Interactive comment on "EARLINET evaluation of the CATS L2 aerosol backscatter coefficient product" by Emmanouil Proestakis et al.
Anonymous Referee #2
Received and published: 17 March 2019

The authors would like to thank the reviewer for the interesting and at the same time substantial comments and suggestions. We tried, and did our best, to incorporate the most suitable proposed changes and corrections in the revised manuscript, aiming at improving the presented paper. Following, you will find our responses, one by one to the comments addressed, in the uploaded supplement pdf file.
Kind regards,
Emmanouil Proestakis

General comments:
This manuscript compares EARLINET (ground-based) and CATS (onboard the international spatial station) retrievals of the aerosol backscatter coefficient over 12 European sites and 1 Asian site. The paper is well written, however, I did miss some explanation in the introduction about the importance of CATS product. I believe this could be easily achieved by modifying the order of some paragraphs and including extra information. In particular, I suggest moving the second paragraph of Section 2.2 (page 4, line 30 to page 5, line 12) to the introduction, with the due adjustments.

The authors agree with the reviewer. The science goals of CATS, indeed, were not mentioned in the introduction, leading to issues in the understanding of the scientific importance of the project in the early stages of the manuscript. For this reason the authors have followed the referee's recommendation to rearrange the manuscript, making at the same time all the appropriate modifications to ensure that the adjustments did not have a negative impact to the understanding of the manuscript context. To be more specific, the following section was added to the introduction:

"CATS was developed to meet three main science goals. The primary objective was to measure and characterize aerosols and clouds on a global scale. The space-borne lidar orbited the Earth at an altitude of approximately 405 km and 51-degree inclination. The use of the ISS as an observation platform facilitated for the first time global lidar-based climatic studies of aerosols and clouds at various local times (Noel et al., 2018, Lee et al., 2018). In addition, near-real-time data acquisition of the CATS observations was developed towards the improvement of aerosol forecast models (Hughes et al, 2016). A secondary objective was related to the need of long-term and continuous satellite-based lidar observations to be available for climatic studies. The first spaceborne lidar mission, the Lidar In-space Technology Experiment (LITE; McCormick et al., 1993) in 1994, was succeeded by the joint NASA and Centre National d'Études Spatiales (CNES) Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) mission in June, 2006 (Winker et al., 2007). Since 2009 the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument (Winker et al., 2009) onboard CALIPSO operates on the secondary backup laser. The launch of the post-CALIPSO missions, the joint European Space Agency (ESA) and JAXA satellite Earth Cloud Aerosol and Radiation Explorer (EarthCARE; Illingworth et al., 2015) and the NASA's Aerosols, Clouds, and Ecosystems (ACE) are planned for 2021 and post-2020 respectively. The CATS project was partially intended to fill a potential gap on global lidar observations of vertical aerosols and clouds profiling. The third scientific objective of CATS was to serve as a low-cost technological demonstration for future satellite lidar missions (McGill et al., 2015). Its science goal to explore different technologies was fulfilled through the use of photon-counting detectors and of two low energy (1-2 mJ) and high repetition rate (4-5 kHz) Nd:YVO4 lasers (Multi-Beam and HSRL - UV demonstrations), aiming to provide simultaneous multiwavelength observations (355, 532 and 1064 nm). Additional gains of the CATS were related to the exploitation and risk reduction of newly applied laser technologies, to pave the

way for future spaceborne lidar missions (high repetition rate, injection seeding, wavelength tripling at 355 nm).”

I also suggest comparing some scenes of coincident vertical profiles of CATS and CALIOP. Would that be possible? I believe this would dramatically improve the visibility of the paper. Also, it wasn’t clear to me whether CATS should only be used to fill a gap in space-based lidar observations or if it is as reliable as CALIOP. I believe this should be further clarified in the text.

The suggested evaluation study between CATS and CALIOP has been already performed by Lee et al. (2018), Rajapakshe et al. (2017), Noel et al. (2018) and Yorks et al., 2016, reporting also on the good agreement of the intercomparison studies. However, the authors agree with the reviewer regarding the necessity of including the findings of the aforementioned studies and have adjusted the manuscript accordingly. To be more specific, the following paragraph was added to the manuscript (Section 1 - Introduction):

“*CATS performance has been validated against ground-based AErosol RObotic NETwork (AERONET; Holben et al., 1998) measurements and evaluated against satellite-based Atmospheric Optical Depth (AOD) retrievals of Aqua and Terra Moderate Imaging Spectroradiometer (MODIS; Levy et al., 2013) and active CPL (McGill et al., 2002) and CALIPSO CALIOP (Winker et al., 2009) profiles of extinction coefficient and AOD at 1064 nm. Lee et al. (2018) compared daytime quality-assured CATS V2-01 vertically integrated extinction coefficient profiles (1064 nm) and AERONET AOD (1020 nm) values, spatially (within 0.4^o Longitude and Latitude) and temporally (±30 minutes) collocated, and found a reasonable agreement with a correlation of 0.64. A comparative analysis of CATS and MODIS C6.1 Dark Target (DT) AOD retrievals, through spectral interpolation between 0.87 and 1.24 μm channels, reported correlation of 0.75 and slope of 0.79, over ocean. In addition, Lee et al., (2019) evaluated AOD and extinction coefficient profiles from CATS through intercomparison with CALIOP. With regard to AOD, analysis a total of 2681 CATS and CALIOP collocated observation cases (within 0.4^o Longitude/Latitude and ±30 minutes ISS and CALIPSO overpass difference), showed correlation of 0.62 and 0.52 over land and ocean respectively during daytime (1342 cases), and 0.84 and 0.81 over land and ocean respectively during nighttime (1339 cases). Comparison of CATS and CALIOP collocated extinction coefficient profiles based on the closest Euclidian distance on the earth’s surface, shows also good shape agreement, despite an apparent CALIOP underestimation in the lowest 2 km height. CATS and CALIOP observations were used by Rajapakshe et al. (2017) to study the seasonally transported aerosol layers over the SE Atlantic Ocean. The performed comparative analysis reported on similar geographical patterns regarding Above Cloud Aerosols (ACA), Cloud Fraction (CF) and ACA occurrence frequency (ACA_F) between CATS and CALIOP retrievals. However, the authors reported also on differences between CATS and CALIOP vertical aerosol distributions, with ACA bottom height identified by CATS lower than the respective of CALIOP. CATS retrievals were used to document the diurnal cycle and variations of clouds, with CALIOP complementarily used. Noel et al. (2018) showed that both CATS and CALIOP profiles of CF agree well on both the vertical patterns and values at 01:30 and 13:30 LT, over both land and ocean, with minor differences of the order of 2-7% throughout the entire profiles of cloud fraction. CATS depolarization measurements, which are critical in the processing algorithms of aerosol subtype classification, were investigated in the case of desert dust, smoke from biomass burning and cirrus clouds (Yorks et al., 2016), and were found consistent and in good agreement with depolarization measurements from previous studies and historical datasets implementing CPL (Yorks et al., 2011) and CALIOP (Liu et al., 2015).*

I also believe a final paragraph stating the main conclusion is needed (that is, what are your suggestions for future studies: should we use CATS or not, under which conditions these retrievals are reliable, what are their advantages and disadvantages and how could future studies benefit - or not - from CATS).

The authors agree with the reviewer and a final paragraph stating suggestions related the use of the unique CAST dataset was included. To be more specific, the following section was added to the "Summary and Conclusions section":

"The qualitative and quantitative agreement between CATS and EARLINET reported in this study is encouraging, especially during nighttime, agreement that will hopefully facilitate further studies implementing CATS observations in the future. CATS, for a period of almost three years, provided an unprecedented global dataset of vertical profiles of aerosols and clouds, much like CALIOP, taking though advantage of the unique orbital characteristics of the ISS. ISS enabled CATS to provide for the first time satellite-based lidar measurements of the diurnal evolution of aerosols and clouds over the tropics and midlatitudes, and to be more specific to latitudes below 52°. Since CALIPSO and Aeolus (and in the future also EarthCARE) are polar sun-synchronous satellites of fixed equatorial crossing time (01:30 and 13:30 LT for CALIOP, 06:00 and 18:00 for ALADIN), it is expected that, at least for the near future, CATS dataset will remain the only available satellite-based lidar source of nearly global diurnal measurements of atmospheric aerosols and clouds. In addition, while CALIOP is a two-wavelength lidar system operating at 532 nm and 1064 nm with depolarization capabilities at 532 nm, CATS provided satellite-based aerosol and cloud depolarization profiles at 1064 nm, thus in a different wavelength. This dataset, much like CALIOP dataset, is especially useful for studies of the three-dimensional distribution of non-spherical aerosol particles in the atmosphere (e.g. mineral dust and volcanic ash), and especially since it is an active sensor, over regions of high reflectivity (e.g. deserts, ice). Future studies including the exploitations of CATS unique observations may help the scientific community to shed new light on physical processes of aerosols and clouds in the Earth's atmosphere."

Specific comments:
page 2, line 3 - Please modify "Physic" to "Physics".

The text is modified according to the reviewer's recommendation.

page 2, line 20 - Please reformulate the sentence (suggestion: "Quality assessment of CATS...").

The text is modified according to the reviewer's recommendation.

page 2, line 24 - Please modify "consists" to "consists of".

The text is modified according to the reviewer's recommendation.

page 3, line 15 - What is the difference between capacity and capability?

The text is reformulated according to the reviewer's recommendation:

"Since the beginning of the initiative in 2000, EARLINET has significantly increased its observing and operational capacity"

page 3, line 16 - Please reformulate or remove the sentence "EARLINET stations are classified as active on condition of...".

According to the reviewer's comment, the sentence was reformulated to:

"EARLINET stations are classified between "active", "not permanent", "joining" and "not active". An EARLINET station is classified as active when on condition of performing regularly and simultaneously measurements with the other stations composing the lidar network, and

accordingly, contributing with uploading the performed measurements to the EARLINET database (https://www.earlinet.org/, last access: 20 December 2018)."

page 4, line 32 - Please modify "space-borne" to "spaceborne".

The text is modified according to the reviewer's recommendation.

page 6, line 16 - It's not clear to me if observations more than 90 minutes apart were compared or not. Could you clarify this?

The study follows the CALIPSO CALIOP validation methodology developed in the framework of a collaboration between ESA and EARLINET collaboration (Pappalardo et al., 2010). The ESA dedicated program of collocated and concurrent EARLINET observations with CALIOP observations was developed prior to the launch of CALIPSO and is planned with a duration until the end-of-mission of the mission. On the contrary of the well-established CALIPSO-EARLINET validation activity, but also to the ESA-Aeolus and to the upcoming ESA-EarthCARE satellite missions, a similar CATS-EARLINET validation strategy was not established.
The participating EARLIENT stations in the study contributed to the evaluation of CATS through measurements performed during the fixed-scheduled program of EARLINET operation. As described in Pappalardo et al (2014), the EARLINET scheduled program of measurements includes three measurements per week, one during daytime around local noon (Monday, 14:00 ± 1h) and two during nighttime (Monday/Thursday, sunset + 2/3h), to enable Raman extinction retrievals. In addition, EARLIENT operates a small number of lidar systems capable for 24/7 continuous measurements (Engelmann et al., 2016).
The absence of an established dedicated validation activity between NASA and EARLINET prior to the operation of CATS, in combination with the fixed measurements schedule of EARLINET, the high variable overpass-time of CATS (bounded by the orbital characteristics of ISS) and the frequently cloud-contaminated cases led to a low number of collocated and concurrent EARLINET-CATS cases to be available for the study. Eventually, this obstacle was tackled through the cooperative effort of a large number of EARLINET stations, contributing through the already performed measurements. The increasing number of EARLINET stations showing interest to contribute to the study led to an overall of forty-seven (47) available cloud-free EARLINET-CATS collocated cases to implement for the evaluation of CATS.
The EARLINET-CATS correlative study considers the collocation criteria established in the validation plan of CALIPSO. Regarding the spatial collocation, EARLINET participating stations contributed with measurements when the ISS overpass was within 50 km horizontal radius from their location.

Regarding the temporal collocation, the study implemented ground-based measurements with a temporal window of EARLINET performed measurements with starting time, or stop time as close in time as possible to the ISS overpass. Accordingly, all the identified EARLINET cases where studied, through case-by-case inspection of the Range-Corrected-Signal quicklooks, for atmospheric homogeneity was of high importance, and additionally for other constrains (e.g. cirrus-clouds). During the first twenty months of CATS operation, based on thirteen EARLINET contributing stations, only 47 cases were found suitable to be used in the comparison. From the total of 47 cases, 44 where performed with "*starting time*", or "*stop time*" within 90 minutes of the ISS overpass. For this reason why the phrase *"typically within 90 minutes of the ISS overpass"* was used in the manuscript. In addition, it has to be mentioned that in the majority of the EARLINET cases encompasses the ISS overpass. The length of the temporal window was variable, based-on the expertise of the EARLINET teams, the homogeneity of the atmospheric scenes and the unique cloud constrains of each case, in order to allow retrievals of high-quality EARLINET backscatter coefficient profiles.

The authors agree though with the reviewer that this part of the manuscript was not clear, therefore the manuscript was revised in the 2.3.1 section referring to the "Comparison methodology", and in addition the manuscript was updated with the following table (*"Table 2"* in the manuscript) that includes information on the correlative cases used in the study. The table provides the "Day-Night Flag" of the study case, "Date" and "Time" of the ISS overpass, the corresponding EARLINET station and the minimum distance between the ISS orbit-track and the station location, and finally the EARLINET temporal window of measurements.

In Section 2.3.1, the following part of the manuscript was reformulated according to the reviewer's recommendation, from:
"In addition, the correlative measurements should be as close in time as possible. EARLINET contributed with performed measurements as close in time as possible, typically within 90 min of the ISS station overpass."
to:
"EARLINET contributed with performed measurements as close in time as possible, typically with starting time or stop time of the performed measurements widow within 90 min of the ISS station overpass. The EARLINET-CATS cases considered to the assessment of the accuracy and representativeness of CATS backscatter coefficient profiles are provided in Table 2, including the name of the EARLINET station, the EARLINET measurements window, the ISS overpass time, the ISS minimum distance between the corresponding EARLINET station and the lidar footprint of CATS and the Daytime/Nighttime information."

Table 2: ISS-CATS and EARLINET cases considered in the evaluation process of CATS backscatter coefficient profiles at 1064 nm.

Day-Night Flag	Date yyyy/mm/dd	Time hh:mm:ss (UTC)	EARLINET station	min Distance (km)	EARLINET Date (yyyy/mm/dd) | measuring time cloud-free window (UTC)
N	2015/11/25	03:44:09	Athens	40.42	2015/11/25 | 03:30:00 – 04:30:00
N	2016/01/29	01:46:08	Athens	46.84	2016/01/29 | 01:00:00 – 02:30:00
N	2016/02/01	17:23:36	Athens	23.29	2016/02/01 | 17:45:00 – 19:30:00
N	2016/02/01	17:23:39	Athens_NTUA	18.58	2016/02/01 | 18:20:51 – 19:57:41
D	2016/05/03	06:45:15	Barcelona	45.93	2016/05/03 | 08:59:00 – 09:59:00
D	2015/08/13	17:29:18	Belsk	2.39	2015/08/13 | 18:02:10 – 18:45:40
N	2016/08/08	17:34:50	Belsk	6.56	2016/08/08 | 17:31:08 – 18:12:05
N	2016/07/28	19:15:24	Bucharest	45.35	2016/07/28 | 17:41:22 – 18:41:22
N	2016/09/14	04:21:09	Cabauw	21.01	2016/09/14 | 05:27:25 – 06:00:03
N	2015/08/03	21:40:39	Dushanbe	42.64	2015/08/03 | 20:00:00 – 22:00:00
N	2016/08/14	15:39:07	Dushanbe	22.08	2016/08/14 | 15:57:00 – 17:19:00
D	2015/06/20	08:38:33	Dushanbe	13.33	2015/06/20 | 08:54:00 – 09:07:00
D	2015/07/12	06:47:07	Dushanbe	33.46	2015/07/12 | 06:25:00 – 07:10:00
D	2016/05/02	07:35:38	Evora	47.27	2016/05/02 | 07:58:50 – 08:00:21
D	2016/05/31	19:43:41	Evora	39.42	2016/05/31 | 19:29:56 – 19:59:35
N	2016/01/30	00:50:16	Hohenpeissenberg	13.36	2016/01/30 | 00:20:00 – 01:20:00
N	2016/03/17	02:12:09	Hohenpeissenberg	43.40	2016/03/17 | 01:42:00 – 02:42:00
D	2015/10/31	12:56:05	Hohenpeissenberg	34.41	2015/10/31 | 12:26:00 – 13:26:00
D	2016/04/12	15:29:18	Hohenpeissenberg	12.77	2016/04/12 | 14:55:00 – 16:05:00
D	2016/08/07	16:49:29	Hohenpeissenberg	31.81	2016/08/07 | 16:19:30 – 17:19:30
D	2016/08/23	10:42:43	Hohenpeissenberg	36.11	2016/08/23 | 10:12:30 – 11:12:30
D	2016/09/14	05:58:59	Hohenpeissenberg	28.37	2016/09/14 | 04:59:00 – 05:59:00
N	2015/07/27	21:14:35	Lecce	34.69	2015/07/27 | 20:42:00 – 21:09:00
N	2016/08/04	22:44:06	Lecce	4.72	2016/08/04 | 20:50:00 – 21:20:00
N	2015/07/30	00:18:19	Leipzig	41.16	2015/07/30 | 00:34:00 – 01:04:00
N	2015/08/03	21:29:44	Leipzig	15.81	2015/08/03 | 21:31:00 – 22:00:00
N	2015/09/24	01:13:34	Leipzig	25.05	2015/09/24 | 01:01:00 – 01:30:00
N	2015/09/29	00:05:33	Leipzig	36.49	2015/09/28 | 22:42:00 – 23:12:00

N	2015/09/29	23:13:24	Leipzig	48.46	2015/09/28 | 22:55:00 – 23:24:00
N	2015/09/30	22:21:13	Leipzig	12.89	2015/09/30 | 21:25:00 – 21:34:00
N	2016/06/05	20:14:01	Leipzig	36.93	2016/06/05 | 20:02:00 – 20:31:00
N	2016/09/13	03:37:49	Leipzig	3.79	2016/06/05 | 00:00:00 – 02:30:00
N	2016/09/12	04:29:46	Leipzig	45.08	2016/09/12 | 00:00:00 – 02:30:00
N	2016/09/15	03:30:25	Leipzig	48.36	2016/09/15 | 00:00:00 – 02:30:00
D	2015/04/21	14:54:35	Leipzig	6.73	2015/04/21 | 16:04:00 – 16:33:00
D	2015/04/21	16:31:00	Leipzig	31.28	2015/04/21 | 16:34:00 – 17:04:00
D	2015/04/24	15:25:13	Leipzig	47.83	2015/04/24 | 14:03:00 – 14:32:00
D	2015/08/13	17:27:54	Leipzig	1.36	2015/08/13 | 19:01:00 – 19:30:00
D	2016/08/24	11:26:39	Leipzig	3.46	2016/08/24 | 10:00:00 – 12:00:00
D	2016/08/24	13:03:12	Leipzig	48.97	2016/08/24 | 10:00:00 – 12:00:00
N	2015/07/21	00:13:26	Potenza	2.01	2015/07/21 | 00:00:00 – 02:52:19
D	2015/11/06	10:54:52	Thessaloniki	19.46	2015/11/06 | 11:57:03 – 12:27:20
N	2016/01/28	19:17:11	Thessaloniki	39.54	2016/01/28 | 20:08:40 – 20:38:57
D	2015/08/13	17:29:20	Warsaw	42.95	2015/08/13 | 17:00:00 – 17:22:00
D	2015/08/19	15:22:30	Warsaw	44.47	2015/08/19 | 15:25:00 – 15:47:00
D	2016/06/07	18:29:46	Warsaw	41.22	2016/06/07 | 18:15:00 – 18:43:00
N	2016/08/08	17:34:53	Warsaw	46.99	2016/08/08 | 17:00:00 – 17:23:00

page 6, line 24 - What does "including cirrus clouds" mean? Cirrus clouds scenes were used or not?

The authors acknowledge that the sentence was not clearly written, thus the sentence was reformulated from:
"In addition, to account for contamination effects of multiple-scattering and specular reflection in the intercomparison process, only cloud-free (including cirrus clouds) atmospheric scenes are used."
to:
"In addition, to account for contamination effects of multiple-scattering and specular reflection in the intercomparison process, only cloud-free atmospheric scenes are used. Cases with detected cirrus either at the EARLINET Range-Corrected-Signal quicklooks or at the ISS-CATS backscatter coefficient profiles or the feature type profiles are not considered in the study."

page 7, line 19 - Please modify "participated" to "participating".

The text is modified according to the reviewer's recommendation.

page 7, line 20 - "exited". Did you mean "excited"?

The reviewer is correct, the text is modified according to the reviewer's comment.

page 9, line 14 – Please modify "in details" to "in detail".

The text is modified according to the reviewer's recommendation.

page 9, line 24 - Please modify "below" to "of".

The text is modified according to the reviewer's recommendation.

Page 9, line 37 - Please modify "over-lying" to "overlaying".

The text is modified according to the reviewer's recommendation.

page 11, line 22 - Has the new product already been released? How does the new algorithm differ from the previous one? What kind of improvements does it present?

CATS V3-00 replaced CATS V2-05 on October 1st, 2018. Initially, the changes in CATS Level 1 and Level 2 algorithms corresponding to CATS Version 3-00 data was planned to be the final algorithm release for the CATS project, though observed issues in the CATS products led to the modifications of V3-00 and the release of the V3-01 later on in the beginning of 2019. Since CATS products are provided in different levels of processing, the made changes in the algorithms correspond to both L1B and L2O products.
To be more specific, the changes in the L1B algorithms include:
(1) improvement of the nighttime attenuated total backscatter (ATB) profiles due to improvements in the calibration of CATS, thus improvement also in the daytime ATB profiles, since nighttime ATB is implemented in the calculations of the daytime calibration.
(2) changes to the “Depolarization_Quality_Flag”, and
(3) implementation of MERRA-2 Reanalysis data instead of GMAO forecasts, for the meteorology in V3-00 and V3-01.
The changes made in the algorithms of CATS L1B reflect on improvements on CATS L2O products. Though additional changes in CATS L2O algorithms include also:
(1) updates in number of profiles in the L2O datasets
(2) improvements in the calculations of uncertainties in the L2O layer-integrated parameters
(3) changes to the “Depolarization_Quality_Flag”
(4) improvements of the Cloud Aerosol Discrimination (CAD) through the implementation of an additional parameter, namely the “Cloud_350m_Fraction_XXX_FOV”, to report of the number of 350 L1B profiles within each 5 km L2O bin of the L2O layer product with attenuated total backscatter values greater than 0.03 $km^{-1}sr^{-1}$, thus atmospheric features of high probability of being a cloud. In addition, the parameter “Num_Profs_Avg_LRatio_XXX_FOV” was added to the L2O Layer data product.
(5) improvements in CATS Feature Type and Feature Type Score variables, but also in the Aerosol Subtype classification (replace of “volcanic” with “UTLS Aerosol”) and addition of the parameters “Opaque_Feature_Optical_Depth_1064_XXX_FOV” and “Opaque_Feature_Optical_Depth_Uncertainty_1064_XXX_FOV” in Mode 7.2 L2O datasets.
(6) Updates in the Lidar Ratio (LR) values for cirrus clouds
(7) update of the effective multiple scattering factor for ice clouds values to 0.52.
The above changes in the CATS V3-00 and V3-01 algorithms and the respective products are extensively presented and in-depth discussed in the CATS official website (https://cats.gsfc.nasa.gov/; last visit on: 22/05/2019), in the “Publications” section.

pages 11 and 12, Section 3.2 and Table 2: It would be interesting to show the mean relative bias (that is bias over mean value).

According to the referee’s comment we have computed and included in the table of comparison statistics between CATS and EARLINET the Mean Relative Bias (MRB), calculated as follows:

$$MRB = \left(\frac{1}{n}\sum_{1}^{n}\frac{(b_{CATS} - b_{EAR})}{b_{EAR}}\right) * 100$$

The MRB were found equal to -24.06% and -19.84% for daytime and nighttime CATS observations respectively, and the results were included to the table.

page 14, line 10, Please modify "discrepancies" to "discrepancy".

The text is modified according to the reviewer’s recommendation.

page 15, line 1, Please modify "based to" to "based on".

The text is modified according to the reviewer’s recommendation.

Figs. 3 to 5: Please use "b) CATS backscatter coefficient at 1064 nm", or "(1064 nm)".

The text is modified according to the reviewer's recommendation.

Fig. 5: I would guess topography influence CATS coefficient quite significantly. Could it be causing the spykes shown in this figure? Could you provide a quantitative estimate of the contributing effect of topography on the discrepancy observed in this figure? What about an estimate of the other contributing effects?

Regarding the question of the reviewer on the discussed constrains on the dataset, Figures 1-4 show quantitatively the effects of (i) distance between the EARLINET station and the closest profile of the CATS-ISS overpass for each correlative case, (ii) CATS Feature Type, (iii) number of CATS Level 2 (L2) Aerosol Profiles (APro) used in the CATS horizontal average, and the effect of (iv) topography of EARLINET stations. The comparison exercise examines the effect of one discussed constrain at a time, while keeping all the other parameters in the methodology constant, and considers various evaluation metrics, as discussed in the following sections.

(i) Effect of distance between the EARLINET station and the closest profile of the CATS-ISS overpass

Figure 1 shows the effect of distance between the closest CATS L2 APro and the respective EARLINET station matchup, for different upper Euclidean distance thresholds (i.e.: 5n km, $n \in N=\{1,10\}$). To be more specific, the Mean Bias (MB; [$Mm^{-1}sr^{-1}$]) - (Fig.1a), Root Mean Square Error (RMSE; [$Mm^{-1}sr^{-1}$]) - (Fig.1b), Correlation Coefficient (Fig.1c), and the number of CATS-EARLINET correlative cases per each upper distance threshold are considered. For each upper distance threshold, all the available CATS-EARLINET cases of Euclidean distance lower or equal to the respective upper limit are considered in the computation of the aforementioned evaluation metrics. This cumulative approach is selected due to the limited number of CATS-EARLINET correlative cases, and is applied separately for daytime and nighttime ISS overpasses, due to the different CATS measurement conditions.
Based on the analysis, during nighttime (daytime), the CATS-EARLINET MB is increasing (decreasing) starting from the 5 km upper distance threshold, to reach -0.0300 (-0.123) $Mm^{-1}sr^{-1}$, for the radius threshold of 50km shown in the study. The computed RMSE values are in the range between 0.447 and 0.343 $Mm^{-1}sr^{-1}$ for nighttime and between 0.357 and 0.448 $Mm^{-1}sr^{-1}$ for daytime, for the distance thresholds of 5km and 50km respectively. The minimum RMSE values are observed when considering ISS overpass cases of closer than 40 km distance to the EARLINET stations during nighttime, corresponding to MB of 0.018 $Mm^{-1}sr^{-1}$. The Correlation Coefficient is decreasing with increasing distance between the ISS overpass and the EARLINET stations. Notably, the Correlation Coefficient is not changing considerably for thresholds between 15 and 40 km for nighttime (~0.8) and between 15 and 30 km for daytime (~ 0.7). Sharp decreases in the Correlation Coefficient are observed during daytime (0.547), for distances closer to the EARLINET stations than during nighttime (0.693), for 35 and 40 km distance respectively.
The observed tendencies can be explained in terms of the distance thresholds and number of available cases, since the distance thresholds define the number of cases that are used in the analysis and the number of case is critical to assess the performance of CATS. Consequently, the MB, RMSE and Correlation Coefficient are all subject to both the number and the characteristics of the CATS-EARLINET cases used. In the study the authors use the maximum number of available EARLINET cases, to avoid any possible selection effect resulting from a poor sample of correlative cases, when strict collocation filters are applied. Using the

maximum number of available correlative cases, i.e. twenty six (26) and twenty one (21) for nighttime and daytime respectively, for ISS overpasses within 50km radius from the EARLINET stations, the authors envisage to quantitatively address the question of CATS performance and the representativeness of the aerosol backscatter coefficient profiles, over various atmospheric, illumination and ISS overpass conditions.

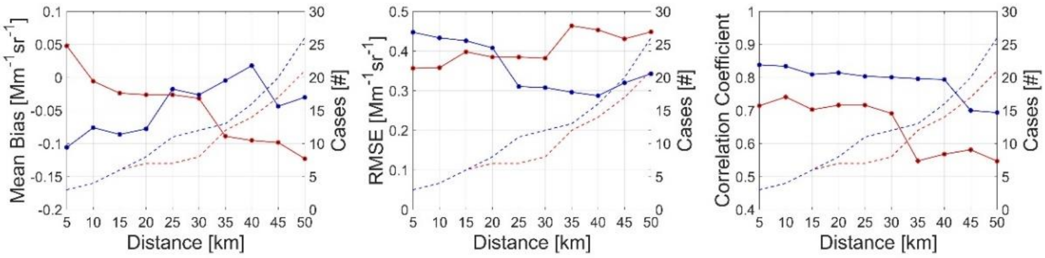

Figure 1: CATS backscatter coefficient at 1064nm with respect to EARLINET ground-based measurements, as a function of distance (km) between the closest CATS Level 2 Aerosol Profile and the respective "collocated" EARLINET station, for daytime (red line) and nighttime (blue line) ISS overpasses. Left: Mean Bias [$Mm^{-1}sr^{-1}$], center: RMSE [$Mm^{-1}sr^{-1}$] and right: Correlation Coefficient. Dashed lines correspond to the number of CATS-EARLINET correlative cases considered per each upper distance threshold between the CATS footprint and the locations of EARLINET stations.

(ii) Effect of Feature Type Score

The main objective of the CATS Cloud Aerosol Discrimination (CAD) score, or Feature Type Score, is to provide to the Feature Type classification a level of confidence. In the case of CATS, the Feature Type score is an integer number ranging between -10 and 10. The values of CATS Feature Type score correspond to classified aerosol atmospheric layers (negative values) and cloud atmospheric layers (positive values), while the magnitude of the Feature Type score corresponds to the confidence level of the classification. A value of -10 indicates complete confidence that the layer is an aerosol layer, while Feature Type score equal to 0, indicates an atmospherics layer with equal probability to be cloud or aerosol.

Figure 2 shows the effect of Feature Type Score, for different values, between -8 and 0 (i.e. for atmospheric layers classified as aerosol layers). The Mean Bias (MB; [$Mm^{-1}sr^{-1}$]) - (Fig.2a), Root Mean Square Error (RMSE; [$Mm^{-1}sr^{-1}$]) - (Fig.2b) and Correlation Coefficient (Fig.2c) are shown per each Feature Type Score. For each Feature Type score, cases of lower classification confidence level are not considered in the assessment of CATS performance and representativity, indicating the effect of the selected Feature Type thresholds.

Based on the MB, RMSE and Correlation Coefficient, a similar tendency is observed for different Feature Type Scores. To be more specific, not considerable changes are observed for different Feature Type Scores, regardless of the selected Feature Type threshold. This effect is due to the atmospheric characteristics of the CATS-EARLINET cases considered in the analysis. In the framework of the study, to account for contamination effects of multiple-scattering and specular reflection in the intercomparison process, only cloud-free atmospheric scenes are used. Furthermore, cases with detected cirrus, either at the EARLINET Range-Corrected-Signal quicklooks or at the ISS-CATS backscatter coefficient profiles or the feature type profiles, are not considered in the study. Initially, the presence of clouds was investigated through the implementation of CATS backscatter coefficient and depolarization time-height images and EARLINET range-corrected-signal. Cases for which the retrieval of EARLINET temporally-averaged profile was not feasible due to the presence of clouds, and/or CATS cases that the presence of clouds propagated into the CATS spatial-averaged profile were discarded from the analysis. Consequently, the lack of dependence shown in Figure 2 (a-c) is the result from the a priory selection of cloud free conditions selected in the analysis. However, a notably characteristic is the nighttime performance of CATS, which as shown from the lower absolute MB and lower RMSE, but in addition from the higher Correlation Coefficient values, due to higher SNR, is more representative than the corresponding daytime performance.

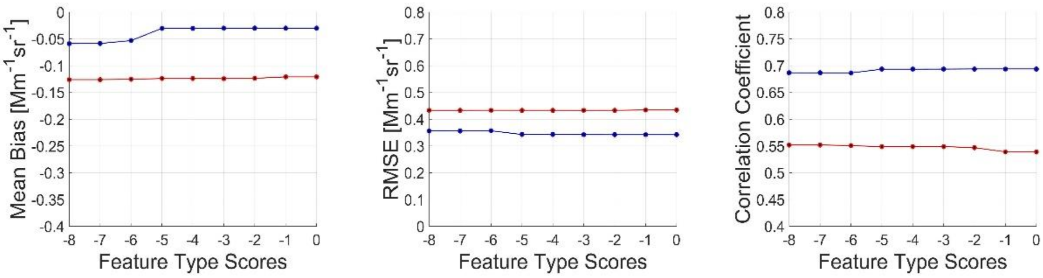

Figure 2: CATS backscatter coefficient at 1064nm with respect to EARLINET ground-based measurements, as a function of Feature Type score, for daytime (red line) and nighttime (blue line) ISS overpasses. Left: Mean Bias [Mm^{-1}sr^{-1}], center: RMSE [Mm^{-1}sr^{-1}] and right: Correlation Coefficient.

(iii) Effect of number of CATS-ISS L2 aerosol profiles used in the spatial averaging

Similarly to the analysis presented and discussed above, Figure 3 shows the effect of different number of aerosol profiles used when spatially averaging to retrieve the CATS aerosol profiles used in the framework of the study. In Figure 3, the acronym “CPro” corresponds to the closest CATS profiles to the corresponding EARLINET station. Accordingly, the Mean Bias (MB; [Mm^{-1}sr^{-1}]) - (Fig.3a), Root Mean Square Error (RMSE; [Mm^{-1}sr^{-1}]) - (Fig.3b), Correlation Coefficient (Fig.3c), are computed for different number of profiles used (i.e. CPro±1Profile, CPro±2Profiles, …).

Based on the MB, RMSE and Correlation Coefficient, the representativeness of CATS spatial profile is increasing with increasing number of aerosol profiles used in the horizontal averaging. To be more specific nighttime MB is almost constant, showing a low dependence on the number of profiles used, while for daytime CATS cases the opposite effect is observed, with improvement of CATS performance though increasing number of profiles used. Regarding RMSE no significant changes are observed, though a slight decreasing tendency in the RMSE is observed for both daytime and nighttime cases. Regarding the Correlation Coefficient, increasing in the values is also observed, with increasing number of profiles used, both for daytime and nighttime cases, denoting the improvement of the representativeness with increasing number of CATS profiles used in the spatial averaging.

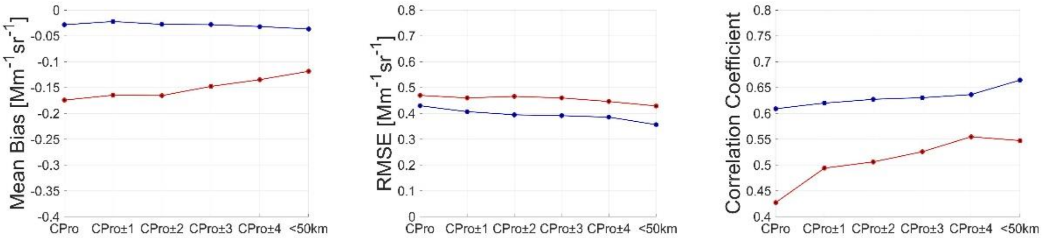

Figure 3: CATS backscatter coefficient at 1064nm with respect to EARLINET ground-based measurements, as a function of the number of L2 Aerosol Profiles used in the CATS spatial averaging, for daytime (red line) and nighttime (blue line) ISS overpasses. Left: Mean Bias [Mm^{-1}sr^{-1}], center: RMSE [Mm^{-1}sr^{-1}] and right: Correlation Coefficient. “CPro” corresponds to the closest CATS profile to the EARLINET station.

(iv) Effect of EARLINET stations topography

In order to study the effect of topography on the CATS profiles the authors separated the participating EARLINET stations into 3 clusters: Continental (Case I – Belsk, Bucharest, Leipzig, and Warsaw), Coastal (Case II – NOA, Athens NTUA, Barcelona, Cabauw, Thessaloniki and Lecce) and Mountainous (Case III – Dushanbe, Evora, Observatory Hohenpeissenberg, Potenza). The three clusters and the characteristics of the stations are given in Table 1. In addition, Figure 4 shows the locations of the participating stations; green circles denote Continental stations, blue circles denote Coastal stations and brown circles denote Mountainous stations. Figure 4 shows, additionally to the geographical distribution of the active EARLINET stations, the daytime/nighttime overpasses of ISS within the evaluation period, between 02/2015 and 09/2016, encompassing the first twenty months of CATS operation. Due to the limited available dataset of CATS-EARLINET cases, the daytime/nighttime approach was not followed in the case of the analysis regarding the effect of topography.

Table 1: Clustering of EARLINET stations with respect to topographical features.

Case I - Continental				
EARLINET Station	Identification Code	Latitude (°N)	Longitude (°E)	Altitude a.s.l. (m)
Belsk	be	51.83	20.78	180
Bucharest	bu	44.35	26.03	93
Leipzig	le	51.35	12.43	90
Warsaw	wa	52.21	20.98	112

Case II - Coastal				
EARLINET Station	Identification Code	Latitude (°N)	Longitude (°E)	Altitude a.s.l. (m)
Athens-NOA	no	37.97	23.72	86
Athens-NTUA	at	37.96	23.78	212
Barcelona	ba	41.39	2.12	115
Cabauw	ca	51.97	4.93	0
Thessaloniki	th	40.63	22.95	50
Lecce	lc	40.33	18.10	30

Case III - Mountainous				
EARLINET Station	Identification Code	Latitude (°N)	Longitude (°E)	Altitude a.s.l. (m)
Dushanbe	du	38.56	68.86	864
Évora	ev	38.57	-7.91	293
Observatory Hohenpeissenberg	oh	47.8	11.01	974
Potenza	po	40.60	15.72	760

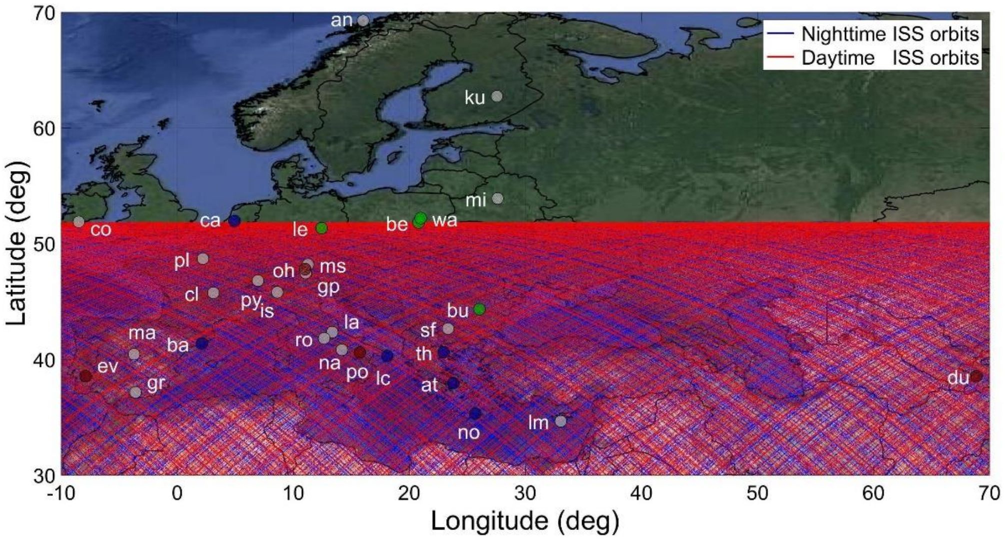

Figure 4: Distribution of EARLINET lidar stations over Europe and West Asia. Green dots: Continental stations used in the inter-comparison. Blue dots: Coastal stations used in the inter-comparison. Brown dots: Mountainous stations used in the inter-comparison. ISS orbits between 02/2015 and 09/2016 are overlaid in red for daytime and in blue for nighttime overpasses.

Figure 5 shows the effect of Topography, for three different clusters of station characteristics, as introduced above (Case I: Continental, Case II: Coastal and Case III: Mountainous). In Figure 5a, the Box and Whisker plot on the $CATS_i$-$EARLINET_i$ residuals is shown, including the lower and upper whiskers which indicate the 10^{th} and 90^{th} percentiles respectively, and the 25^{th} and the 75^{th} quantiles indicated by the lower and upper box boundaries respectively. The horizontal line and the red dot indicate the statistical mean and median values respectively while outliers are indicated by red crosses. According to the results, it is evident that the correlative measurements between the Mountainous EARLINET stations and the ISS overpasses are characterized by higher variability, more extreme differences, higher absolute mean and median biases and higher RMSE than in the Continental and Maritime cases. Complex topography, in terms of geographical characteristics, erroneous mean backscatter coefficient profiles due to the high variability of aerosol load in the Planetary Boundary Layer, the horizontal distance between the CATS lidar footprint and the ground-based lidar stations and surface returns enhance the discrepancies, especially in the lowermost part of the profiles, resulting in higher differences between the EARLINET profiles and CATS profiles. Due to the lack of the aforementioned effects arising from complex topography, CATS representativeness and performance is higher over the Continental cases, while CATS performance over the Coastal stations is characterized by slightly lower absolute value of mean bias and at the same time by lower Correlation Coefficient than in the case of Continental cases. However, it has to be taken into consideration the important factor related to the presented resultsm that is the number of CATS-EARLIENT correlative cases used in the analysis, 23 for Case I - Continental, 10 for Case II - Coastal and 14 for Case III - Mountainous. Analytical evaluation metrics on the effect of topography are given in Table 2.

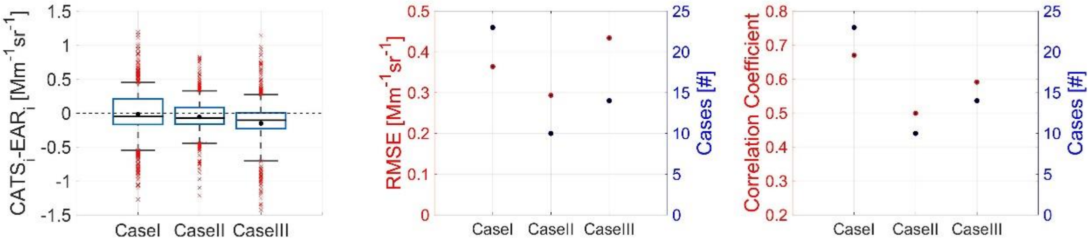

Figure 5: CATS backscatter coefficient at 1064nm with respect to EARLINET ground-based measurements, as a function of different topography of EARLINET stations for three different clusters of station topographical characteristics (Case I: Continental, Case II: Coastal and Case III: Mountainous). In Fig.5a, the Box and Whisker plot on the $CATS_i$-$EARLINET_i$ residuals is shown, including the lower and upper whiskers which indicate the 10^{th} and 90^{th} percentiles respectively, and the 25^{th} and the 75^{th} quantiles indicated by the lower and upper box boundaries respectively. The horizontal line and the red dot indicate the statistical mean and median values respectively while outliers are indicated by red crosses. Fig.5b and Fig.5c show the RMSE and Correlation Coefficient as a function of the different clusters, including the number of available cases per cluster.

Table 2: Clusters of EARLINET stations and CATS evaluation metrics.

	Continental stations	Coastal stations	Mountainous stations
Median	-0.053 [$Mm^{-1}sr^{-1}$]	-0.076 [$Mm^{-1}sr^{-1}$]	-0.106 [$Mm^{-1}sr^{-1}$]
Mean	-0.016 [$Mm^{-1}sr^{-1}$]	-0.058 [$Mm^{-1}sr^{-1}$]	-0.151[$Mm^{-1}sr^{-1}$]
RMSE	0.367 [$Mm^{-1}sr^{-1}$]	0.293 [$Mm^{-1}sr^{-1}$]	0.434 [$Mm^{-1}sr^{-1}$]
Correlation Coefficient	0.673	0.499	0.591
Number of cases	23	10	14