Replies to comments of Reviewer 2

Authors would like to express sincere gratitude to an anonymous reviewer for his/her valuable comments and suggestions. Our manuscript has greatly benefited from these insights. We have carefully revised the manuscript, taking all comments into account. Our responses to the reviewer's comments are provided below.

The proposed cloud-screening approach appears physically consistent with the different spectral behavior of aerosols and clouds. Author found that the spectral slope derived from the logarithmic contrast between zenith radiances at 0.400 and 0.675 µm has strong sensitivity to cloudy-sky conditions but limited capability for detecting clear-sky conditions, whereas slopes based on the 0.500 and 0.675 µm and 0.400 and 0.675 µm wavelength pairs exhibits more balanced detection performance for both clear and cloudy-sky conditions. The proposed method cannot be considered a fundamentally new concept, but represents a useful and practical cloud-screening technique to apply to sun-sky radiometers.

→ We sincerely thank the reviewer for this constructive comment and for recognizing the practical utility of the proposed cloud-screening approach. We agree that the physical basis underlying the proposed method, namely the different spectral scattering characteristics of aerosols and clouds, is well established. However, the use of these spectral differences in zenith-sky radiance measurements for sky-condition classification has received limited attention. To the best of our knowledge, the spectral slope of simultaneously measured zenith-sky radiances has not previously been systematically investigated and utilized for discriminating between clear- and cloudy-sky conditions in ground-based sun–sky radiometer observations.

Previous cloud-screening approaches for sun–sky radiometers have primarily relied on retrieved aerosol optical properties, temporal or spectral variability, statistical quality-control procedures, or complementary irradiance measurements (e.g., Smirnov et al., 2000; Khatri and Takamura, 2009). Zenith-sky radiance measurements have also been used for retrieving cloud optical properties (e.g., Chiu et al., 2006; Khatri et al., 2019); however, these studies used zenith radiances as input to cloud-property retrieval algorithms rather than exploiting their spectral slope as a direct clear/cloudy classification criterion. The methodological contribution of the present study therefore lies in systematically characterizing this spectral behavior and developing it into a simple sky-condition classification approach for sun–sky radiometer observations.

Following the reviewers' comments, we have further revised the methodology to explicitly account for the wavelength-dependent instrumental response. The observed spectral slope is separated into atmospheric and instrumental contributions, $m_{obs} = m_{atm} + m_{inst}$, and the estimated instrumental contribution, m_{inst}, is removed before sky-condition classification. The resulting atmospheric spectral slope, m_{atm}, is therefore used as the basis for evaluating the spectral differences between clear- and cloudy-sky conditions while reducing the influence of the wavelength-dependent instrumental response.

The practical advantage of the proposed approach is that sky-condition classification can be performed using the spectral characteristics of zenith-sky measurements without first conducting aerosol or cloud retrievals, temporal variability analyses, or using ancillary observations. Furthermore, although the wavelength-dependent instrumental correction employed in the present study represents a first-order approximation, the observational results demonstrate that m_{atm} retains a distinct separation between clear- and cloudy-sky conditions. Thus, the proposed approach

provides a simple and computationally efficient preprocessing or postprocessing quality-control procedure for sun–sky radiometer observations.
In response to this comment, we have revised the Methodology section to more clearly explain the physical basis and methodological contribution of the proposed approach.

There are some points that the authors should address:

1) In line 138, it is stated that the SKYNET archive is available at atmos3.cr.chiba-u.jp/skynet. However, this is not the official SKYNET website, which is www.skynet-isdc.org. The latter provides two processing modes, namely CERES and ESR-MRI. Are the products downloaded from the former website exactly the same as those available from the official SKYNET portal? If not, it would not be accurate to state that the database is obtained from SKYNET without further clarification.

→ We agree that the original description was ambiguous and could be interpreted as implying that the data were obtained directly from the official SKYNET website. In this study, the observational data were downloaded from the CEReS SKYNET archive (https://skynet.irie-lab.jp), rather than from the official SKYNET ISDC portal (https://www.skynet-isdc.org). Specifically, we used the SR-CEReS processed products distributed through the CEReS archive. To avoid confusion regarding the data source, we have revised the manuscript to explicitly identify the CEReS SKYNET archive as the source of the observational data and clarified that the analyses were performed using the SR-CEReS processed products (**Page 3, Line 143 – Page 4, Line 144**).

2) To what extent does the value of the epsilon parameter (line 406) depend on the measurement site? Please discuss its site dependence and whether a site-specific tuning is required.

→ In the present study, ε is a tolerance parameter used to define the uncertainty margin associated with the sky-condition classification criterion.The primary role of ε is therefore to control the trade-off between classification coverage and confidence. A smaller ε narrows the uncertain region and allows a larger fraction of observations to be classified, but may retain more ambiguous cases near the classification boundary. Conversely, a larger ε provides a more conservative classification by excluding more observations close to the boundary, at the expense of increasing the fraction of undetermined cases.
In the present study, a common value of ε was applied to all observation sites rather than optimizing it separately for each site. The multi-site analysis, covering diverse atmospheric environments, indicates that this common tolerance provides effective discrimination between clear- and cloudy-sky conditions without site-specific tuning. Nevertheless, some site dependence may remain because the distribution and variability of m_{atm} can differ among sites due to differences in atmospheric conditions and measurement uncertainty. Therefore, its value may be adjusted for particular applications depending on the desired balance between classification coverage and confidence.
We have clarified the role of ε, its possible site dependence, and the use of a common value across all observation sites in the revised manuscript (**Page 13, Line 488 – Page 14, Line 463**).

Here, the parameter ε controls the degree of conservativeness of the classification. A smaller ε results in a narrower undetermined region and allows more observations to be classified as clear or cloudy, whereas a larger ε broadens the undetermined region and provides a more conservative classification. The value of ε may be adjusted according to the desired balance between classification coverage and confidence or to meet specific operational requirements.

3) Instead of the qualitative comparison presented in Figure 9, I recommend providing a quantitative assessment of the different cloud-screening methods. For example, the authors could report the percentage or number of data points rejected by each method.

→ In response to this comment, we have replaced the primarily qualitative comparison in the original manuscript with an expanded quantitative evaluation by including additional Table 4 and Figures 12 and 13 in Section 4.3 of the revised manuscript.

4) The proposed method should be evaluated under a wider range of atmospheric conditions.In particular, it would be useful to compare the performance of the different cloud-screening approaches during representative case studies, such as desert dust outbreaks, wildfire smoke events, and cirrus cloud conditions, in order to demonstrate any improvement achieved by the proposed method.

→ We agree that evaluation under a wider range of atmospheric conditions is important for assessing the robustness and applicability of the proposed cloud-screening method. In response to this comment, we have substantially expanded the evaluation in the revised manuscript.

Specifically, we compared aerosol optical thickness (AOT) at 0.5 μm and Ångström exponent (AE) obtained after application of the proposed m_{atm}-based cloud-screening method with those obtained using the SR-CEReS cloud-screening procedure at eight SKYNET sites representing substantially different atmospheric environments. These include five Japanese sites, a tropical site (Burgos), and two polar sites (Syowa and Ny-Ålesund), thereby providing an evaluation over a wide range of aerosol loadings and atmospheric conditions.

In addition to the AOT–AE relationship, we have added probability-density distributions of AOT and AE to provide a more quantitative comparison of the aerosol populations retained by the two screening approaches. Overall, the distributions obtained using the proposed method substantially overlap those obtained using SR-CEReS at most sites, indicating that the proposed screening method retains broadly similar aerosol populations under the different atmospheric environments examined.

We further evaluated the consistency of the two approaches under seasonally varying atmospheric conditions. The revised manuscript now includes a seasonal comparison of the number of retained observations and the mean ± standard deviation of AOT and AE for both screening methods. The five Japanese sites were evaluated separately for DJF, MAM, JJA, and SON, while seasonally appropriate periods were used for Burgos, Syowa, and Ny-Ålesund. In most cases, the seasonal mean AOT and AE values obtained using the proposed method are close to those obtained using SR-CEReS and remain within the corresponding variability. Some differences are observed at individual sites and seasons, particularly at Fukue during periods of relatively high aerosol loading, indicating that the two screening approaches do not classify all observations identically. Nevertheless, the overall AOT and AE distributions remain broadly consistent.

Taken together, these multi-site, distributional, and seasonal comparisons substantially broaden the range of atmospheric conditions under which the proposed method is evaluated and indicate that its performance is generally consistent with SR-CEReS across the different environments examined.

We agree with the reviewer that dedicated case studies of desert-dust outbreaks, wildfire-smoke events, and thin-cirrus conditions would provide an additional and more stringent event-specific evaluation. However, such an assessment requires independent identification of aerosol and cloud types using coincident observations, such as lidar measurements, satellite aerosol/cloud products,

or other ancillary datasets. Such observations are not consistently available across the eight sites and analysis periods considered in the present study. Therefore, we do not interpret the present results as an event-by-event validation for these specific conditions. We consider such dedicated event-specific validation an important direction for future research.
Accordingly, the revised manuscript now includes the expanded multi-site comparison (Fig. 11), the AOT and AE probability-density comparisons (Figs. 12 and 13), the seasonal quantitative comparison (Table 4), and an expanded discussion of the performance and limitations of the proposed method under diverse atmospheric conditions in Section 4.3

5) The title should clearly specify that the proposed cloud-screening method applies only to direct-sun measurements.

→ We thank the reviewer for this comment. We would like to clarify that the proposed cloud-screening method does not use direct-sun measurements. Rather, the method is based on the spectral characteristics of zenith-sky radiances measured by a surface-based sun–sky radiometer. Specifically, sky-condition classification is performed using the atmospheric spectral slope, m_{atm} derived from simultaneous zenith-sky measurements at different wavelengths after accounting for the wavelength-dependent instrumental contribution. Direct-sun measurements are not used as input to the proposed sky-condition classification.
We recognize that the use of direct-sun-derived aerosol products in the subsequent evaluation of the cloud-screening results may have caused this misunderstanding. However, these aerosol products are used only to evaluate the observations retained after cloud screening and are not part of the proposed classification method itself.

References

Chiu, J. C., Huang, C.-H., Marshak, A., Slutsker, I., Giles, D. M., Holben, B. N., Knyazikhin, Y., and Wiscombe, W. J.: Cloud optical depth retrievals from the Aerosol Robotic Network (AERONET) cloud mode observations, J. Geophys. Res., 115, D14202, https://doi.org/10.1029/2009JD013121, 2010.

Khatri, P., Iwabuchi, H., Hayasaka, T., Irie, H., Takamura, T., Yamazaki, A., Damiani, A., Letu, H., and Kai, Q.: Retrieval of cloud properties from spectral zenith radiances observed by sky radiometers, Atmos. Meas. Tech., 12, 6037–6047, https://doi.org/10.5194/amt-12-6037-2019, 2019.

Khatri, P. and Takamura, T.: An algorithm to screen cloud-affected data for sky radiometer data analysis, Journal of the Meteorological Society of Japan, 87, 189–204, https://doi.org/10.2151/jmsj.87.189, 2009.

Smirnov, A., Holben, B. N., Eck, T. F., Dubovik, O., and Slutsker, I.: Cloud-screening and quality control algorithm for the AERONET database, Remote Sens. Environ., 73, 337–349, https://doi.org/10.1016/S0034-4257(00)00109-7, 2000.