title gives the times of the first and last CALIOP points plotted.

data by the BIRA-IASB (top plot, Fig. ??2), DLR (second plot, Fig. ??2), LMD (third plot, Fig. ??2), and LISA (bottom plot, Fig. ??2) algorithms, ~~the KNMI GOME-2 (Fig. ??) and the IUP SCIAMACHY dust data (Fig. ??)~~ are overlaid by

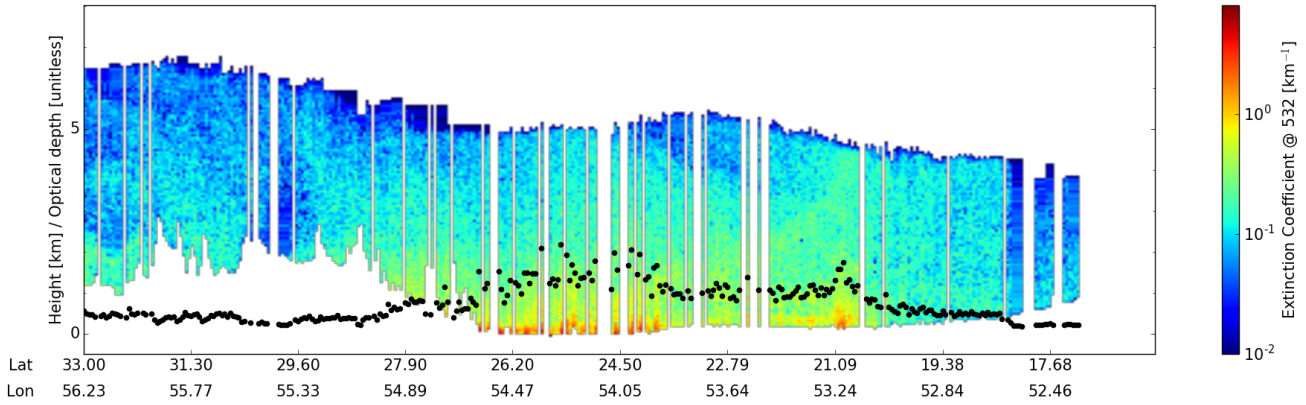

Figure 3. ~~Similar to Fig. ??, but GOME-2~~ Curtain plot of the CALIOP extinction coefficent for heights identified as dust~~height from KNMI analysis. Note difference in time from Fig~~. The black dots are the column optical depth at 532 ??nm from CALIOP. The curtain ~~in the bottom plot~~ is for the CALIOP data between ~~0-20~~40-60°E in the top plot.

CALIOP cumulative extinction heights, derived from profiles identified as dust (step 3, ~~grey~~ red dots), that are within the temporal and spatial requirements. CALIOP data are recorded after the IASI overpass. To account for possible movements of the dust between the overpasses the CALIOP dust heights were moved in longitude, latitude and height using the FLEXTRA model ~~(step 4, Stohl et al., 1995; Stohl and Seibert, 1998) . FLEXTRA calculates mean-wind trajectories using data from the ECMWF. It~~ (step 4, Stohl et al., 1995) . FLEXTRA does not include turbulence or loss processes. ~~The blue dots in the upper plots of Figs. ??-??~~ FLEXTRA calculated mean-wind trajectories with meterological input data from the ECMWF. Here operational data with a 1° latitude × 1° longitude resolution, 91 vertical levels and a time resolution of 3 hrs were used. Quantification of trajectory errors is always difficult, due to a general lack of ground-truth data. However, FLEXTRA has been quantitatively evaluated in the past. Comparisons of FLEXTRA trajectories driven with ECMWF data with balloon trajectories have revealed typical horizontal transport errors of about 20% of the travel distance, but with large variability from case to case (Baumann and Stohl, 1997; Stohl and Koffi, 1998; Riddle et al., 2006) . Evaluation against meteorological tracers such as potential vorticity suggests errors of a similar magnitude (Stohl and Seibert, 1998) . Thanks to improvements in the meteorological analysis data, slightly smaller errors may be assumed for more recent years, but the order of magnitude of the errors is likely still similar.

The black dots in Fig. 2 are CALIOP dust height pixels that have been moved from their original location (~~grey~~ red dots) to the nearest IASI pixel (step 5). As the cumulative extinction and geometric mean CALIOP dust heights are different they will be moved by FLEXTRA to different locations. ~~This~~ An example of this is seen in ~~the curtain plots in Figs. ??-??~~ Fig. 4 where the cumulative extinction (~~purple dots~~black circles) and geometric mean (~~pink triangles~~red circles) heights from the passive instruments sometimes overlap (not moved or moved to same location and height) and sometimes do not overlap (moved

to different location and/or height). It is also seen in the difference in number of colocated ~~points, compare grey and light grey entries in~~ point, Table 2. For the full data period the CALIOP dust heights were on average moved upwards by 0.015 (cumulative extinction) and 0.020 km (geometric mean), both with standard deviations of 0.25 km.

3 Results

The analysis steps 1-5, section 2.3, were performed for all days and algorithms. The number of dust pixels identified by the various algorithms after step 1 is given in Table 2. The number of pixels identified as dust by the various IASI-algorithms varies by a factor of 4.6. The differences reflect the differences in dust detection methods and it is outside the scope of this study to further investigate the reasons for these differences. As expected the solar algorithms detect far fewer dust pixels due to only day time coverage (factor of 2) and larger pixels size (factor of about 16). The difference between the two solar algorithms (KNMI and IUP) are due to differences in the constraints set to detect dust. In step 2 dust pixels within a given

Table 2. The number of data points (dust heights) at data reducing step of the data analysis chain described in section 2.3. ~~Numbers in parenthesis are in percentage relative to the total number in the previous analysis step.~~ Step number refers to the analysis steps as described in section 2.3.

~~Step~~ Institute (Instrument/Algorithm) Step	BIRA-IASB (IASI/MAPIR)	DLR (IASI/IMARS)	LMD IASI	LISA (IASI/AEROIASI)	KNMI GOME-2	IUP SCIAMACHY
1	2324277	503944	811360	1770793	21535	2710
2[a]	13377 (0.58)	5208 (1.0)	14916 (1.8)	13110 (0.74)	3715 (17.3)	1979 (73.0)
5-cumulative extinction[b]	2620 (19.6)	1420 (27.3)	748 (5.0)	2203 (16.8)	215 (5.8)	34 (1.7)
5-geometric mean[b]	2408 (18.0)	1296 (24.9)	704 (4.7)	1978 (15.1)	91 (2.4)	21 (1.1)

a: Numbers in parenthesis are data points in percentage relative to the total number in the previous analysis step.

b: Numbers in parenthesis are data points in percentage relative to the total number in analysis step number 2, for example 18.0=2408/13377 (column 2, row 4).

time and distance from the CALIOP detected dust pixels are selected. This step reduces the number of IASI data points to between 0.58-1.8 % of those in step 1. The number of GOME-2 and SCIAMACHY points are reduced to 17.3 and 73.0 %, respectively. The movement of CALIOP dust heights to the MetOp-A and Envisat overpass times and the final colocation of CALIOP heights and dust pixels give the final number of dust heights to be compared to CALIOP dust heights, see values for step 5 in Table 2, and ~~grey and light grey entries in~~ Table 3.

Inspection of IASI retrieved dust heights shown in ~~the upper plots of Figs. ?? ??~~ Fig. 2 reveal differences in dust detection and dust height between the various algorithms. While differences in dust detection is not the subject of this paper, we do, however, note that there are substantial differences in the pixels identified as containing dust by the various algorithms. In particular the DLR algorithm detects very little dust over the ocean regions; the BIRA-IASB and LISA algorithms detect dust

over the ocean west of 40°W and north of 20°N whereas the DLR and LMD algorithms do not detect dust in this region; for the example ~~curtain~~ swath plots in Figs. ~~??-??~~ 3 and 4 all algorithms except LMD detects dust north of about 28°N. The BIRA-IASB-algorithm's detection of dust over Himalaya is due to retrievals being undertaken for all non-cloudy scenes, and the final result may never be a true zero due to the method used. These retrievals have a low aerosol optical depth (AOD) and their inclusion indicates that the AOD threshold for the dust flag may be be too permissive. The differences in dust detection is the reason for the different number of pixels available for comparison with CALIOP.

~~The curtain plots in Figs. ??-?? show that~~ For the example CALIOP swath shown in Fig. 3 the BIRA-IASB-algorithm (~~purple dots and pink triangles~~red and black circles, Fig. 4) agrees reasonably with the CALIOP geometric mean heights (~~deep pink plusses~~red triangles) and gives higher dust heights compared to the CALIOP cumulative extinction heights (~~red crosses~~black triangles). For the DLR-algorithm the situation is similar, albeit the DLR-algorithm generally gives larger dust heights. The LMD-algorithm heights are generally similar to CALIOP cumulative extinction heights while LISA-algorithm heights are in better agreement with the geometric mean heights. For this transect, BIRA-IASB and LISA algorithms capture the rather monotonous decrease of dust layers heights from about 4 km of altitude near 30°N to 2 km of altitude at 19°N depicted by CALIOP geometric mean heights. The LMD algorithm retrieves dust heights near 1.5 km of height at 24-27°N as for CALIOP cumulated extinction estimates. The behaviour for this single overpass is also present in the full IASI data set as shown in Figs. 5-6 and Table 3. Do, however, note that there are substantial differences when comparing the passive methods with the CALIOP cumulative extinction and geometric mean methods. Overall, the CALIOP geometric mean method gives a larger CALIOP dust height, Table 4. Thus, the CALIOP minus passive instruments difference is smaller for the geometric mean method compared to the cumulative extinction method. The geometric mean method also gives slightly smaller standard deviations and more dust heights from the passive instrument within the CALIOP dust layer, see ~~the apricot colored entries in~~ Table 3. This may point to a non-symmetrical vertical distribution of the aerosols, with more aerosol in the lower part of the layer, where IASI algorithms usually have less sensitivity (depending on surface temperature).

In the upper row of Fig. 5 the CALIOP cumulative extinction height is plotted against the dust heights from all IASI algorithms for all dates. Fig. 6 is similar but for the CALIOP geometric mean height method. Also included in the plots are the Pearson's correlation coefficient and root mean square error (RMSE). In the centre rows of Figs. 5-6 the differences between the passive algorithms and CALIOP heights are shown versus the CALIOP column extinction. In the upper and centre rows the color indicates the density of points. In the bottom rows of Figs. 5-6 are shown the frequency distribution of the difference between the dust heights from the various IASI algorithms and the CALIOP heights. ~~Also included is a fit to the normal distribution and the mean and standard deviation (σ) of the fit.~~ Similar plots for the KNMI and IUP algorithms are shown in Fig. 7. For the IASI algorithms ocean-day, ocean-night, land-day and land-night data subsets are presented in Figs. A1-A4 for the CALIOP geometric mean heights, and in Figs. A5-A8 for the CALIOP cumulative extinction heights. The mean and standard deviation and the number of data points are also listed in Table 3. It is noted that an analysis in terms of "bias" is correct as a mean analysis only when the difference distribution is at least symmetrical (if not Gaussian). This is not always the case as shown for example for the ocean day subset in Fig. A8. Thus, while the mean of the difference may appear good the histogram sometimes shows something very different.

Table 3. The mean (, and boxes) ± the standard deviation of the dust height difference between the passive sensors and CALIOP. Also given is the standard deviation (σ) and , and the number (#) of colocated points. The inlay () is the percentage of heights that are within the CALIOP layer. Bright colored numbers are for the cumulative extinction heights while the dimmed colors are for the geometric mean CALIOP dust heights. For BIRA-IASB, DLR, LISA and LMD statistics are given for all () data and subgroups of data recorded during day and night and over land () and ocean()and ocean. For KNMI-GOME2 only day comparisons are possible, hence the lack of comparisons with CALIOP night overpasses. Note that for KNMI Aslo note that for KNMI-GOME2 the number of land and ocean pixels does not add up to the total due to some pixels covering coastal regions (mixed pixels).

Institute	BIRA-IASB	DLR	LMD	LISA	
Instrument/Algorithm	mean σ mean σ mean σ mean IASI/MAPIR	σ IASI/IMARS	mean IASI	σIASI/AEROIASI	mo
() () () () () height	CALIOP Day and Night, cumulative extinction heights				
Height difference (km)	0.590±1.213	0.785±1.281	-0.053±1.339	0.507±1.126	-(
() () points (#)	()2620	1420	()748	2203	
All 1.213 1.281 1.339 inlay (%)	83.1	1.126 78.3	77.5	1.45585.8	
	CALIOP Day and Night, geometric mean heights				
Height difference (km)	0.078±1.108	0.243±1.181	-0.607±1.187	-0.045±1.029	-(
points (#)	2408	1296	704	1978	
inlay (%)	81.1	77.5	75.9	84.0	
	CALIOP Day, Land, cumulative extinction heights				
Land Height difference (km)	0.357±1.665	0.405±1.660	-0.102±1.448	-0.225±1.454	-(
points (#)	605	1.572 377	319	1.526 440	
inlay (%)	1.322 58.5	61.0	1.35770.2	71.4	
	CALIOP Day, Land, geometric mean heights				
Height difference (km)	0.087±1.572	-0.044±1.526	-0.496±1.322	-0.635±1.357	-(
points (#)	598	393	322	425	
inlay (%)	57.7	59.8	70.5	70.1	
	CALIOP Day, Ocean, cumulative extinction heights				
Height difference (km)	0.783±0.913	0.913±1.539	-0.501±1.409	0.172±1.389	-(
Ocean points (#)	172	0.913 22	118	1.539 180	
inlay (%)	1.409 74.4	59.1	1.38958.5	62.2	
	CALIOP Day, Ocean, geometric mean heights				
Height difference (km)	0.340±1.187	0.184±1.174	-0.922±1.142	-0.285±1.187	-(
points (#)	170	22	109	170	
inlay (%)	72.4	68.2	55.0	59.4	
	CALIOP Night, Land, cumulative extinction heights				
Height difference (km)	0.567±1.020	0.906±1.062	0.073±1.092	0.663±0.896	
points (#)	1501	996	206	1226	
inlay (%)	91.0	84.8	92.2	91.2	
	CALIOP Night, Land, geometric mean heights				
Land Height difference (km)	0.038±0.903	1.020 0.358±0.964	-0.579±1.058	1.062 0.170±0.855	
points (#)	1330	1.092 854	177	0.896 1064	
inlay (%)	89.4	0.903 85.4	89.8	0.964 89.6	
	CALIOP Night, Ocean, cumulative extinction heights				
Height difference (km)	1.008±0.741	1.058 1.599±1.127	0.352±1.180	0.855 1.043±0.637	
points (#)	342	25	105	357	
inlay (%)	96.5	96.0	92.4	97.2	
	CALIOP Night, Ocean, geometric mean heights				
Height difference (km)	0.094±0.678	0.835±0.720	-0.674±0.878	0.152±0.486	
points (#)	310	27	95	319	
Ocean inlay (%)	0.741 95.8	96.3	1.127 91.7	96.9	

Table 4.

CALIOP mean cumulative and geometric mean dust layer height± standard deviation together with the dust layer thickness± standard deviation. Statistics are given for the full dataset and for subsets divided in the land and ocean for day and night overpasses.

Quantity	~~1.180~~ CALIOP All	CALIOP Land		~~0.637~~	CALIOP Ocean	
		~~0.678~~ Day	Night	~~0.720~~	Day	~~0.878~~ Night
Cumulative extinction height [km]	2.32±1.39	~~0.486~~ 2.75±1.77	2.47±1.13		2.07±1.50	1.05±0.63
Geometric mean height [km]	2.86±1.26	3.04±1.70	3.02±1.02		2.52±1.33	2.01±0.61
Thickness [km]	3.20±1.35	2.35±1.33	3.55±1.20		2.54±1.39	3.70±1.09

For the BIRA-IASB, LISA and LMD algorithms versus the CALIOP dust cumulative extinction (geometric mean) height, the Pearson's correlation coefficient is between 0.408-0.510 (0.414-0.518) . It is smaller for the DLR, KNMI and IUP algorithms, being -0.115-0.120 (-0.238-0.137). For the IASI-algorithms the RMSE is between 1.030-1.334 km when comparing with the CALIOP geometric mean dust heights. It increases to 1.235-1503 km for the CALIOP cumulative extinction dust heights. For the KNMI and IUP algorithms the RMSE is larger, 1.670-3.439 km. The rather large RMSE indicates the difficulty and uncertainty involved when comparing dust heights from very different sensors and data recorded at different times with large differences in footprint size. There appears to be no dependence of height differences on dust column extinction as shown in the centre rows of Figs. 5-7.

For the IASI algorithms both day and night data are included in Figs. 5-6. The mean height difference between the various algorithms and the CALIOP heights are ~~high-lighted in dimmed yellow (geometric mean) and bright yellow (cumulative extinction)~~ given in Table 3. The BIRA-IASB mean height difference is 0.078 km (0.590 km) when compared with the CALIOP geometric mean (cumulative extinction) height. The DLR algorithm mean height difference of 0.243 km (0.785 km) is larger. However, it is noted that the DLR algorithm generally gives the altitude at two distinct modes, Fig. 5. For LMD the magnitude of the mean height difference is smallest when comparing with the CALIOP cumulative extinction height, -0.053 km. It increases to -0.607 km when comparing with the CALIOP geometric mean height. For the LISA algorithm the behaviour is similar to the BIRA-IASB and DLR algorithms with mean height differences of -0.045 km (geometric mean) and 0.507 km (cumulative extinction). Scatter plots in Figs. 5-6 (upper panels) reveal rather elongated clouds of points along (parallel to) the 1:1 straight line for BIRA-IASB and LISA with respect to geometric mean (cumulative extinction) heights from CALIOP, whereas the point cloud is mainly localized below 2 km of altitude for LMD and above 2.5 km for DLR (this last one presenting maxima of occurrences). The standard deviations are similar for the BIRA-IASB, DLR, LMD and LISA algorithms, between 1.029-1.187 km (geometric mean) and 1.126-1.339 km (cumulative extinction), being slightly lower for LISA, intermediate for BIRA-IASB and DLR, and to some extent greater for LMD.

Due to the larger footprint size for the solar sensors, fewer data points are available for dust height comparison of CALIOP with GOME-2 and SCIAMACHY. ~~Data for a single day are given in Figs. ?? (KNMI) and ?? (IUP). Note that the data are from different days.~~ The statistics for all colocated GOME-2 and SCIAMACHY with CALIOP points are summarized in Fig. 7 and Table 3. Both algorithms give lower dust heights compared to CALIOP, with IUP being on the average lower

by -1.097 km (-0.961 km) compared to the CALIOP geometric mean (cumulative extinction) height and KNMI lower by -1.393 km (-0.818 km).

The features seen in the upper rows of Figs. 5-7 reveal that height differences may depend on region and time of day or other variables. It is well known that CALIOP daytime measurements are more noisy due to straylight from the sun. Thus to investigate possible differences between night and day time data the differences between CALIOP heights and the passive algorithm heights were calculated separately for night and day and also for land and ocean. For each IASI-algorithm plots similar to those in Figs. 5-7 for ocean-day, ocean-night, land-day and land-night data subsets are shown in Figs. A1-A4 for the CALIOP cumulative extinction heights and in Figs. A5-A8 for the CALIOP geometric mean heights. The results are also summarized in Table 3.

For BIRA-IASB the mean difference is similar over land during day (0.087 km) and night (0.038 km) when comparing with the CALIOP geometric mean height. For the cumulative extinction height the mean difference increases from 0.357 km during day to 0.567 km during night over land. Over the ocean the mean difference is somewhat larger during the day (0.340 km) than night (0.094 km) for the geometric mean height, while it is vice versa for the cumulative extinction height, being 0.783 km (day) and 1.008 km (night). For DLR few data points are available over the ocean. Over land the mean difference is smaller for the day data than the night data, being -0.044 km (0.405 km) and 0.358 km (0.906 km) respectively for the geometric mean (cumulative extinction) height. For LMD the dust heights over land are found to be smaller than the CALIOP geometric mean (cumulative extinction) height during day than night, -0.496 km (-0.102 km) versus -0.579 km (0.073 km). Over the ocean the behaviour is similar, but the differences somewhat larger, see Table 3. The magnitude of the mean LISA difference is larger during the day (-0.635 km) than night (0.170 km) over land for the geometric mean height. For the cumulative extinction height the behaviour is opposite, being -0.225 km during day and 0.663 km during the night. Over the ocean a similar behaviour is observed. For nearly all comparisons the RMSE is smaller for the night data than the day data, most likely reflecting the lower noise in the CALIOP night data.

~~The plots in Figs. A1-A8 reflect the findings presented in Table 3. The BIRA-IASB algorithm agrees well with the CALIOP geometric mean height over land for day and night, Fig. A1. Over ocean the agreement is better during the night. It is noted that for the ocean day subset the histogram is bimodal. When compared with the cumulative extinction height, Fig. A5, the BIRA-IASB dust height is overestimated over ocean during both day and night; the ocean day subset appears to be bimodal; the agreement appears better over land during day than night, but this may in part be due to a bimodal histogram for the land day subset. This is reflected in the spread in the difference which is smaller during night than day. For DLR, Figs. A2 and A6, there are few data points available over the ocean. Over land the DLR height data are clumped at a single height for the day subset and at two heights for the night data subset. For LMD, the agreement is monomodal for the land day, land night and ocean night subsets when compared with both the CALIOP cumulative extinction and geometric mean heights. For the ocean day subset a bimodal distribution may be present. Overall the agreement is better when compared with the cumulative extinction height. The LISA data, Figs. A3 and A7, also have a bimodal ocean day distribution compared with the CALIOP heights. For the ocean the mean difference with the CALIOP cumulative extinction height is significantly larger during night than day. This difference is nearly a factor 2 smaller when comparing with the geometric mean height. For land the magnitude of the~~

~~difference is smallest when compared with the cumulative extinction height during the day and with the geometric mean height during night~~These findings are further discussed and illustrated with plots in Appendix A. The KNMI-GOME-2 dust heights compares better with the CALIOP cumulative extinction (geometric mean) dust heights over land, difference of -0.229 km (-0.893 km), than over ocean, -1.477 km (-2.015 km) Table 3.

The four dust episodes investigated may have dust with different optical characteristics that may have an effect on the retrieved dust heights. The comparison was therefore further subdivided into four time periods representing the episodes. ~~In Figs. ??-?? is shown box, whisker and flier plots~~ Investigations of the difference between the height from the various algorithms and the CALIOP cumulative extinction ~~height~~ and geometric mean heights for the four episodes ~~.~~ reveal no clear temporal variations.

~~Box, whisker and flier plots of the difference between the height from the various IASI algorithms and the CALIOP cumulative extinction height. The box covers the lower (Q_1) to upper quartile (Q_3) values of the data. The horizontal line in the box is at the median. The whiskers show the range of the data where range is between $Q_1 - 1.5\Delta Q$ and $Q_3 + 1.5\Delta Q$, $\Delta Q = Q_3 - Q_1$. Flier points are those beyond the whiskers. Similar to Fig ??, but for the IUP and KNMI algorithms. There appears to be no clear temporal variations with medians within ±one quartile. The same conclusions are drawn when comparing to CALIOP geometric mean heights.~~

For the full data set mean height differences vary between -0.607 and 0.243 km (geometric mean) and -0.053 and 0.785 km (cumulative extinction) for the IASI algorithms and -1.393 and -1.097 km (geometric mean) and -0.961 and -0.818 km (cumulative extinction) for the solar algorithms, Table 3. ~~In apricot color boxes in Table 3 are given the~~ The percentage of retrieved heights from the passive sensors that are within the dust layer as seen by CALIOP, are given in Table 3. Here the CALIOP dust layer is identified as the lowermost and uppermost heights identified as dust. ~~The highest percentage for~~ For the IASI algorithms between 75.9 and 85.8% of the retrieved heights are within the CALIOP dust layer. The highest percentage is achieved by ~~LMD~~ LISA during night over the ocean with up to ~~68.8~~96.9 % (geometric mean) and ~~48.5~~97.2 % (cumulative extinction) dust heights located within the CALIOP dust layer ~~during night over the land~~ (for a subset of respectively ~~206 and 177~~ 319 and 357 points).

The average CALIOP dust layer thickness is ~~0.925~~2.35 km over land during day and ~~0.859~~2.54 km over the ocean, Table 4. For the night the layer thickness is ~~1.292~~3.55 km over land and ~~0.924~~3.70 km over ocean. Thus, over land the dust layer thickness is larger during night than day by about ~~0.367~~1.12 km and 1.16 km ~~while it is similar~~ over the ocean. The dust layer over land is about ~~0.700-1.400~~0.680-1.42 km higher over land than ocean~~and it~~ . It is lower by ~~about 0.700~~0.52-1.02 km during the night than day over the ocean. This is mainly caused by different regions being sampled during night and day overpasses. Most of the concurrent IASI and CALIOP night data are from the Persian Gulf and the Red Sea (lower dust height) while the day time data are more evenly distributed over the study area.

4 Discussion

We have compared dust layer heights from various passive sensors with CALIOP derived heights. The CALIOP heights are considered as the "true" values. However, the CALIOP heights are not unique as described in section 2.1.1, thus we have used two different CALIOP derived heights. For the cumulative extinction CALIOP height method, the lidar ratio is involved. This may be different for different regions and time of year. Thus adding to the uncertainty in the comparison. The CALIOP analysis may also ~~misclassify~~ mis-classify aerosol as discussed by Kim et al. (2013). The latter is largely avoided in this study by focusing on dust aerosol which have a relatively large depolarization ratio. Different methods to calculate CALIOP heights are compared in Fig. 1. The RMSEs for the height methods compared are 0.652 km and 0.182 km. These numbers should be kept in mind for the comparison results presented above.

While comparable, the heights retrieved from CALIOP and IASI are not the same quantities due to the instruments different sensitivities to various aerosol particle sizes and the assumptions of aerosol optical properties (lidar ratio, refractive index, particle shape) used in the retrieval. A full understanding of the reason for the differences requires a detailed algorithm comparison which is beyond the scope of this study. It is noted that infrared sensors have lower sensitivity to low height dust caused by the small temperature difference between the temperature of the surface and the temperature of the dust. For example for the BIRA-IASB algorithm the lowest possible retrieval height is around 1.2 km due to low sensitivity to dust at lower height. For the DLR algorithm a positive bias with respect to the CALIOP cumulative extinction height was predicted, section 2.2.2. A positive bias between 0.0405 (day) and 0.906 km (night) is indeed found over land surfaces, see Table 3.

Overall the standard deviation of the difference between CALIOP heights and the passive sensor heights, is smaller for the night time data than the day time data, Table 3. This is most likely due to less noise in the CALIOP night data. Standard deviations are generally similar for ocean and land data, but there are differences for individual algorithms indicating opportunities for future improvements.

There is quite a large difference between day and night over the ocean and all algorithms overestimate more at night than day over ocean, Table 3. Due to the differences in satellite overpass times different regions are sampled during night and day overpasses. Most of the concurrent IASI and CALIOP night data are from the Persian Gulf and the Red Sea (lower dust height) while the day time data are more evenly distributed over the study area. The differences seen between night and day data may thus be caused by differences in optical properties of the dust between the two regions, which is not accounted for by the retrieval algorithms.

The CALIOP heights are moved to the SCIAMACHY and ~~MetOp-A~~ GOME-2 overpass times. On average the vertical shift is small being between 0.015 and 0.020 km with a standard deviation of 0.25 km. For individual data points the shift may be larger, compare ~~red and green crosses and lime and deep pink plusses in bottom panels of Figs. ??-??~~shifted and unshifted black and red triangles in Fig. 4. This suggests that when comparing data sets from satellite sensors with different overpass times, transport processes should be accounted for in the analysis.

~~The~~ Moreover, the spatial resolutions of the IASI, GOME-2 and SCIAMACHY instruments are much coarser than CALIOP. The impact of differences in spatial ~~reolution~~ resolution has not been investigated, but it is assumed to be small within large dust clouds as studied here.

It is not straigthforward to estimate an uncertainty for the IASI height retrievals as this would require a sensitivity study that is beyond the scope of this work. A best-guess estimate would be that the uncertainty is on the order of 1-1.5 km.

Vandenbussche et al. (2013) found that for low dust loads, the BIRA-IASB algorithm placed the aerosol layer 1-2 km above the CALIOP retrieved layer. The algorithm has since then undergone several revisions and improvements and the average overestimate for all data is 0.078 km (0.590 km) when compared with the CALIOP geometric mean (cumulative extinction) height. Possible reasons for this overestimate is discussed above. Vandenbussche et al. (2013) reported better agreement for moderate to higher dust loads compared to low dust loads. In the present study no effect of dust load on dust height agreement appears to be present, see centre row plots of Figs. A1 and A5.

For monthly mean $1^\circ \times 1^\circ$gridded IASI data covering the period July 2007-June 2013 Capelle et al. (2014) reported a systematic IASI-CALIOP bias of 0.4 km with a standard deviation of 0.48 km over the ocean. Peyridieu et al. (2013) reported similar values for the same data set. In this study for LMD over ocean a bias of -0.922 km (-0.501 km) against the CALIOP geometric mean (cumulative extinction) height is found for data recorded during the day overpasses with a standard deviation of 1.142 km (1.409 km). For the night overpasses the differences are -0.674 km (0.352 km) and the standard deviation 0.878 (1.180), Table 3. One reason for the larger spread in this study may be due to the use of monthly and spatially averaged data by Peyridieu et al. (2013) and Capelle et al. (2014) while here the comparison is made on a pixel by pixel basis. Hence extreme values are not averaged out.

Overall, for the IASI algorithms, two algorithms (BIRA-IASB and LISA) agree better with the CALIOP geometric mean height while LMD agree better with the CALIOP cumulative extinction height, Table 3. The DLR algorithm generally gives altitudes at two distinct modes, but overall agrees better with the geometric mean height. This may indicate that the IASI algorithms do not provide the same height information. The BIRA-IASB and LISA algorithms retrieve an aerosol profile from which a dust height is calculated. The LMD algorithm, on the other hand, use a single layer aerosol in the retrieval while DLR estimate the altitude from the retrieved dust layer temperature. The comparison with the two CALIOP heights suggests that the profile retrieval is generally more sensitive to the actual dust layer vertical location. Both the BIRA-IASB and LISA algorithms use 1 km vertical steps. But with 1.5-2 degrees of freedom there is a significant correlation between the layers, and therefore a low sensitivity to the actual high resolution vertical distribution represented in the cumulative extinction height. Mean altitudes from those retrievals would then be something resembling a "geometric mean height". Contrary, the LMD algorithm, which places the aerosol in a single homogeneous layer, is more sensitive to the aerosol layer radiatively “effective” height. It is also important to note that the sensitivity of the IASI algorithms does not only depend on aerosol load, but also on the temperature profile.

The BIRA-IASB algorithm use CALIOP profiles as a priori which implies that the BIRA-IASB altitude data include information about the CALIOP data against which it is compared. However, the (monthly) a priori is averaged over a large spatial area ($5^\circ \times 5^\circ$) therefore including measurements from different days and most probably even different dust events. Furthermore,

the retrievals usually have a significant departure with respect to the a priori. Thus the a priori profile used for a single retrieval is only vaguely related to the exact CALIOP profile used for the validation. The LISA algorithm also uses an a priori profile of dust derived from a CALIOP climatology, but it is a unique a priori profile for all retrievals. Therefore, it is not related with the CALIOP measurements used in the validation.

It is noted that all IASI algorithms assume the dust particles to be spherical. Klüser et al. (2016) compared optical properties of spherical and non-spherical dust particles. They found the values of the dust single scattering albedo to be different for spherical and non-spherical dust particles. This may potentially affect the dust height retrieval. It is beyond this study to investigate and quantify this effect.

The heights from the passive solar IUP-SCIAMACHY and KNMI-GOME-2 algorithms are generally low compared with the CALIOP height, Table 3. While IASI is mainly sensitive to the aerosol coarse mode, SCIAMACHY and GOME-2 are sensitive to both the fine and coarse modes. The height retrieved from these sensors depends on whether the surface albedo is retrieved simultaneously or not as shown by Sanders et al. (2015). They found that fixing the albedo in the retrieval gave a lower dust height than when retrieving both the albedo and the dust height. Fixing the albedo also gave better agreement with lidar measurements for the 16 scenes they analysed. ~~We~~ Dubuisson et al. (2009) made sensitivity studies for the retrieval of aerosol height from POLarization and Directionality of the Earth's Reflectances (POLDER) and MEdium Resolution Imaging Spectrometer (MERIS) oxygen A-band measurements. They showed that aerosol height estimates vary with aerosol optical depth (AOD), single scattering albedo, aerosol phase function, aerosol layer height, and the underlying surface albedo. For low surface albedo theoretical analysis gave errors of about $\pm$0.5 km and $\pm$0.2 km for POLDER and MERIS, respectively. Comparison between POLDER and CALIOP gave standard deviations less than about 0.55 km consistent with the theoretical analysis. However, for parts of the three cases of coast Africa, the aerosol height was underestimated by up to 1-2 km . Dubuisson et al. (2009) attributed this to either a more complex vertical aerosol structure including a layer near the surface or the presence of low clouds under the aerosol layer. The theoretical sensitivity results of Dubuisson et al. (2009) was confirmed by Kokhanovsky and Rozanov (2010) whose modelling sensitivity study indicated aerosol heights within $\pm$0.5 km if the aerosol single scattering albedo used in the retrieval deviated by less then 0.01 from the actual value of 0.99. They also reported that error increased with dust layer top height.

It must be stressed that the reported errors on dust height, based on synthetic sensitivity studies, are estimated with the assumption that either AOD, optical model, or surface reflectance are perfectly known beforehand or are derived from independent instrument channels. For instance, Dubuisson et al. (2009) make first use of the official POLDER and MERIS AOD products, while focusing on dark ocean surfaces only. Then, they find the most accurate aerosol model (being this pure dust or a mixture with sea salt or biomass burning) by perturbing the reflectance in a non-absorbing channel. Conversely, the solar algorithms of this work have been designed to fit the oxygen spectrum to concurrently infer dust height and optical thickness together, so that AOD and height uncertainties cannot be decoupled and deviations of the assumed optical model and climatological surface reflectivity from the actual ones contribute to the overall error budget. As such, the evaluation of the presented dust cases can be regarded as a more comprehensive test bed for operational dust height retrievals.

A similar approach has been devised by Xu et al. (2017) . Combining oxygen A and B-band measurements from the Earth Polychromatic Imaging Camera (EPIC) on the Deep Space Climate Observatory (DSCOVR), they retrieved AOD and altitude, finding that 71.5% and 98.7% aerosol heights were respectively within $\pm$0.5 km and $\pm$1.0 km envelopes, when compared to CALIOP for two overpasses of Saharan dust events over water only. Their reported RMS error, in this case, amounts to 0.45 km, pointing to the advantage of adding in the retrieval the information concealed in the B-band.

Generally, we found the IUP-SCIAMACHY and KNMI-GOME-2 algorithm retrieved heights to be low by -1.097 km (-0.961) and -1.393 km (-0.818) respectively when compared with CALIOP geometric mean (cumulative extinction) heights. For the KNMI-GOME-2 algorithm the underestimate is larger over ocean -2.015 km (-1.477) than over land -0.893 km (-0.229), Table 3. These differences are larger than those reported in the above cited studies. Still, it is found that for KNMI-GOME-2 (IUP-SCIAMACHY) between 63.7-67-0% (40.9-45.7%) of the aerosol heights are within the CALIOP aerosol layer. For KNMI more of the retrieved heights are within the CALIOP aerosol layer over land than ocean, Table 3.

In general, possible reasons for the underestimation of layer height by the solar sensors are the local optical dust properties and the surface reflectivity assumed in the forward model. While it has been already demonstrated that a positive deviation of the true surface albedo from the assumed prior value leads to an underestimation of layer height (Sanders et al., 2015), the similar tendency of lower retrievals by GOME-2 and SCIAMACHY ~~suggest~~ suggests that the influence of a wrongly prescribed aerosol model can be ruled out. This is because the KNMI/GOME-2 algorithm uses a Henyey-Greenstein phase function whereas IUP-SCIAMACHY ingests spectrally resolved T-matrix calculations of the phase matrix representing aspherical dust particles (see Table 1).

To this end, we note that the algorithms of the solar spectrometers assume that satellite pixels are fully covered by dust. Because of their coarse footprint sizes, this condition can ~~be frequently not satisfied.~~ frequently not be satisfied. The EPIC pixel size is 12$\times$12 km^2 compared to 80$\times$40 km^2 for GOME-2 and 30$\times$60 km^2 for SCIAMACHY, thus the results of Xu et al. (2017) are likely less affected by cloud contamination and aerosol inhomogeneities. A situation of partially aerosol-covered pixels implies that less oxygen molecules are shielded by the intervening scatterers, with the effect of increasing absorption inside the A-band sensed by the instruments. This effect is even more pronounced closer to the ground, where the majority of oxygen molecules reside. Since most of the information content on the height of the aerosol layer is carried by the in-band wavelengths of the A-band (about 760 nm) ratioed to the continuum (758 nm), it can be deduced that a dust pixel fraction smaller than one will lead also to an additional underestimation of layer height.

We end the discussion by listing several questions left open by this study. These questions may broadly be divided into two sets: 1) questions requiring analysis of larger dataset to consolidate findings; and 2) questions requiring a more detailed analysis to better understand reasons for differences. Specific open questions are:

1A: How will the results change when including other types of aerosol in the analysis?

1B: How will a larger data set in time and space affect the results?

1C: May an optimal aerosol height algorithm covering all situations be developed?

2A: What are the physical reasons for the differences between the IASI-algorithms?

2B: What are the physical reasons for the differences between the solar algorithms?

2C: What are the physical reasons for differences between the quantities estimated by the IR and solar algorithms?

2D: How may synthetic data be used to understand and evaluate the various algorithms?

2E: Which pixel-level uncertainties can we estimate to the layer height results of each algorithm (based on studies 2A-2D)?

5 Conclusions

As part of the ESA Aerosol_cci project dust aerosol heights retrieved from passive infrared and solar sensors using different algorithms have been compared with two different CALIOP derived dust layer heights. The comparison was made on a pixel by pixel basis for the IASI, GOME-2 and SCIAMACHY sensors for four dust episodes in 2010. Time differences between the overpass of CALIOP and the passive sensors were accounted for by shifting the CALIOP heights to the location of the pixels of the passive sensors using the FLEXTRA trajectory model.

As it is not possible to construct a unique dust layer height from CALIOP data, two CALIOP derived layer heights were used: the cumulative extinction height which is set to the height where the CALIOP extinction column is half of the total extinction column; and the geometric mean height which is defined to be the geometrical mean of the top and bottom heights of the dust layer.

Four algorithms (BIRA-IASB, DLR, LMD, LISA) retrieved dust heights from IASI spectra. The mean difference between the IASI heights and the CALIOP geometric mean (cumulative extinction) heights were found to vary between -0.635 and 0.087 km (-0.225 and 0.405 km) over land during day. For night time overpasses the values were between -0.579 ~~-0.358~~and 0.358 km (0.073 ~~-0.906~~and 0.906 km). Over the ocean day differences were between -0.922 ~~-0.340~~and 0.340 km (-0.501 ~~-0.913~~and 0.913 km) and night time differences between -0.674 ~~-0.835~~and 0.835 km (0.352 ~~-1.599~~and 1.599 km). Standard deviations were between 1.322-1.572 km (1.448-1.665 km) over land during day and decreasing to 0.855-1.058 km (0.896-1.092 km) during night. Over the ocean the standard deviation decreased from 1.142-1.187 km (0.913-1.539 km) during the day to 0.486-0.878 km (0.637-1.180 km) at night.

Two of the IASI algorithms (BIRA-IASB and LISA) were found to agree better with the CALIOP geometric mean height (BIRA-IASB: 0.078 km (cumulative extinction 0.590 km); LISA: -0.045 km (0.507 km)) while the LMD algorithm agreed better with the CALIOP cumulative extinction height: -0.053 km (geometric mean: -0.607 km). This is believed to be caused by the differences in the aerosol profile used for the radiative transfer simulations: BIRA-IASB and LISA use and retrieve vertically extended and resolved profiles while LMD place all the aerosols in one single homogeneous layer.

Far fewer data points were available for the solar sensors due to their larger pixel size and lack of night time data. The heights retrieved from the solar sensors on the mean underestimate the CALIOP geometric mean (cumulative extinction) heights by -1.393 km (-0.818 km) (KNMI, GOME-2) and -1.097 km (-0.961 km) (IUP, SCIAMACHY). This may be caused by the large pixel size and the assumption in the retrieval that the pixels are fully covered by aerosol.

The IASI instrument was first flown in 2006 and was the first of several to be launched. Thus data from IASI has the potential to provide long global time series of ECVs. There is considerable variation between the IASI retrieved dust heights. Nevertheless, if careful consideration is taken of differences in temporal and spatial characteristics of the observations, it might be feasible to construct a global data set of IASI retrieved heights quality controlled against CALIOP. The quality control will allow uncertainties on a pixel by pixel basis which again may be used for sensitive studies. Such a dust height data set may be used to further our understanding of dust on the climate system. However, several open questions should be answered to have a better understanding of the quantities measured and their accuracy. A list of open questions are given at the end of the discussion section and includes both studies requiring large data sets and time periods, and studies looking at algorithm specifics.

Finally, the various algorithms and instruments are different in their approaches to retrieve the dust height. In the comparison with CALIOP no single algorithm is found to stand as the best overall. Different methodologies may give best results at different locations and situations. Thus it seems fruitful to continue development of all algorithms and encourage comparison exercises.

Code and data availability. All data are available to registered users from http://www.icare.univ-lille1.fr/. The FLEXTRA model is available from https://www.flexpart.eu/.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was supported by the European Space Agency as part of the Aerosol_cci project (ESA Contract No. 4000109874/14/I-NB). Comments to the manuscript by Andreas Stohl is greatly appreciated. Thanks to Sabine Eckhardt for help with running of the FLEXTRA model. Two anonymous referees are thanked for their constructive comments.

References

Amiridis, V., Marinou, E., Tsekeri, A., Wandinger, U., Schwarz, A., Giannakaki, E., Mamouri, R., Kokkalis, P., Binietoglou, I., Solomos, S., Herekakis, T., Kazadzis, S., Gerasopoulos, E., Proestakis, E., Kottas, M., Balis, D., Papayannis, A., Kontoes, C., Kourtidis, K., Papagiannopoulos, N., Mona, L., Pappalardo, G., Le Rille, O., and Ansmann, A.: LIVAS: a 3-D multi-wavelength aerosol/cloud database based on CALIPSO and EARLINET, Atmospheric Chemistry and Physics, 15, 7127–7153, https://doi.org/10.5194/acp-15-7127-2015, http://www.atmos-chem-phys.net/15/7127/2015/, 2015.

Anderson, G. P., Clough, S. A., Kneizys, F., Chetwynd, J. H., and Shettle, E. P.: AFGL atmospheric constituent profiles (0-120km), Environmental research papers (Hanscom AFB, Mass.) no 954, 1986.

August, T., Klaes, D., Schlüssel, P., Hultberg, T., Crapeau, M., Arriaga, A., O'Carroll, A., Coppens, D., Munro, R., and Calbet, X.: IASI on Metop-A: Operational Level 2 retrievals after five years in orbit, Journal of Quantitative Spectroscopy and Radiative Transfer, 113, 1340 – 1371, https://doi.org/http://dx.doi.org/10.1016/j.jqsrt.2012.02.028, //www.sciencedirect.com/science/article/pii/S0022407312000921, three Leaders in Spectroscopy, 2012.

Balkanski, Y., Schulz, M., Claquin, T., and Guibert, S.: Reevaluation of Mineral aerosol radiative forcings suggests a better agreement with satellite and AERONET data, Atmospheric Chemistry and Physics, 7, 81–95, https://doi.org/10.5194/acp-7-81-2007, http://www.atmos-chem-phys.net/7/81/2007/, 2007.

Banks, J., Brindley, H., Flamant, C., Garay, M., Hsu, N., Kalashnikova, O., Klüser, L., and Sayer, A.: Intercomparison of satellite dust retrieval products over the west African Sahara during the Fennec campaign in June 2011, Remote Sensing of Environment, 136, 99–116, https://doi.org/http://dx.doi.org/10.1016/j.rse.2013.05.003, 2013.

Baumann, K. and Stohl, A.: Validation of a Long-Range Trajectory Model Using Gas Balloon Tracks from the Gordon Bennett Cup 95, Journal of Applied Meteorology, 36, 711–720, https://doi.org/10.1175/1520-0450-36.6.711, https://doi.org/10.1175/1520-0450-36.6.711, 1997.

Capelle, V., Chédin, A., Péquignot, E., Schlüssel, P., Newman, S. M., and Scott, N. A.: Infrared Continental Surface Emissivity Spectra and Skin Temperature Retrieved from IASI Observations over the Tropics, Journal of Applied Meteorology and Climatology, 51, 1164–1179, https://doi.org/10.1175/JAMC-D-11-0145.1, http://dx.doi.org/10.1175/JAMC-D-11-0145.1, 2012.

Capelle, V., Chédin, A., Siméon, M., Tsamalis, C., Pierangelo, C., Pondrom, M., Crevoisier, C., Crepeau, L., and Scott, N. A.: Evaluation of IASI-derived dust aerosol characteristics over the tropical belt, Atmospheric Chemistry and Physics, 14, 9343–9362, https://doi.org/10.5194/acp-14-9343-2014, http://www.atmos-chem-phys.net/14/9343/2014/, 2014.

Chédin, A., Scott, N. A., Wahiche, C., and Moulinier, P.: The Improved Initialization Inversion Method: A High Resolution Physical Method for Temperature Retrievals from Satellites of the TIROS-N Series, Journal of Climate and Applied Meteorology, 24, 128–143, https://doi.org/10.1175/1520-0450(1985)024<0128:TIIIMA>2.0.CO;2, http://dx.doi.org/10.1175/1520-0450(1985)024<0128:TIIIMA>2.0.CO;2, 1985.

Chevallier, F., Chéruy, F., Scott, N. A., and Chédin, A.: A Neural Network Approach for a Fast and Accurate Computation of a Longwave Radiative Budget, Journal of Applied Meteorology, 37, 1385–1397, https://doi.org/10.1175/1520-0450(1998)037<1385:ANNAFA>2.0.CO;2, http://dx.doi.org/10.1175/1520-0450(1998)037<1385:ANNAFA>2.0.CO;2, 1998.

Cuesta, J., Marsham, J. H., Parker, D. J., and Flamant, C.: Dynamical mechanisms controlling the vertical redistribution of dust and the thermodynamic structure of the West Saharan atmospheric boundary layer during summer, Atmospheric Science Letters, 10, 34–42, https://doi.org/10.1002/asl.207, http://dx.doi.org/10.1002/asl.207, 2009.

Cuesta, J., Eremenko, M., Flamant, C., Dufour, G., Laurent, B., Bergametti, G., Höpfner, M., Orphal, J., and Zhou, D.: Three-dimensional distribution of a major desert dust outbreak over East Asia in March 2008 derived from IASI satellite observations, Journal of Geophysical Research: Atmospheres, 120, 7099–7127, https://doi.org/10.1002/2014JD022406, http://dx.doi.org/10.1002/2014JD022406, 2015.

de Graaf, M., Stammes, P., Torres, O., and Koelemeijer, R. B. A.: Absorbing Aerosol Index: Sensitivity analysis, application to GOME and comparison with TOMS, J. Geophys. Res., 110, https://doi.org/10.1029/2004JD005178, 2005.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, http://dx.doi.org/10.1002/qj.828, 2011.

Deirmendjian, D., Clasen, R., and Viezee, W.: Mie Scattering with Complex Index of Refraction, J. Opt. Soc. Am., 51, 620–633, https://doi.org/10.1364/JOSA.51.000620, http://www.osapublishing.org/abstract.cfm?URI=josa-51-6-620, 1961.

Di Biagio, C., Boucher, H., Caquineau, S., Chevaillier, S., Cuesta, J., and Formenti, P.: Variability of the infrared complex refractive index of African mineral dust: experimental estimation and implications for radiative transfer and satellite remote sensing, Atmospheric Chemistry and Physics, 14, 11 093–11 116, https://doi.org/10.5194/acp-14-11093-2014, https://www.atmos-chem-phys.net/14/11093/2014/, 2014.

Dubovik, O., Holben, B. N., Lapyonok, T., Sinyuk, A., Mishchenko, M. I., Yang, P., and Slutsker, I.: Non-spherical aerosol retrieval method employing light scattering by spheroids, Geophysical Research Letters, 29, 54–1–54–4, https://doi.org/10.1029/2001GL014506, http://dx.doi.org/10.1029/2001GL014506, 2002.

Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B., Mishchenko, M., Yang, P., Eck, T., Volten, H., Muñoz, O., Veihelmann, B., van der Zande, W., Léon, J.-F., Sorokin, M., and Slutsker, I.: Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust, Journal of Geophysical Research: Atmospheres, 111, D11 208, https://doi.org/10.1029/2005JD006619, 2006.

Dubuisson, P., Frouin, R., Dessailly, D., Duforêt, L., Léon, J.-F., Voss, K., and Antoine, D.: Estimating the altitude of aerosol plumes over the ocean from reflectance ratio measurements in the O_2 A-band, Remote Sens. Environ., 113, 1899–1911, 2009.

Hollmann, R., Merchant, C. J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., de Leeuw, G., Forsberg, R., Holzer-Popp, T., Paul, F., Sandven, S., Sathyendranath, S., van Roozendael, M., and Wagner, W.: The ESA Climate Change Initiative: Satellite Data Records for Essential Climate Variables, Bulletin of the American Meteorological Society, 94, 1541–1552, https://doi.org/10.1175/BAMS-D-11-00254.1, http://dx.doi.org/10.1175/BAMS-D-11-00254.1, 2013.

Jacquinet-Husson, N., Crepeau, L., Armante, R., Boutammine, C., Chédin, A., Scott, N., Crevoisier, C., Capelle, V., Boone, C., Poulet-Crovisier, N., Barbe, A., Campargue, A., Benner, D. C., Benilan, Y., Bézard, B., Boudon, V., Brown, L., Coudert, L., Coustenis, A., Dana, V., Devi, V., Fally, S., Fayt, A., Flaud, J.-M., Goldman, A., Herman, M., Harris, G., Jacquemart, D., Jolly, A., Kleiner, I., Kleinböhl, A., Kwabia-Tchana, F., Lavrentieva, N., Lacome, N., Xu, L.-H., Lyulin, O., Mandin, J.-Y., Maki, A., Mikhailenko, S., Miller, C., Mishina, T., Moazzen-Ahmadi, N., Müller, H., Nikitin, A., Orphal, J., Perevalov, V., Perrin, A., Petkie, D., Predoi-Cross, A., Rinsland, C., Remedios, J., Rotger, M., Smith, M., Sung, K., Tashkun, S., Tennyson, J., Toth, R., Vandaele, A.-C., and Auwera, J. V.: The 2009 edition of the {GEISA} spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 112, 2395 – 2445, https://doi.org/http://doi.org/10.1016/j.jqsrt.2011.06.004, http://www.sciencedirect.com/science/article/pii/S0022407311002160, 2011.

Kerker, M.: The Scattering of Light, and Other Electromagnetic Radiation: Physical Chemistry, vol. 666 of *Physical Chemistry, Monogr. Ser.*, Academic Press, New York, 1969.

Kim, M.-H., Kim, S.-W., Yoon, S.-C., and Omar, A. H.: Comparison of aerosol optical depth between CALIOP and MODIS-Aqua for CALIOP aerosol subtypes over the ocean, Journal of Geophysical Research: Atmospheres, 118, 13,241–13,252, https://doi.org/10.1002/2013JD019527, http://dx.doi.org/10.1002/2013JD019527, 2013JD019527, 2013.

Klüser, L., Martynenko, D., and Holzer-Popp, T.: Thermal infrared remote sensing of mineral dust over land and ocean: a spectral SVD based retrieval approach for IASI, Atmospheric Measurement Techniques, 4, 757–773, https://doi.org/10.5194/amt-4-757-2011, http://www.atmos-meas-tech.net/4/757/2011/, 2011.

Klüser, L., Kleiber, P., Holzer-Popp, T., and Grassian, V.: Desert dust observation from space - Application of measured mineral component infrared extinction spectra, Atmospheric Environment, 54, 419 – 427, https://doi.org/http://dx.doi.org/10.1016/j.atmosenv.2012.02.011, http://www.sciencedirect.com/science/article/pii/S1352231012001264, 2012.

Klüser, L., Banks, J., Martynenko, D., Bergemann, C., Brindley, H., and Holzer-Popp, T.: Information content of space-borne hyperspectral infrared observations with respect to mineral dust properties, Remote Sensing of Environment, 156, 294 – 309, https://doi.org/http://dx.doi.org/10.1016/j.rse.2014.09.036, http://www.sciencedirect.com/science/article/pii/S0034425714003927, 2015.

Klüser, L., Biagio, C. D., Kleiber, P. D., Formenti, P., and Grassian, V. H.: Optical properties of non-spherical desert dust particles in the terrestrial infrared - An asymptotic approximation approach, Journal of Quantitative Spectroscopy and Radiative Transfer, 178, 209 – 223, https://doi.org/http://dx.doi.org/10.1016/j.jqsrt.2015.11.020, //www.sciencedirect.com/science/article/pii/S0022407315300996, electromagnetic and light scattering by nonspherical particles XV: Celebrating 150 years of Maxwell's electromagnetics, 2016.

Koffi, B., Schulz, M., Br??on, F.-M., Griesfeller, J., Winker, D., Balkanski, Y., Bauer, S., Berntsen, T., Chin, M., Collins, W. D., Dentener, F., Diehl, T., Easter, R., Ghan, S., Ginoux, P., Gong, S., Horowitz, L. W., Iversen, T., Kirkev??g, A., Koch, D., Krol, M., Myhre, G., Stier, P., and Takemura, T.: Application of the CALIOP layer product to evaluate the vertical distribution of aerosols estimated by global models: AeroCom phase I results, Journal of Geophysical Research: Atmospheres, 117, n/a–n/a, https://doi.org/10.1029/2011JD016858, http://dx.doi.org/10.1029/2011JD016858, 2012.

Kokhanovsky, A. A. and Rozanov, V. V.: The determination of dust cloud altitudes from a satellite using hyperspectral measurements in the gaseous absorption band, International Journal of Remote Sensing, 31, 2729–2744, https://doi.org/10.1080/01431160903085644, https://doi.org/10.1080/01431160903085644, 2010.

Lelli, L., Sanders, A., Rozanov, V., Kokhanovsky, A., Jäger, M., and Burrows, J.: Science verification of aerosol layer height for Sentinel-5 Precursor: a multi-sensor volcanic case study, Atmospheric Measurement Techniques, in preparation, 2017.

Maes, K., Vandenbussche, S., Kl??ser, L., Kumps, N., and de Mazi??re, M.: Vertical Profiling of Volcanic Ash from the 2011 Puyehue Cord??n Caulle Eruption Using IASI, Remote Sensing, 8, 103, https://doi.org/10.3390/rs8020103, http://www.mdpi.com/2072-4292/8/2/103, 2016.

Massie, S.: Indices of refraction for the Hitran compilation, Journal of Quantitative Spectroscopy and Radiative Transfer, 52, 501–513, 1994.

Massie, S. and Goldman, A.: The infrared absorption cross-section and refractive-index data in HITRAN, Journal of Quantitative Spectroscopy and Radiative Transfer, 82, 413–428, 2003.

Merucci, L., Zak??ek, K., Carboni, E., and Corradini, S.: Stereoscopic Estimation of Volcanic Ash Cloud-Top Height from Two Geostationary Satellites, Remote Sensing, 8, 206, https://doi.org/10.3390/rs8030206, http://www.mdpi.com/2072-4292/8/3/206, 2016.

Metzig, G.: OPTIMA—Computation of the optical properties of single homogeneous or coated Mie-particles, vol. 87, Kernforschungszentrum Karlsruhe GmbH, Lab. fuer Aerosolphysik und Filtertechnik, 1984.

Newman, S. M., Smith, J. A., Glew, M. D., Rogers, S. M., and Taylor, J. P.: Temperature and salinity dependence of sea surface emissivity in the thermal infrared, Quarterly Journal of the Royal Meteorological Society, 131, 2539–2557, https://doi.org/10.1256/qj.04.150, http://dx.doi.org/10.1256/qj.04.150, 2005.

Omar, A. H., Winker, D. M., Vaughan, M. A., Hu, Y., Trepte, C. R., Ferrare, R. A., Lee, K.-P., Hostetler, C. A., Kittaka, C., Rogers, R. R., Kuehn, R. E., and Liu, Z.: The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm, Journal of Atmospheric and Oceanic Technology, 26, 1994–2014, https://doi.org/10.1175/2009JTECHA1231.1, http://dx.doi.org/10.1175/2009JTECHA1231.1, 2009.

Paul, M., Aires, F., Prigent, C., Trigo, I. F., and Bernardo, F.: An innovative physical scheme to retrieve simultaneously surface temperature and emissivities using high spectral infrared observations from IASI, Journal of Geophysical Research: Atmospheres, 117, n/a–n/a, https://doi.org/10.1029/2011JD017296, http://dx.doi.org/10.1029/2011JD017296, d11302, 2012.

Peyridieu, S., Chédin, A., Tanré, D., Capelle, V., Pierangelo, C., Lamquin, N., and Armante, R.: Saharan dust infrared optical depth and altitude retrieved from AIRS: a focus over North Atlantic –; comparison to MODIS and CALIPSO, Atmospheric Chemistry and Physics, 10, 1953–1967, https://doi.org/10.5194/acp-10-1953-2010, http://www.atmos-chem-phys.net/10/1953/2010/, 2010.

Peyridieu, S., Chédin, A., Capelle, V., Tsamalis, C., Pierangelo, C., Armante, R., Crevoisier, C., Crépeau, L., Siméon, M., Ducos, F., and Scott, N. A.: Characterisation of dust aerosols in the infrared from IASI and comparison with PARASOL, MODIS, MISR, CALIOP, and AERONET observations, Atmospheric Chemistry and Physics, 13, 6065–6082, https://doi.org/10.5194/acp-13-6065-2013, http://www.atmos-chem-phys.net/13/6065/2013/, 2013.

Pierangelo, C., Chédin, A., Heilliette, S., Jacquinet-Husson, N., and Armante, R.: Dust altitude and infrared optical depth from AIRS, Atmospheric Chemistry and Physics, 4, 1813–1822, https://doi.org/10.5194/acp-4-1813-2004, http://www.atmos-chem-phys.net/4/1813/2004/, 2004.

Pierangelo, C., Mishchenko, M., Balkanski, Y., and Chédin, A.: Retrieving the effective radius of Saharan dust coarse mode from AIRS, Geophysical Research Letters, 32, n/a–n/a, https://doi.org/10.1029/2005GL023425, http://dx.doi.org/10.1029/2005GL023425, l20813, 2005.

Popp, C., Wang, P., Brunner, D., Stammes, P., Zhou, Y., and Grzegorski, M.: MERIS albedo climatology for FRESCO+ O_2 A-band cloud retrieval, Atmospheric Measurement Techniques, 4, 463–483, https://doi.org/10.5194/amt-4-463-2011, http://www.atmos-meas-tech.net/4/463/2011/, 2011.

Popp, T., de Leeuw, G., Bingen, C., Br??hl, C., Capelle, V., Chedin, A., Clarisse, L., Dubovik, O., Grainger, R., Griesfeller, J., Heckel, A., Kinne, S., Kl??ser, L., Kosmale, M., Kolmonen, P., Lelli, L., Litvinov, P., Mei, L., North, P., Pinnock, S., Povey, A., Robert, C., Schulz, M., Sogacheva, L., Stebel, K., Stein Zweers, D., Thomas, G., Tilstra, L. G., Vandenbussche, S., Veefkind, P., Vountas, M., and Xue, Y.: Development, Production and Evaluation of Aerosol Climate Data Records from European Satellite Observations (Aerosol_cci), Remote Sensing, 8, 421, https://doi.org/10.3390/rs8050421, http://www.mdpi.com/2072-4292/8/5/421, 2016.

Riddle, E. E., Voss, P. B., Stohl, A., Holcomb, D., Maczka, D., Washburn, K., and Talbot, R. W.: Trajectory model validation using newly developed altitude-controlled balloons during the International Consortium for Atmospheric Research on Transport and Transformations 2004 campaign, Journal of Geophysical Research: Atmospheres, 111, n/a–n/a, https://doi.org/10.1029/2006JD007456, http://dx.doi.org/10.1029/2006JD007456, d23S57, 2006.

Rodgers, C. D.: Inverse methods for atmospheric sounding, Theory and practice, World Scientific Publishing Co. Ptc. Ltd., 2000.

Rothman, L., Gordon, I., Barbe, A., Benner, D., Bernath, P., Birk, M., Boudon, V., Brown, L., Campargue, A., Champion, J.-P., Chance, K., Coudert, L., Dana, V., Devi, V., Fally, S., Flaud, J.-M., Gamache, R., Goldman, A., Jacquemart, D., Kleiner, I., Lacome, N., Lafferty, W., Mandin, J.-Y., Massie, S., Mikhailenko, S., Miller, C., Moazzen-Ahmadi, N., Naumenko, O., Nikitin, A., Orphal, J., Perevalov, V., Perrin,

A., Predoi-Cross, A., Rinsland, C., Rotger, M., Šimečková, M., Smith, M., Sung, K., Tashkun, S., Tennyson, J., Toth, R., Vandaele, A., and Auwera, J. V.: The HITRAN 2008 molecular spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 110, 533 – 572, https://doi.org/10.1016/j.jqsrt.2009.02.013, 2009.

Rozanov, V.: Adjoint radiative transfer equation and inverse problems, In Light Scattering Reviews, A.A. Kokhanovsky (Ed.), pp. 339-392 (Berlin: Springer-Praxis)., 2006.

Rozanov, V., Rozanov, A., and Kokhanovsky, A.: Derivatives of the radiative field and their application to the solution of inverse problems, In Light Scattering Reviews, A.A. Kokhanovsky (Ed.), pp. 205-268 (Berlin: Springer-Praxis)., 2007.

Rozanov, V. V., Rozanov, A. V., Kokhanovsky, A. A., and Burrows, J. P.: Radiative transfer through terrestrial atmosphere and ocean: Software package SCIATRAN, Journal of Quantitative Spectroscopy and Radiative Transfer, 133, 13 – 71, https://doi.org/10.1016/j.jqsrt.2013.07.004, 2014.

Sanders, A. F. J. and de Haan, J. F.: Retrieval of aerosol parameters from the oxygen A band in the presence of chlorophyll fluorescence, Atmospheric Measurement Techniques, 6, 2725–2740, https://doi.org/10.5194/amt-6-2725-2013, http://www.atmos-meas-tech.net/6/2725/2013/, 2013.

Sanders, A. F. J., de Haan, J. F., Sneep, M., Apituley, A., Stammes, P., Vieitez, M. O., Tilstra, L. G., Tuinder, O. N. E., Koning, C. E., and Veefkind, J. P.: Evaluation of the operational Aerosol Layer Height retrieval algorithm for Sentinel-5 Precursor: application to O_2 A band observations from GOME-2A, Atmospheric Measurement Techniques, 8, 4947–4977, https://doi.org/10.5194/amt-8-4947-2015, http://www.atmos-meas-tech.net/8/4947/2015/, 2015.

Schlundt, C., Kokhanovsky, A. A., von Hoyningen-Huene, W., Dinter, T., Istomina, L., and Burrows, J. P.: Synergetic cloud fraction determination for SCIAMACHY using MERIS, Atmospheric Measurement Techniques, 4, 319–337, https://doi.org/10.5194/amt-4-319-2011, http://www.atmos-meas-tech.net/4/319/2011/, 2011.

Scott, N. A. and Chédin, A.: A Fast Line-by-Line Method for Atmospheric Absorption Computations: The Automatized Atmospheric Absorption Atlas, Journal of Applied Meteorology, 20, 802–812, https://doi.org/10.1175/1520-0450(1981)020<0802:AFLBLM>2.0.CO;2, http://dx.doi.org/10.1175/1520-0450(1981)020<0802:AFLBLM>2.0.CO;2, 1981.

Spurr, R., de Haan, J. D., van Oss, R., and Vasilkov, A.: Discrete ordinate radiative transfer in a stratified medium with first order rotational Raman scattering, J. Quant. Spectrosc. Radiat. Transfer, 109, 404–425, 2008.

Stiller, G. P., ed.: The Karlsruhe Optimized and Precise Radiative Transfer Algorithm (KOPRA), vol. FZKA 6487 of *Wissenschaftliche Berichte*, Forschungszentrum Karlsruhe, 2000.

Stohl, A. and Koffi, N. E.: Evaluation of trajectories calculated from ecmwf data against constant volume balloon flights during etex, Atmospheric Environment, 32, 4151 – 4156, https://doi.org/https://doi.org/10.1016/S1352-2310(98)00185-X, http://www.sciencedirect.com/science/article/pii/S135223109800185X, 1998.

Stohl, A. and Seibert, P.: Accuracy of trajectories as determined from the conservation of meteorological tracers, Quarterly Journal of the Royal Meteorological Society, 124, 1465–1484, https://doi.org/10.1002/qj.49712454907, http://dx.doi.org/10.1002/qj.49712454907, 1998.

Stohl, A., Wotawa, G., Seibert, P., and Kromp-Kolb, H.: Interpolation Errors in Wind Fields as a Function of Spatial and Temporal Resolution and Their Impact on Different Types of Kinematic Trajectories, Journal of Applied Meteorology, 34, 2149–2165, https://doi.org/10.1175/1520-0450(1995)034<2149:IEIWFA>2.0.CO;2, http://dx.doi.org/10.1175/1520-0450(1995)034<2149:IEIWFA>2.0.CO;2, 1995.

Tikhonov, A.: On the solution of incorrectly stated problems and a method of regularization, Dokl. Acad. Nauk SSSR, 151, 501–504, 2003.

Vandenbussche, S. and De Mazière, M.: African mineral dust sources: a combined analysis based on 3D dust aerosols distributions, winds and surface parameters, Atmospheric Chemistry and Physics Discussions, 2017, 1–37, https://doi.org/10.5194/acp-2017-809, https://www.atmos-chem-phys-discuss.net/acp-2017-809/, 2017.

Vandenbussche, S., Kochenova, S., Vandaele, A. C., Kumps, N., and De Maziére, M.: Retrieval of desert dust aerosol vertical profiles from IASI measurements in the TIR atmospheric window, Atmospheric Measurement Techniques, 6, 2577–2591, https://doi.org/10.5194/amt-6-2577-2013, http://www.atmos-meas-tech.net/6/2577/2013/, 2013.

Veefkind, J., Aben, I., McMullan, K., F??rster, H., de Vries, J., Otter, G., Claas, J., Eskes, H., de Haan, J., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P.: {TROPOMI} on the {ESA} Sentinel-5 Precursor: A {GMES} mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sensing of Environment, 120, 70 – 83, https://doi.org/http://dx.doi.org/10.1016/j.rse.2011.09.027, http://www.sciencedirect.com/science/article/pii/S0034425712000661, the Sentinel Missions - New Opportunities for Science, 2012.

Virtanen, T. H., Kolmonen, P., Rodríguez, E., Sogacheva, L., Sundström, A.-M., and de Leeuw, G.: Ash plume top height estimation using AATSR, Atmospheric Measurement Techniques, 7, 2437–2456, https://doi.org/10.5194/amt-7-2437-2014, http://www.atmos-meas-tech.net/7/2437/2014/, 2014.

Volz, F. E.: Infrared absorption by atmospheric aerosol substances, Journal of Geophysical Research, 77, 1017–1031, https://doi.org/10.1029/JC077i006p01017, http://dx.doi.org/10.1029/JC077i006p01017, 1972.

Volz, F. E.: Infrared Optical Constants of Ammonium Sulfate, Sahara Dust, Volcanic Pumice, and Flyash, Appl. Opt., 12, 564–568, 1973.

Wang, P., Tuinder, O. N. E., Tilstra, L. G., de Graaf, M., and Stammes, P.: Interpretation of FRESCO cloud retrievals in case of absorbing aerosol events, Atmospheric Chemistry and Physics, 12, 9057–9077, https://doi.org/10.5194/acp-12-9057-2012, http://www.atmos-chem-phys.net/12/9057/2012/, 2012.

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms, Journal of Atmospheric and Oceanic Technology, 26, 2310–2323, https://doi.org/10.1175/2009JTECHA1281.1, http://dx.doi.org/10.1175/2009JTECHA1281.1, 2009.

Winker, D. M., Tackett, J. L., Getzewich, B. J., Liu, Z., Vaughan, M. A., and Rogers, R. R.: The global 3-D distribution of tropospheric aerosols as characterized by CALIOP, Atmospheric Chemistry and Physics, 13, 3345–3361, https://doi.org/10.5194/acp-13-3345-2013, http://www.atmos-chem-phys.net/13/3345/2013/, 2013.

Xu, X., Wang, J., Wang, Y., Zeng, J., Torres, O., Yang, Y., Marshak, A., Reid, J., and Miller, S.: Passive remote sensing of altitude and optical depth of dust plumes using the oxygen A and B bands: First results from EPIC/DSCOVR at Lagrange-1 point, Geophysical Research Letters, 44, 7544–7554, https://doi.org/10.1002/2017GL073939, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL073939, 2017.

Xu, X., Wang, J., Wang, Y., and Kokhanovsky, A.: Chapter 1 - Passive Remote Sensing of Aerosol Height, in: Remote Sensing of Aerosols, Clouds, and Precipitation, edited by Islam, T., Hu, Y., Kokhanovsky, A., and Wang, J., pp. 1 – 22, Elsevier, https://doi.org/https://doi.org/10.1016/B978-0-12-810437-8.00001-3, https://www.sciencedirect.com/science/article/pii/B9780128104378000013, 2018.

Zhou, D., Larar, A., Liu, X., Smith, W., Strow, L., Yang, P., Schlüandssel, P., and Calbet, X.: Global Land Surface Emissivity Retrieved From Satellite Ultraspectral IR Measurements, Geoscience and Remote Sensing, IEEE Transactions on, 49, 1277–1290,

https://doi.org/10.1109/TGRS.2010.2051036, http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5523979&isnumber=5738422, 2011.

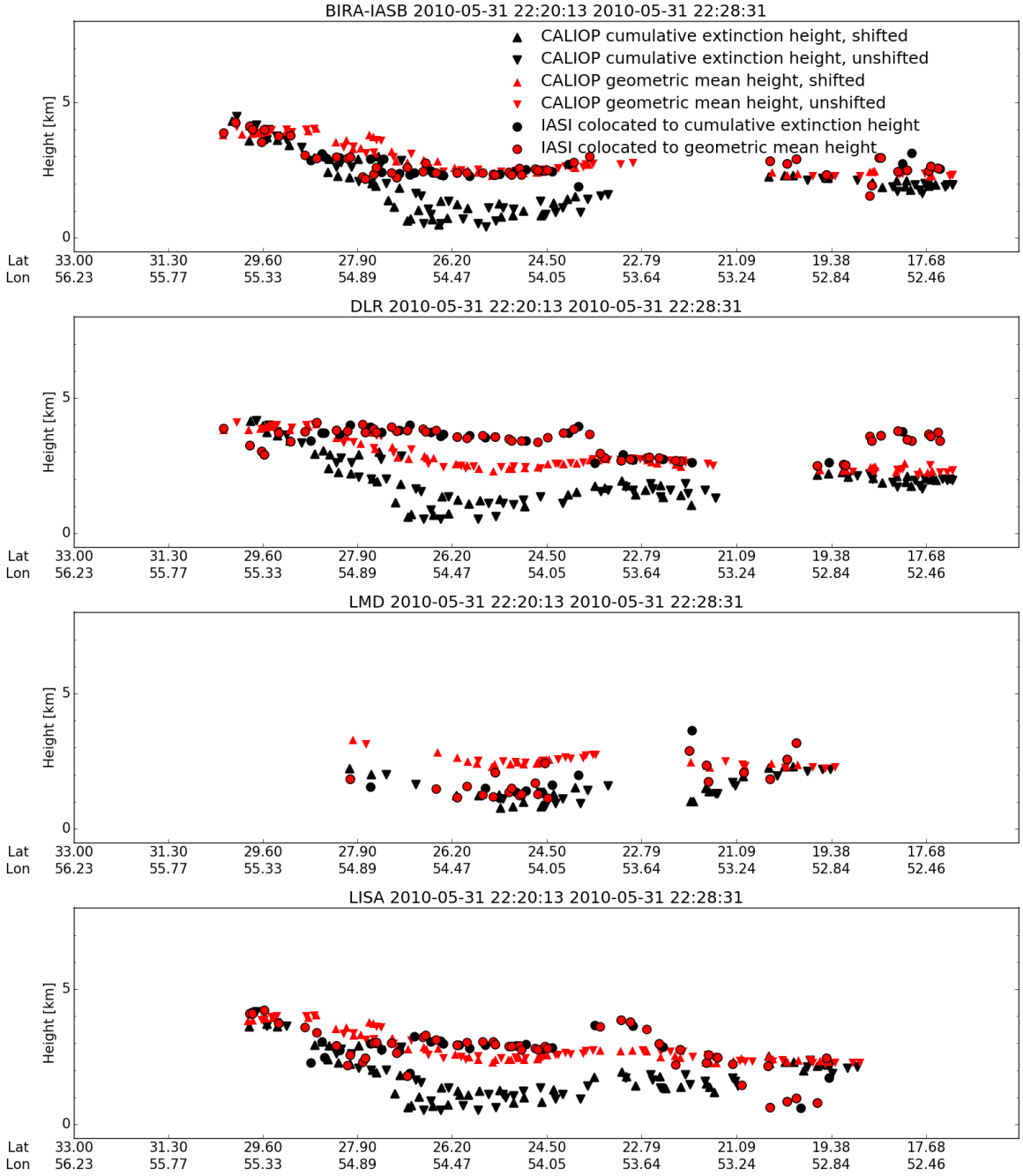

Figure 4. ~~Similar to Fig. ??, but SCIAMACHY~~ IASI dust ~~height from IUP analysis. Note difference in~~ layer heights colocated to CALIOP cumulative extinction heights (black circles) and CALIOP geometric mean heights (red circles) for the same time ~~from~~ and location as in Fig. ~~??~~3. ~~The curtain in the bottom plot is for the~~ Also shown are shifted (upward triangles) and unshifted (downward triangles) CALIOP ~~data between 20-40°E in top plot~~cumulative extinction and geometric mean heights.

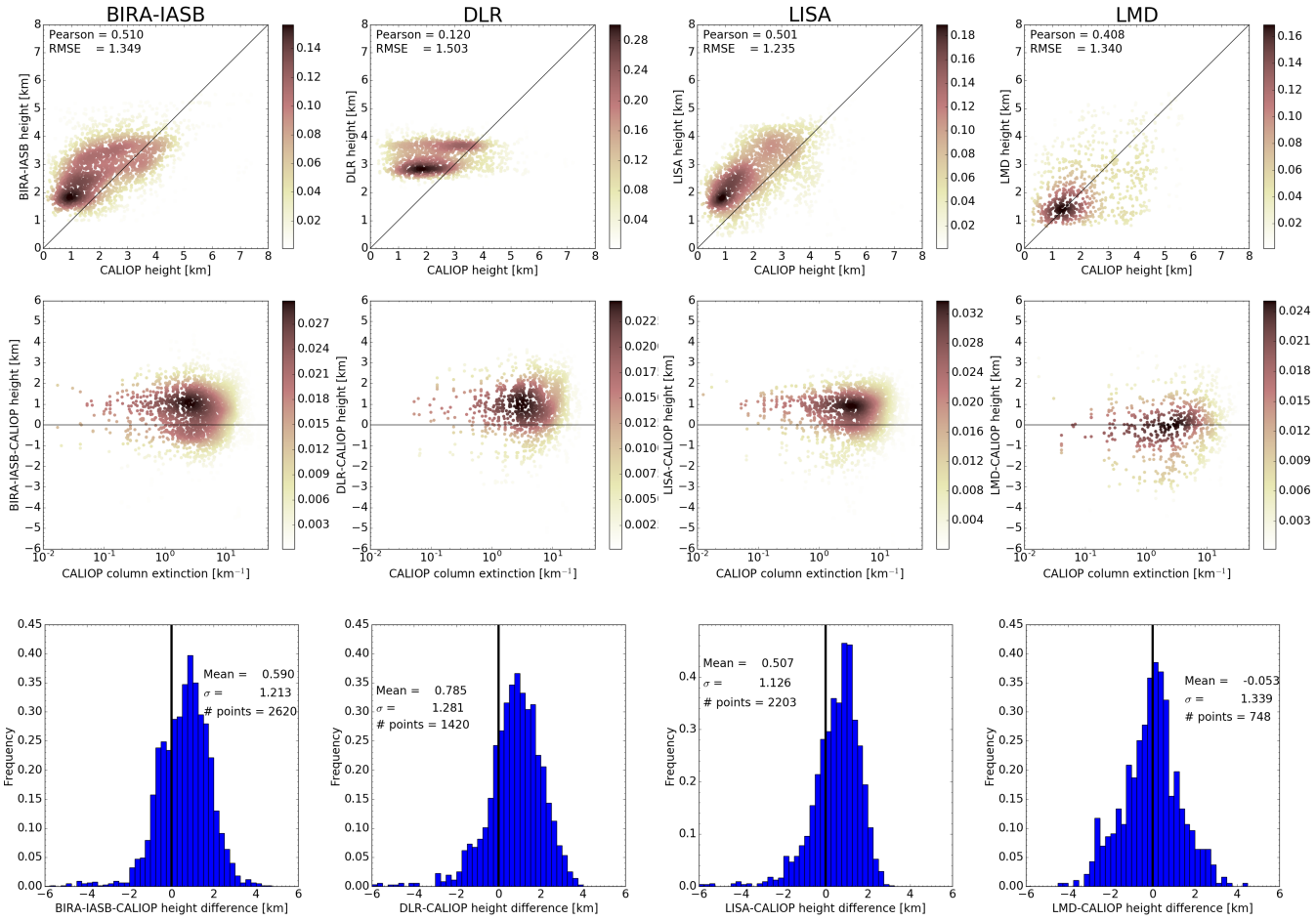

Figure 5. (Upper row) ~~Scatter plots~~ The probability density, using kernel density estimation, of the CALIOP cumulative extinction height versus height from the various algorithms. ~~The color indicates the density of points.~~ Also given are the Pearson's correlation coefficient and root mean square error (RMSE). (Centre row) The probability density, using kernel density estimation, of the difference between the passive algorithm and the CALIOP cumulative extinction heights versus the CALIOP column extinction. ~~The color indicates the density of points.~~ (Bottom row) Frequency distribution of the difference between the height from the various algorithms and the CALIOP cumulative extinction height. ~~Included is also a normal distribution fit to the difference.~~ The mean and standard deviation (σ) ~~of the normal distribution~~ together with the number data points are given in each plot. This information is also provided in Table 3. Data are shown for the BIRA-IASB (first column), DLR (second column), LISA (third column) and LMD (fourth column) algorithms.

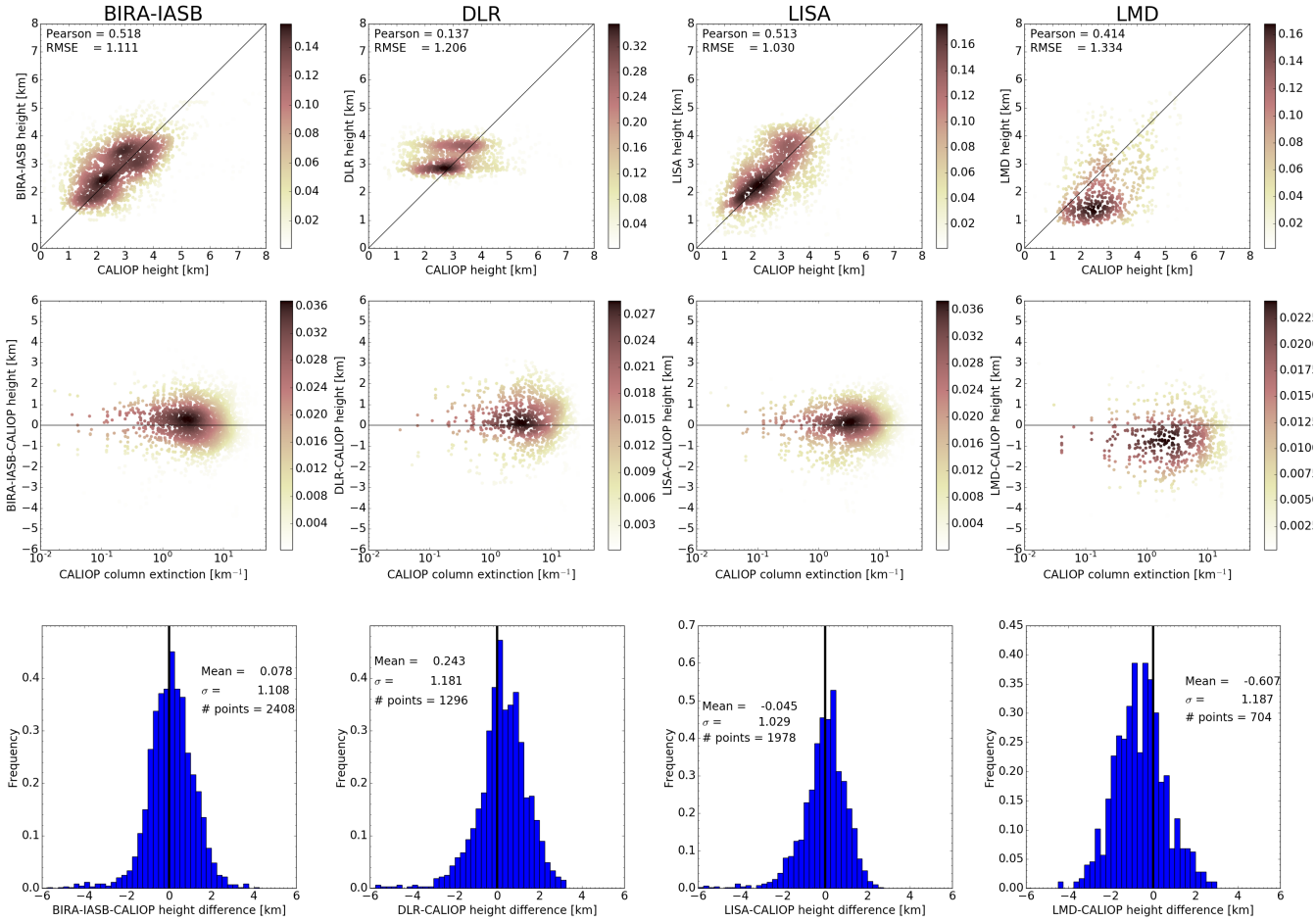

Figure 6. Similar to Fig. 5, but for the CALIOP geometric mean height~~This information is also provided in Table 3, numbers within parentheses~~.

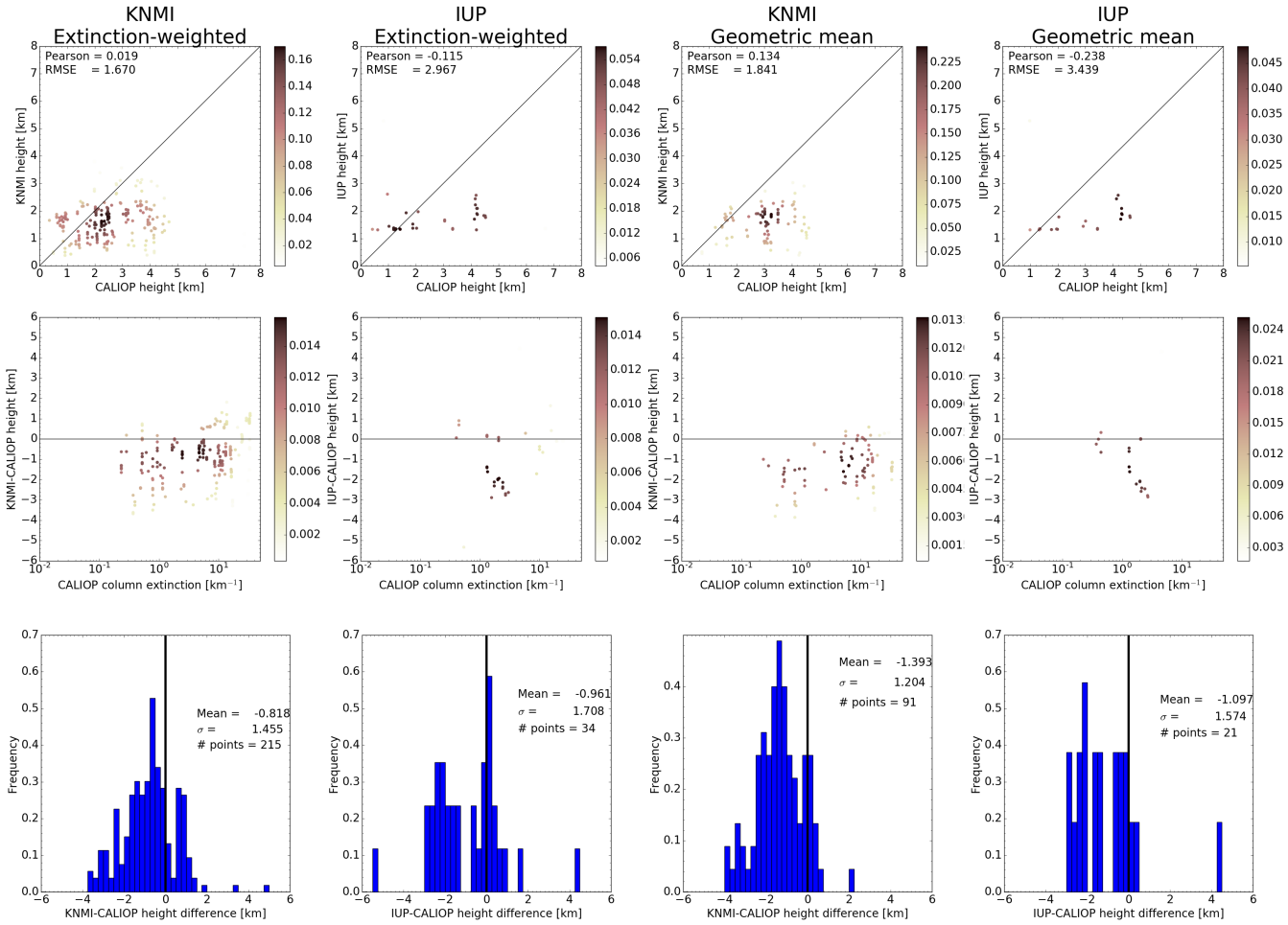

Figure 7. Similar to Figs. 5-6, but for the KNMI (first and third column) and IUP (second and fourth column) algorithms versus the CALIOP cumulative extinction (first and second column) and geometric mean (third and fourth column) dust heights.

Appendix A: Additional figures

In Figs. A1-A4 statistics are shown for all data, and land-day, ocean-day, land-night and land-day subsets for all IASI dust height retrieval methods as compared with the CALIOP geometric mean heights. Figs. A5-A8 show similar data but using the CALIOP cumulative extinction heights.

The plots in Figs. A1-A8 reflect the findings presented in Table 3. The BIRA-IASB algorithm agrees well with the CALIOP geometric mean height over land for day and night, Fig. A1. Over ocean the agreement is better during the night. It is noted that for the ocean day subset the histogram is bimodal. When compared with the cumulative extinction height, Fig. A5, the BIRA-IASB dust height is overestimated over ocean during both day and night; the ocean day subset appears to be bimodal; the agreement appears better over land during day than night, but this may in part be due to a bimodal histogram for the land day subset. This is reflected in the spread in the difference which is smaller during night than day. For DLR, Figs. A2 and A6, there are few data points available over the ocean. Over land the DLR height data are clumped at a single height for the day subset and at two heights for the night data subset. For LMD, Figs. A3 and A7, the agreement is mono-modal for the land day, land night and ocean night subsets when compared with both the CALIOP cumulative extinction and geometric mean heights. For the ocean day subset a bimodal distribution may be present. Overall the agreement is better when compared with the cumulative extinction height. The LISA data, Figs. A4 and A8, also have a bimodal ocean day distribution compared with the CALIOP heights. For the ocean the mean difference with the CALIOP cumulative extinction height is significantly larger during night than day. This difference is nearly a factor 2 smaller when comparing with the geometric mean height. For land the magnitude of the difference is smallest when compared with the cumulative extinction height during the day and with the geometric mean height during night.

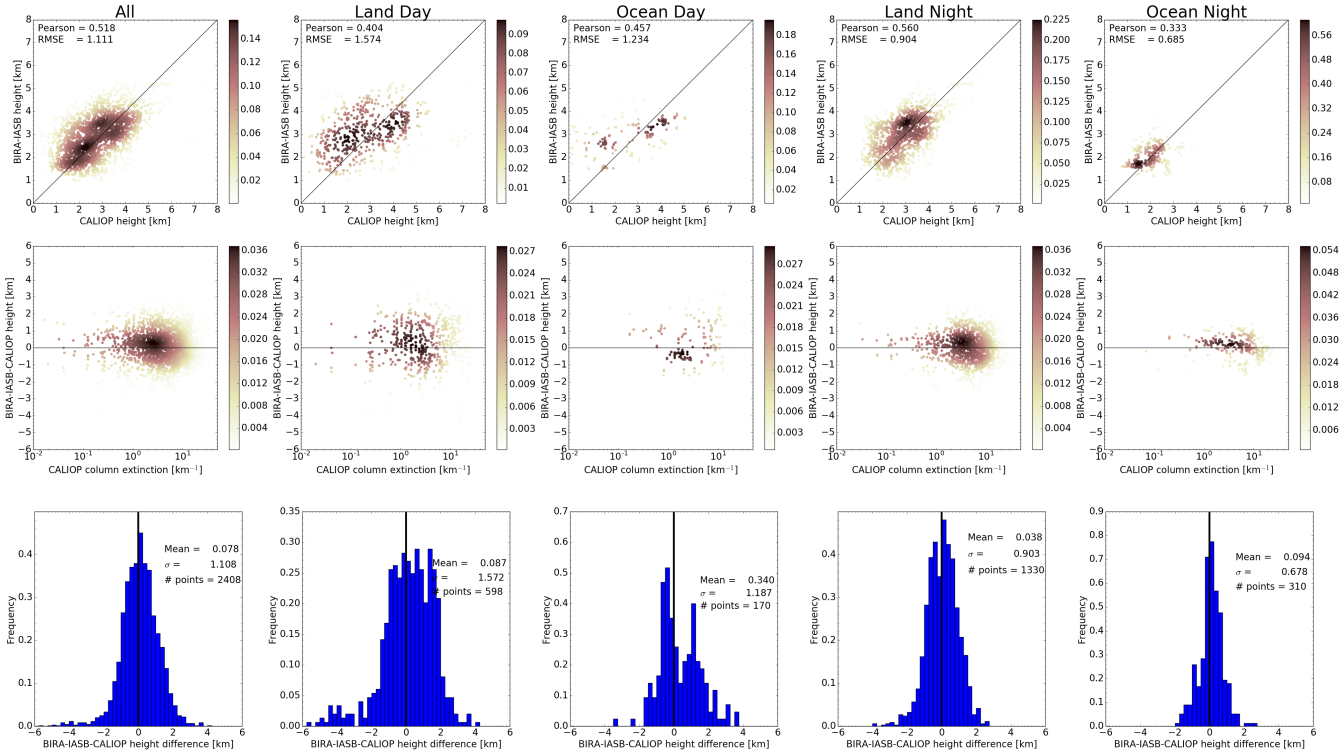

Figure A1. (Upper row) Scatter plots of the CALIOP geometric mean height versus height from the BIRA-IASB algorithm. (Centre row) Scatter plot of the difference between the BIRA-IASB and CALIOP heights versus the CALIOP column extinction. (Bottom row) Frequency distribution of the difference between the height from the BIRA-IASB algorithm and the CALIOP height. ~~Included is also a normal distribution fit to the difference.~~ The mean and standard deviation of the normal distribution are given in each plot. (Column 1) All data points. (Column 2) CALIOP day data and IASI land pixels. (Column 3) CALIOP day data and IASI ocean pixels. (Column 4) CALIOP night data and IASI land pixels. (Column 5) CALIOP night data and IASI ocean pixels.

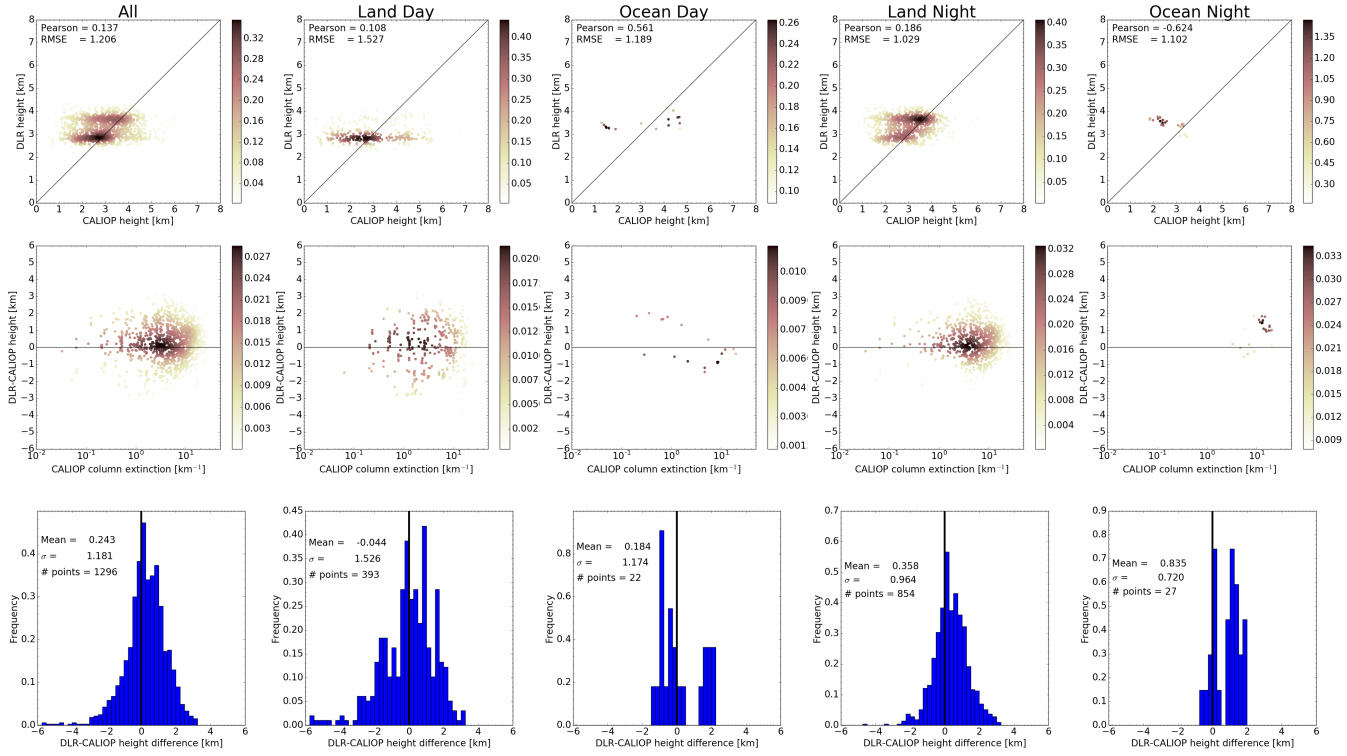

Figure A2. Similar to Fig. A5 but for the DLR algorithm.

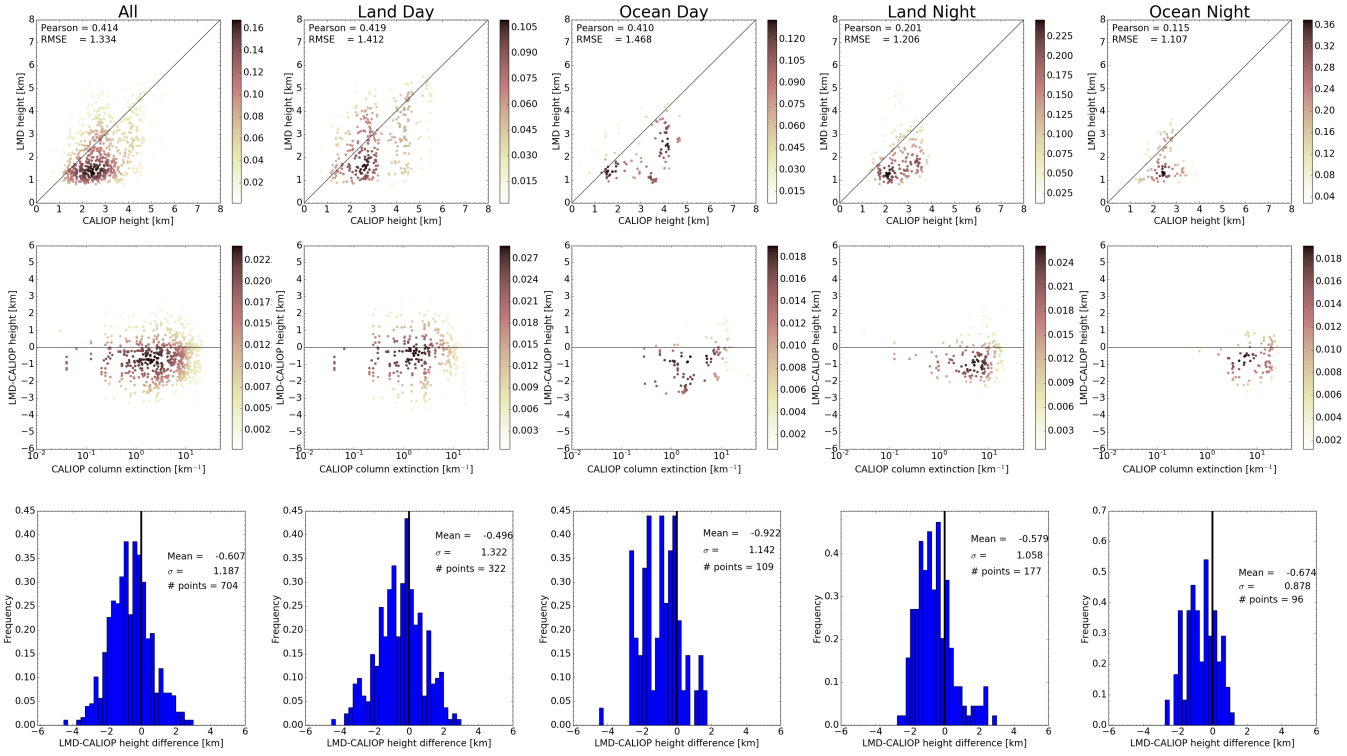

Figure A3. Similar to Fig. A5 but for the LMD algorithm.

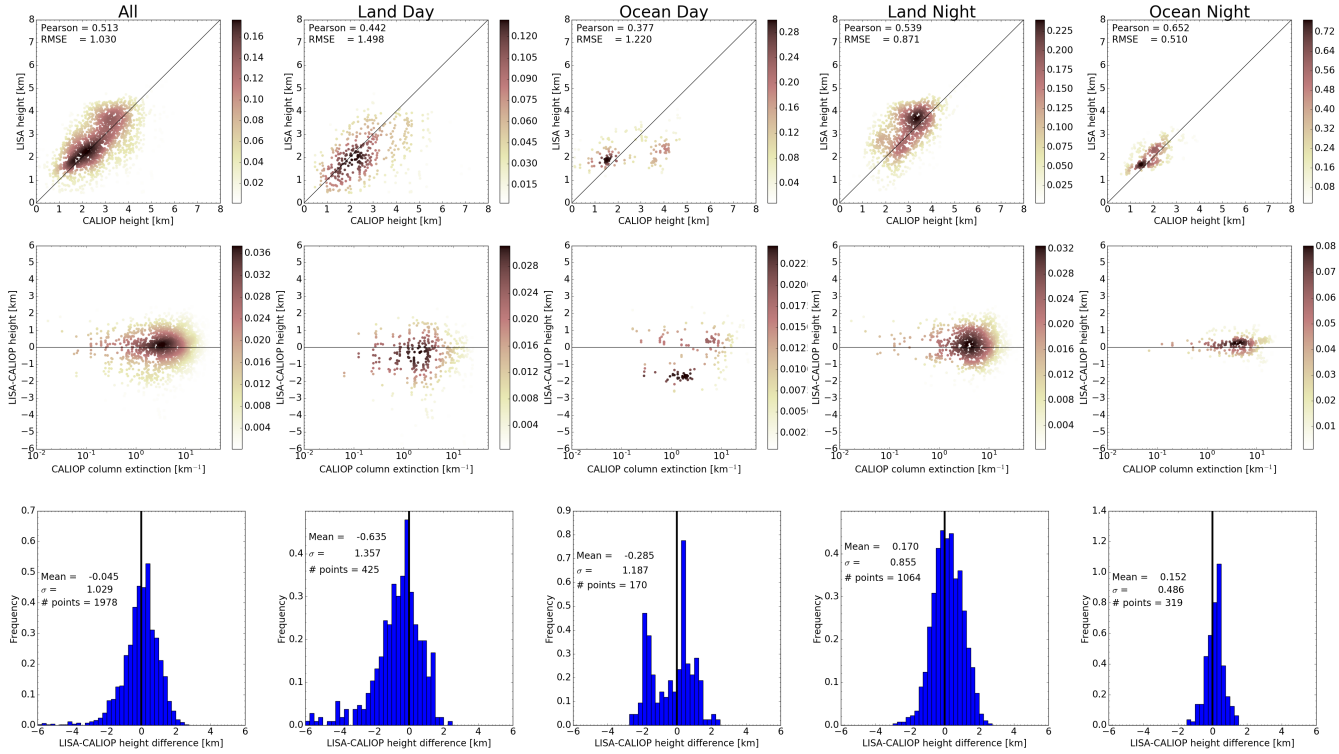

Figure A4. Similar to Fig. A5 but for the LISA algorithm.

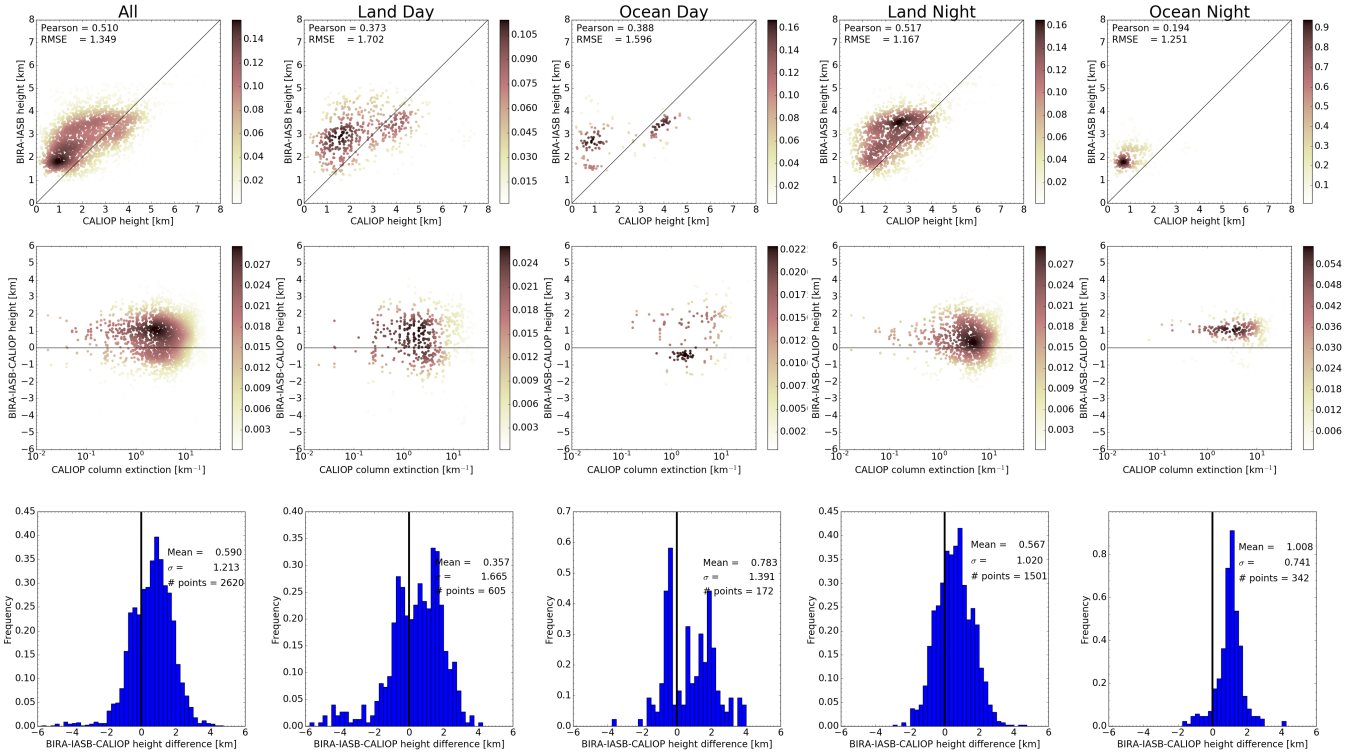

Figure A5. (Upper row) Scatter plots of the CALIOP cumulative extinction height versus height from the BIRA-IASB algorithm. (Centre row) Scatter plot of the difference between the BIRA-IASB and CALIOP heights versus the CALIOP column extinction. (Bottom row) Frequency distribution of the difference between the height from the BIRA-IASB algorithm and the CALIOP height. ~~Included is also a normal distribution fit to the difference.~~ The mean and standard deviation of the normal distribution are given in each plot. (Column 1) All data points. (Column 2) CALIOP day data and IASI land pixels. (Column 3) CALIOP day data and IASI ocean pixels. (Column 4) CALIOP night data and IASI land pixels. (Column 5) CALIOP night data and IASI ocean pixels.

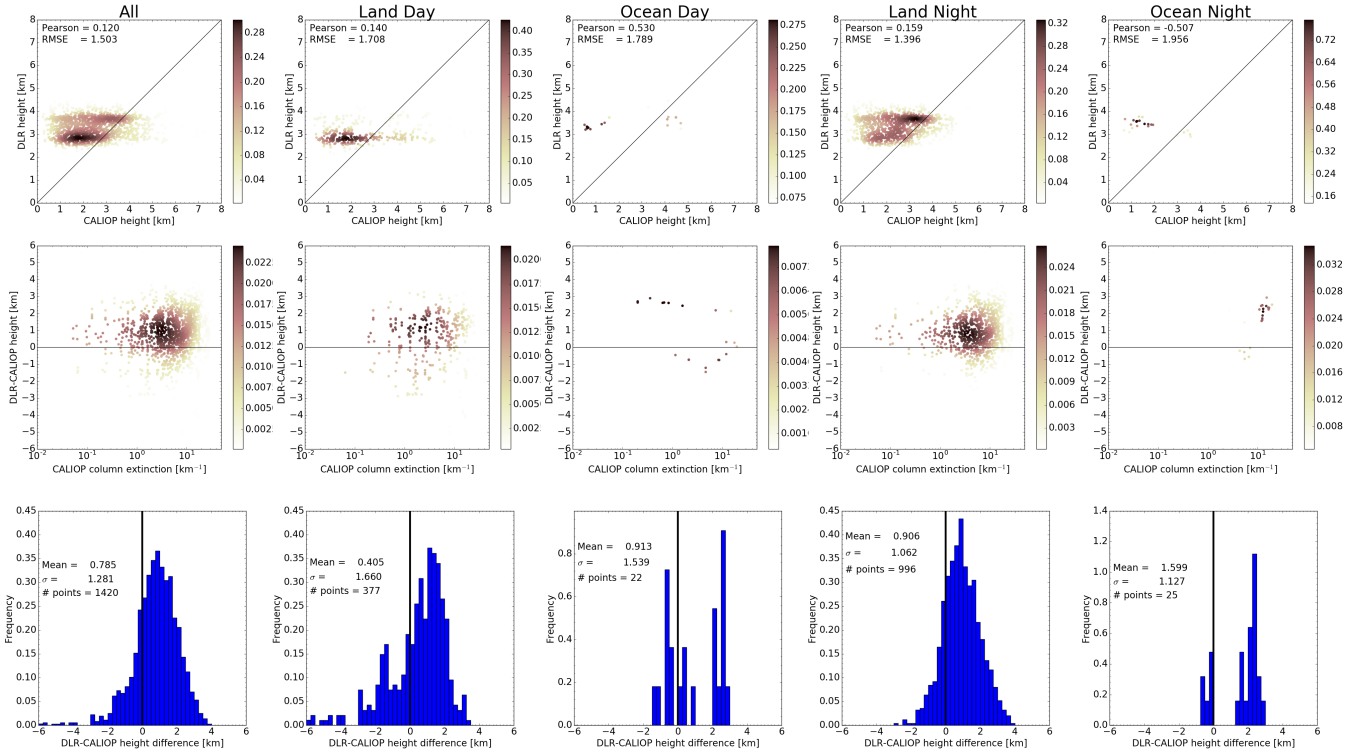

Figure A6. Similar to Fig. A5 but for the DLR algorithm.

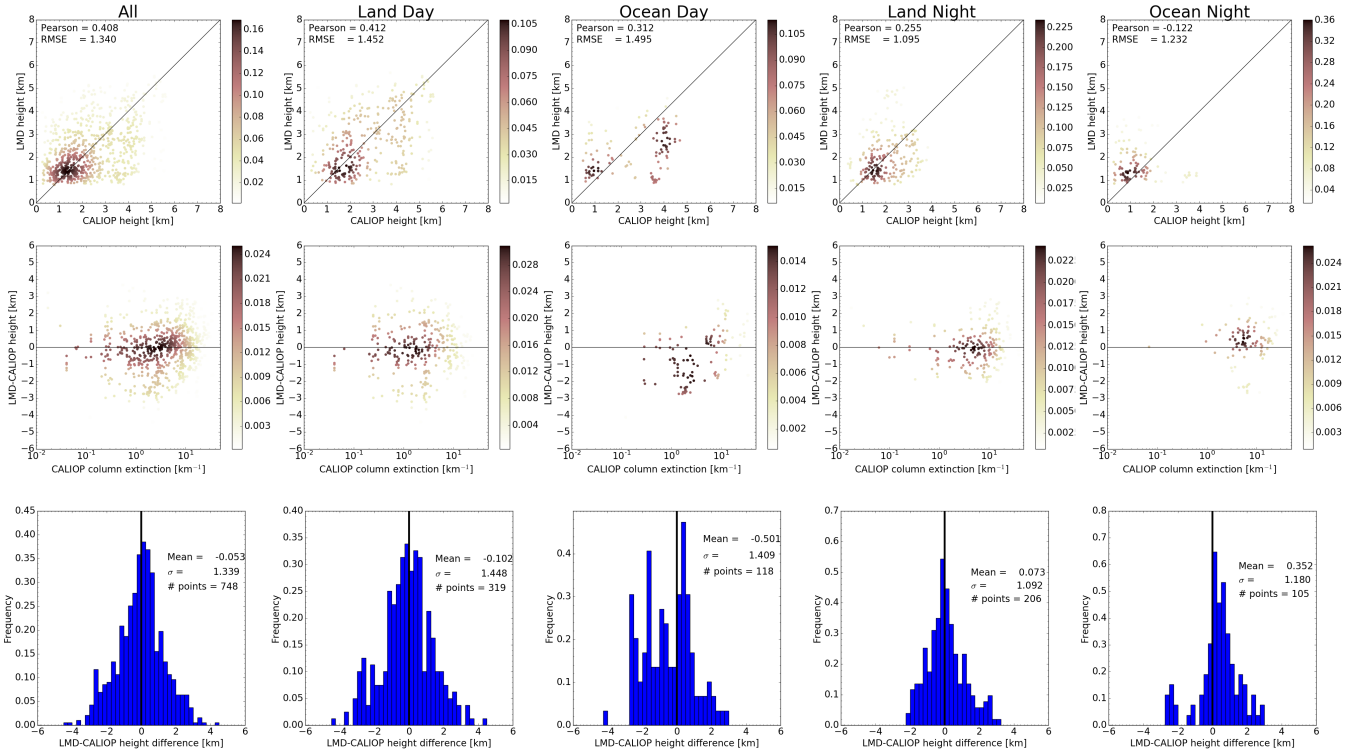

Figure A7. Similar to Fig. A5 but for the LMD algorithm.

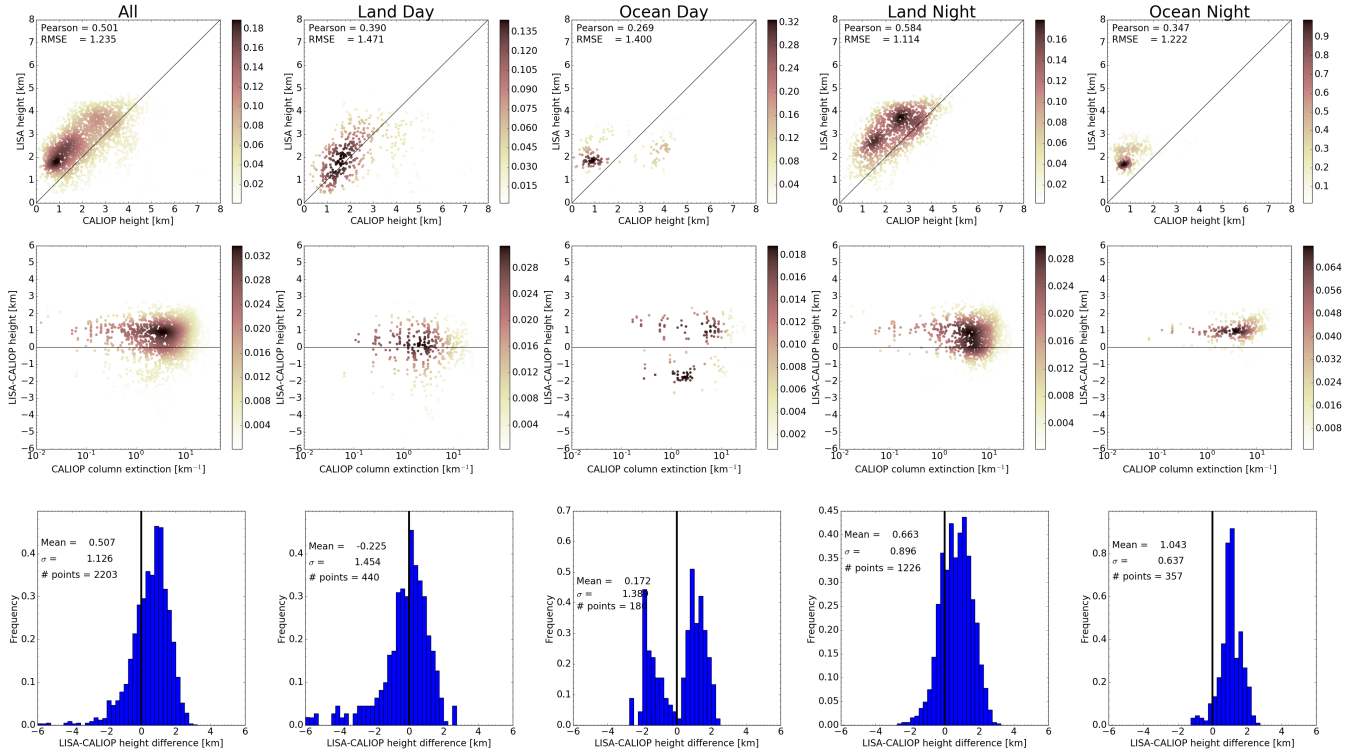

Figure A8. Similar to Fig. A5 but for the LISA algorithm.