We thank the reviewer for their detailed assessment as well as the numerous specific suggestions for improving clarity and quantitative rigor throughout the manuscript.

This manuscript presents an extensive assessment of AERONET nighttime aerosol optical depth (AOD) retrievals based on lunar photometry and empirical correction of the ROLO lunar irradiance model. The work is timely and important because nighttime AOD observations are increasingly valuable for aerosol monitoring, air quality applications, wildfire smoke tracking, and studies of diurnal aerosol variability. The manuscript is generally well organized and contains a large amount of useful analysis based on long-term global AERONET observations. The effort to evaluate continuity between solar and lunar retrievals across hundreds of sites is particularly valuable.

Major comments:

The current validation strategy relies heavily on continuity between solar and lunar AOD during day–night transition periods under low and stable aerosol conditions (AOD440 ≤ 0.1 with low variability). While this is a reasonable approach for isolating calibration consistency and minimizing atmospheric variability, it may introduce a selection bias toward relatively homogeneous and low-loading atmospheric states. Consequently, the presented analysis primarily demonstrates agreement under favorable retrieval conditions rather than across the broader range of nighttime aerosol environments that are often of greatest scientific and operational interest.

This rationale has been clarified with the following addition to the text.

"There are many environmental reasons why day and night AOD may differ that are unrelated to calibration such as changes in cloud cover, relative humidity and dominant emission sources. In situations where these factors notably change over the 2-3 hour intercomparison window it is not expected that the solar and lunar AOD will agree over this transition. Therefore, these confounding factors have been minimized by looking at low AOD, stable conditions so any bias in calibration can be assessed properly."

The paper also details the absence of curvature check at night (less effective cloud-screening) which could also introduce differences in the lunar data relative to the solar data. Assuming no differences due to clouds, the accuracy of solar and lunar will be equivalent (within the small differences noted) for the same airmass regardless of AOD. The question about consequences of lower signal for lunar data is well considered however, AERONET AOD is only calculated when a measurement has sufficient signal, defined as V_o/1500 for both solar and lunar data (using appropriate Vo for each). The effective $V_0/V_{\min}$ is smaller for lunar observations than for solar observations, indicating that lunar retrievals become noise-limited at lower optical depths particularly for largest lunar phase angle. Thus, as described in the paper, a smaller percentage of lunar AOD may be retained which must be recognized when looking at site statistics.

Although the empirical correction is applied at the irradiance calibration level and is therefore not explicitly AOD-dependent, it is unclear whether retrieval performance remains equivalent under high-AOD conditions where signal levels are substantially reduced and atmospheric variability is often larger. Additional discussion of how calibration uncertainty, low-signal effects, and retrieval bias may vary with aerosol loading would strengthen the manuscript.

In addition, the manuscript later notes that high-AOD lunar observations are more likely to be excluded because of low-signal limitations, particularly at shorter wavelengths. This raises the possibility of an additional sampling bias in the validation dataset.

The validation dataset (includes comparisons from 474 sites) was selected to exclude conditions when sampling bias is a concern. There certainly is a differential upper truncation bias in the full dataset, that is, data systematically missing due to low signal at high AOD for both solar and lunar AOD (Fig 12). This inevitably produces statistical disparity (assymmetric truncation rates) as the solar AOD threshold is higher than that for lunar observations (which varies with air mass).

It would strengthen the manuscript if the authors discussed more explicitly how these filtering criteria may limit the representativeness of the validation framework, particularly for high-AOD, polluted, smoke-affected, or rapidly evolving nighttime conditions. Relatedly, would the reported agreement statistics change materially under elevated AOD conditions?

The following section will be added to the text:

"We are checking for systematic bias between lunar and solar AOD. Even with stability thresholds applied to the 1-hour day/night AOD means, the random component of AOD differences increases substantially with increasing AOD due to real atmospheric aerosol variability over the ~3 hour transition interval."

"Means of AOD differences across the transition window were binned by the paired measurement mean and associated 95% confidence intervals were computed by two methods: a standard normal approximation based on the sample mean and standard error, and bootstrap resampling. The bootstrap method is distribution-free, does not assume Gaussian errors, and accommodates asymmetric sampling distributions — considerations particularly relevant at higher AOD where transition pair counts are lower and AOD difference distributions become increasingly skewed due to episodic aerosol loading. The two methods yield nearly identical confidence intervals across all bins through AOD ~ 0.8, providing mutual validation, with meaningful divergence only at the highest AOD bins ($\geq$ 0.9) where sample sizes are greatly reduced."

"At low AOD ($<$ 0.2), the mean solar-lunar AOD differences are negligibly different, ~0.002 for both 3-hr solar-lunar transition interval directions (day to night, night to day), much less than the estimated AERONET solar AOD instrumental uncertainty of 0.01. Both diurnal transition scatter plots for day AOD vs night AOD inclusive of all AOD $<$ 1.0 (Figs. 7+8) have nearly identical slope ~ 0.95 and CCC (concordance correlation coefficient) ~ 0.96. A slope slightly below unity is consistent with a tendency for daytime AOD to exceed nighttime AOD at elevated aerosol loading. This binned mean difference analysis shows a small but statistically significant higher daytime AOD of 0.013 (night-to-day, CI: [0.011, 0.017]) and 0.016 (day-to-night, CI: [0.012, 0.020]) for AOD_{440nm} bin= [0.6-0.8] which is the highest AOD bin with adequate sample size (n = 2,138 and n = 2,988 for day to night and night respectively). This should not be interpreted as a calibration or measurement difference between solar and lunar instruments, but rather as a plausible real diurnal aerosol signal, consistent with known daytime aerosol production mechanisms (photochemistry, greater daytime emission of anthropogenic and biomass burning aerosols, etc.) and the absence of any corresponding effect at low AOD."

Minor:

Line 43: Vo should be V0

Figure 2: The legend titles should be made consistent (e.g., use either "channel" or "wavelength" for both figures). Additionally, why do the phase-angle dependence patterns differ between the two wavelengths?

The labeling inconsistency has been fixed.

The lunar reflectance varies spectrally (by a factor of ~3x from 400nm-2500nm) and the ROLO lunar irradiance model has an absolute accuracy of 5% or worse with only 10nm spectral resolution that will match to the 10nm bandpass of these channels to differing levels of agreement. The recorded CIMEL signal at 440nm is also significantly less (2-3x) than at 675nm. The omission of an explicit ROLO correction for libration may also introduce some of this small variability. Ultimately the accuracy of the correction is driven by the precision of the solar/lunar langley V_0 ratios at each wavelength. Also, see AOD differences over transition binned by LPA below.

Section 4: The manuscript would benefit from clearer specification of the data quality levels used throughout each analysis section. While Section 5 explicitly references Level 2 (L2) AOD data, it is less clear whether the transition analyses in Section 4 also rely exclusively on L2 observations. It would be helpful to clarify whether "L2" solar and lunar products should be interpreted as equivalent in quality assurance rigor and discuss any implications for the validation results.

A line has been added to the Introduction indicating that only L2 data are used in this paper.

The section 4.2 "Potential effects of differences in cloud-screening for lunar data" describes the absence of the 'cross-scan scenario' (sky radiance measurements across the aureole during solar operation) which is used to identify clouds via the 'curvature' check in the cloud-screening algorithm and the potential for this to result in poorer cloud-screening at night. Thus, the lunar L2 AOD product when considered across all possible conditions would not be uniformly equivalent in quality to L2 solar AOD though the likelihood of lesser accuracy (greater cloud contamination) is seasonally and regionally dependent.

Line 137: The statement that deviations "diminish with increasing wavelength" would be stronger if quantitative wavelength-dependent results or statistics were provided.

The 'deviations' discussed here just refer to the apparent slight oscillation with lunar phase angle that is apparent visually in Figure 5 for 440nm. This effect has been quantified further and the following statements added to the text.

"When assessed in 10° lunar phase angle bins, the mean AOD 440$_{\text{nm}}$ transition differences were within $\pm$ 0.005 for all bins except 80°-90° (0.0072). Mean AOD differences by LPA bin for each wavelength are provided in table S3 in the supplemental section."

Table S3: Mean AOD differences and sample sizes (N) by lunar phase angle (LPA) bin and wavelength channel.

	440 nm		500 nm		675 nm		870 nm		1020 nm		1640 nm	
LPA bin (°)	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
$(-90, -80]$	-0.0019	7650	-0.0001	8104	0.0043	8324	0.0022	8358	0.0004	8177	0.0007	5812
$(-80, -70]$	-0.0007	10117	0.0004	10633	0.0026	10956	0.0006	10961	-0.0013	10683	0.0010	10172
$(-70, -60]$	-0.0014	11633	0.0004	12108	0.0025	12403	0.0002	12358	-0.0015	12234	0.0002	11466
$(-60, -50]$	-0.0016	11973	-0.0001	12403	0.0023	12429	0.0005	12427	-0.0008	12320	0.0005	11549
$(-50, -40]$	-0.0020	13087	-0.0016	13368	0.0014	13528	0.0002	13552	-0.0015	13428	0.0003	12636
$(-40, -30]$	-0.0009	12990	-0.0006	13172	0.0018	13297	0.0008	13266	-0.0007	13253	0.0009	12330
$(-30, -20]$	-0.0013	13645	-0.0006	13815	0.0010	13901	0.0002	13873	-0.0008	13757	0.0004	13075
$(-20, -10]$	-0.0005	18416	-0.0002	18500	0.0004	18651	0.0003	18605	-0.0001	18472	0.0002	17457
$(-10, 0]$	0.0006	20707	0.0008	20787	0.0004	20933	0.0005	20932	0.0005	20702	0.0006	19782
$(0, +10]$	0.0004	22668	0.0002	22677	-0.0004	22842	-0.0002	22853	0.0003	22586	0.0003	21403
$(+10, +20]$	0.0011	20739	0.0013	20819	-0.0012	21036	-0.0001	21018	0.0010	20764	0.0003	19718
$(+20, +30]$	0.0017	15949	0.0020	16180	-0.0015	16312	0.0003	16288	0.0019	16029	0.0005	15518
$(+30, +40]$	0.0020	15489	0.0011	15777	-0.0018	15990	0.0003	15949	0.0022	15757	-0.0001	14892
$(+40, +50]$	0.0042	15492	0.0013	15825	-0.0016	16074	-0.0006	16032	0.0026	15729	0.0007	15155
$(+50, +60]$	0.0000	13994	0.0002	14587	-0.0017	14764	-0.0012	14811	0.0014	14489	-0.0005	13891
$(+60, +70]$	-0.0002	13058	0.0000	13842	-0.0016	14161	-0.0002	14162	0.0008	13761	-0.0008	13353
$(+70, +80]$	0.0045	11522	0.0037	12110	-0.0006	12391	0.0016	12403	0.0017	12060	0.0003	11849
$(+80, +90]$	0.0072	6933	0.0059	7307	-0.0008	7463	0.0025	7455	0.0025	7230	-0.0001	5809

Line146: Figures 7 and 8 would benefit from additional discussion and interpretation. In particular, the spread of the day–night AOD differences appears to increase substantially at higher AOD values. This may reflect more rapidly evolving aerosol conditions during transition periods, but it could also indicate increasing retrieval uncertainty or residual bias under high-loading conditions where lunar signal limitations become more important. Please discuss the likely causes of the increased scatter and whether retrieval performance degrades systematically at elevated AOD.

The following additional paragraphs will be added to the text and the 'Paired AOD scatter by comparison type' figure will be added as supplemental material:

The AOD measurement (not a retrieval) absolute accuracy does not depend on AOD level for CIMEL observations. However, the solar/lunar comparison is made after an interval of several hours when aerosol conditions can change. Higher AOD almost without exception occurs in more variable conditions – often much more variable– so the spread of AOD differences across this time window would be expected to increase (greater standard deviation of AOD difference) though this is mitigated to some degree by the stability check on the computed means. Even a small fractional change in AOD at higher aerosol loading over this interval would manifest as larger individual solar/lunar AOD differences – which would not be the case at low AOD (barring cloud-contamination).

The spread of day-night AOD differences, assessed as the standard deviation of linear regression residuals (observed AOD of the lagged interval minus the value predicted by the linear fit), was computed for 0.1 AOD bins for the combined day/night and night/day transitions as well as the corresponding 3-hour lagged 'within day' and 'within night' comparisons. Notably, all three groups show a progressively increasing standard deviation of AOD differences as AOD increases strongly suggesting this is fundamentally an effect of the increasing aerosol variability. The standard deviation

of the AOD differences for the transition data is higher but quite similar to the lagged 'within day' and 'within night' values (see Table S4 and Fig S2 below which are included in the supplemental section) The lagged AOD 'night only' trend is comparable to the 'day only' trend suggesting the increased scatter is not a product of low lunar signal at higher AOD. Note: The sample sizes are much larger for the 'within day'/'within night' than the transition group and sample size drops off dramatically for the highest AOD bins.

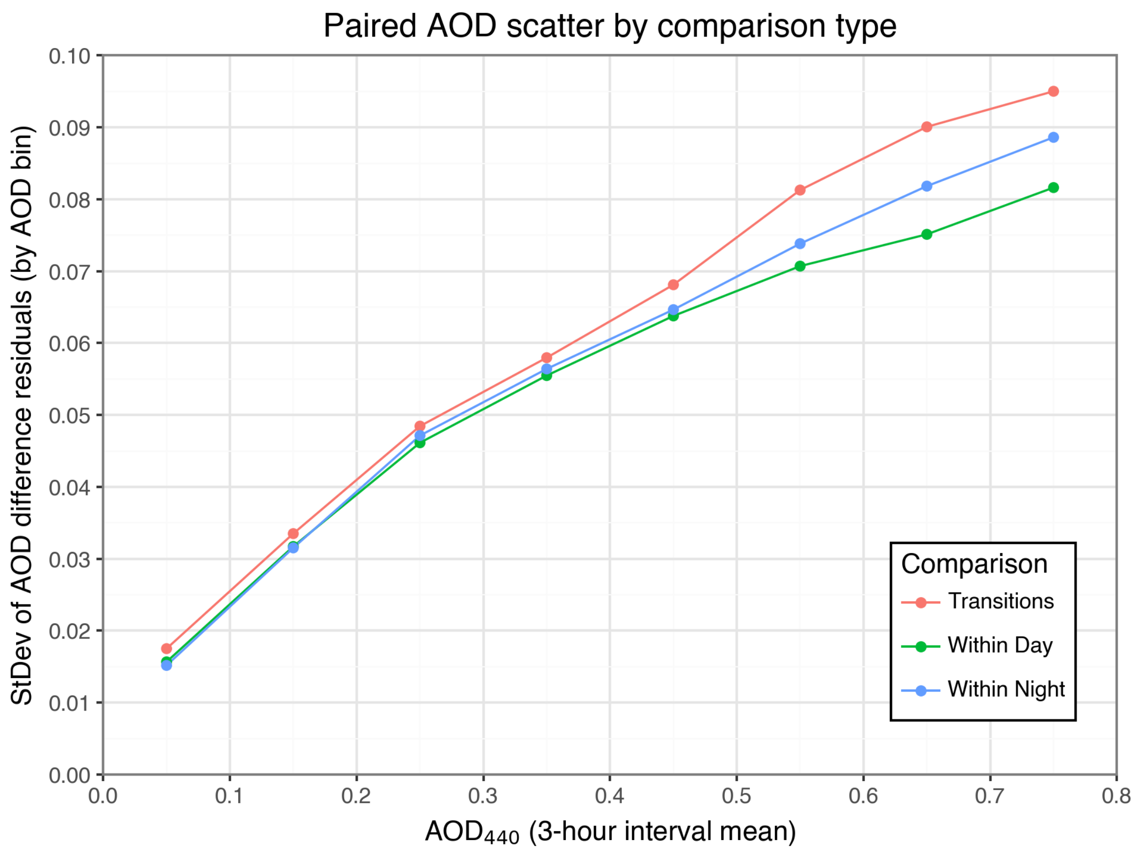

Table S4: Average standard deviation of AOD differences by AOD bin and measurement group.

AOD bin	Transition	Within-day	Within-night
0.0–0.1	0.018	0.016	0.015
0.1–0.2	0.034	0.032	0.032
0.2–0.3	0.048	0.046	0.047
0.3–0.4	0.058	0.056	0.056
0.4–0.5	0.068	0.064	0.065
0.5–0.6	0.081	0.071	0.074
0.6–0.7	0.091	0.075	0.082
0.7–0.8	0.096	0.080	0.086

Line 252: Please provide a quantitative description for "the great majority." More generally, several statements throughout the manuscript rely on qualitative descriptors such as "good agreement," "minor," "very small," "not notably dependent," and "the great majority" without quantitative definition. I recommend providing explicit numerical thresholds, percentages, or statistical metrics wherever possible to improve interpretability and reproducibility.

Quantitative values were added wherever a specific threshold or metric was absent which included the following:

'Despite this selective reduction in higher AOD data, the majority of observations at a given site are for the airmass range from 1-2 (Solar: 58%, Lunar: 65.5%) when the available signal is maximal and sampling bias is not impactful.'

'Good agreement of the solar and lunar AOD (within $\pm$ 0.0075) is also observed for individual sites as depicted in Figure 9'

’The degree of agreement between solar and lunar AOD is not notably dependent on LZA despite lower signal at large lunar zenith while there is some varying structure to the differences in Figure 5 as a function of LPA suggesting the complexities of lunar irradiance are not always perfectly bias-corrected though the deviations are minor, and diminish with increasing wavelength. When assessed in 10° lunar phase angle bins, the mean AOD 440_{nm} transition differences were within $\pm$ 0.005 for all bins except 80°-90° (0.0069). Mean AOD differences for 675nm and longer wavelengths were with $\pm$ 0.002 for all LPA bins.

Some figure labels and color-bar annotations are difficult to read because of small font sizes (e.g., Figures 7 and 8).

Fonts have been increased for Fig 7 and 8 and several others.