

S1 Identification of unavailable measurements in the raw spectra

During the quality control of the raw spectra, two types of measurements that were unsuitable for subsequent retrievals were identified and removed.

5 The first type corresponds to a persistent abnormal decrease in the UV spectral signal. From 9 September to 13 October 2025, the DN values measured by the instrument in the UV spectral region were clearly lower than those during the adjacent normal periods. Fig. S1(a) compares raw spectra measured at the same time of day before and during this abnormal period. Compared with the normal measurement at 11:30 LT on 8 September 2025, the spectrum measured at 11:30 LT on 9 September 2025 shows a systematic decrease in detector signal over approximately 300–430 nm, with a stronger reduction at shorter UV wavelengths. Fig. S1(b) further shows that, during this continuous abnormal period, the DN values at 330 and 400 nm both
10 deviated substantially from their normal levels, and the 330 nm signal remained close to the low-DN range for an extended period. Therefore, the measurements from 9 September to 13 October 2025 were regarded as unavailable and were excluded from subsequent aerosol and trace-gas retrievals.

15 The second type corresponds to sudden decreases in detector signal within individual days. One example occurred on 18 May 2025. As shown in Fig. S1(c), the signal at 400 nm decreased abruptly to near-zero values after about 09:00 LT and did not recover to the normal diurnal level during the following measurement sequences. This abnormal decrease was caused by a malfunction of the instrument control software, which resulted in incorrect shutter/light-path control and prevented normal light collection by the spectrometer. For such cases, only the valid observations before the onset of the abnormal decrease were retained, while the low-DN measurements after the abnormal decrease were removed.

S2 Example of QDOAS spectral fitting

20 To illustrate the DOAS spectral fitting procedure, Fig. S2 shows a typical QDOAS fitting example in the 338–370 nm fitting window. The figure presents the matching between the measured and reference spectra, the fitting residual, the slit-function parameter (SFP), and the fitted differential absorption structures of O_4 , NO_2 , and the Ring effect. The measured and reference spectra show generally consistent spectral structures within the fitting window, indicating acceptable wavelength alignment and reference-spectrum matching in this example. The fitting residual is 1.18×10^{-3} , and the O_4 and NO_2 components show
25 identifiable differential absorption structures. This indicates that the fitting result is suitable for subsequent dSCD calculation and aerosol retrieval analysis.

S3 Pre-processing of derived quantities

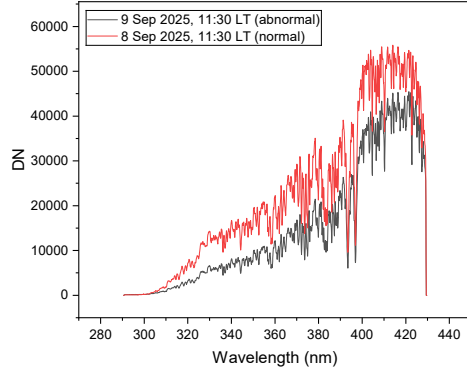
S3.1 Smoothing of the zenith signal before TSI calculation

30 The measured zenith signal $I_{\text{obs}}(400 \text{ nm})$ is defined as the zenith DN normalized by the integration time in milliseconds (DN ms^{-1}). A small residual tilt of the instrument can introduce systematic differences among zenith observations acquired at different azimuth settings. During the eight-azimuth scan sequence, these differences can produce a repeating high–low signal pattern associated with the installation-related zenith-pointing deviations. This pattern is particularly evident around noon under small-SZA conditions and can artificially increase the TSI, leading to false cloud identification.

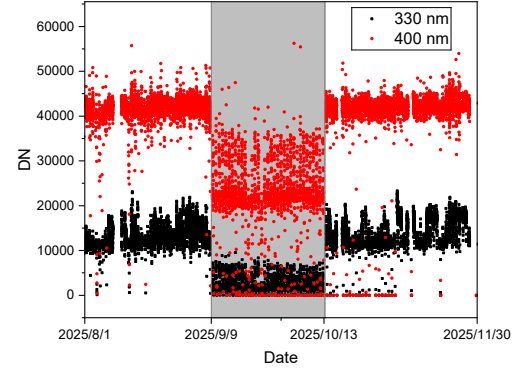
35 Fig. S3 shows an example of the diurnal variation of the zenith signal at 400 nm on 9 February 2025. During the high-signal period around noon, the unsmoothed signal sequence shows a clear eight-azimuth structural pattern. The amplitude of this structure is usually smaller than 100 DN ms^{-1} , but it can vary with SZA and aerosol loading. Although this variation is generally weaker than cloud-induced abrupt changes or persistent signal reductions, it can increase the short-term variability used in the TSI calculation and may therefore lead to cloud-free conditions being misclassified as cloudy conditions.

40 To reduce this instrumental contribution, preprocessing was restricted to periods in which the repeating azimuth-related high–low pattern was identified. Within these selected periods, an 11-point centred moving average was applied to the zenith signal before calculating the TSI. Each complete window includes the current measurement and five measurements on either side; fewer available measurements are used at the sequence boundaries. Observations outside the selected periods were left

(a)



(b)



(c)

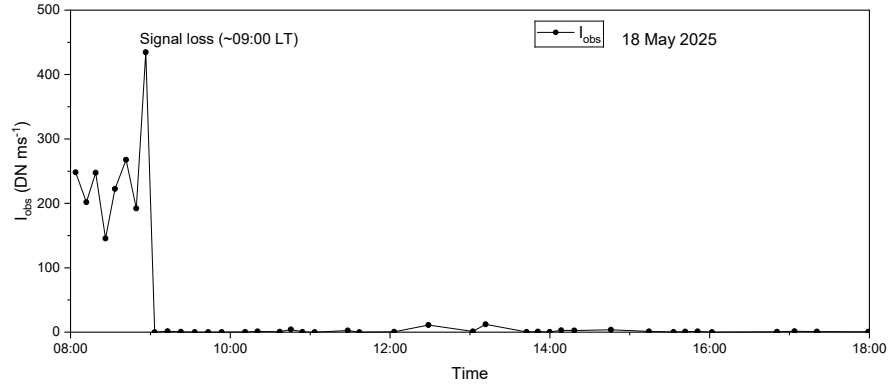


Figure S1. Examples of unavailable MAX-DOAS measurements identified during raw-spectra quality control. (a) Raw spectra before and during the persistent UV-signal abnormality. (b) DN time series at 330 and 400 nm, with the grey shading indicating the affected period from 9 September to 13 October 2025. (c) Example of a sudden within-day signal decrease on 18 May 2025, shown by the integration-time-normalized zenith signal I_{obs} at 400 nm (DN ms⁻¹).

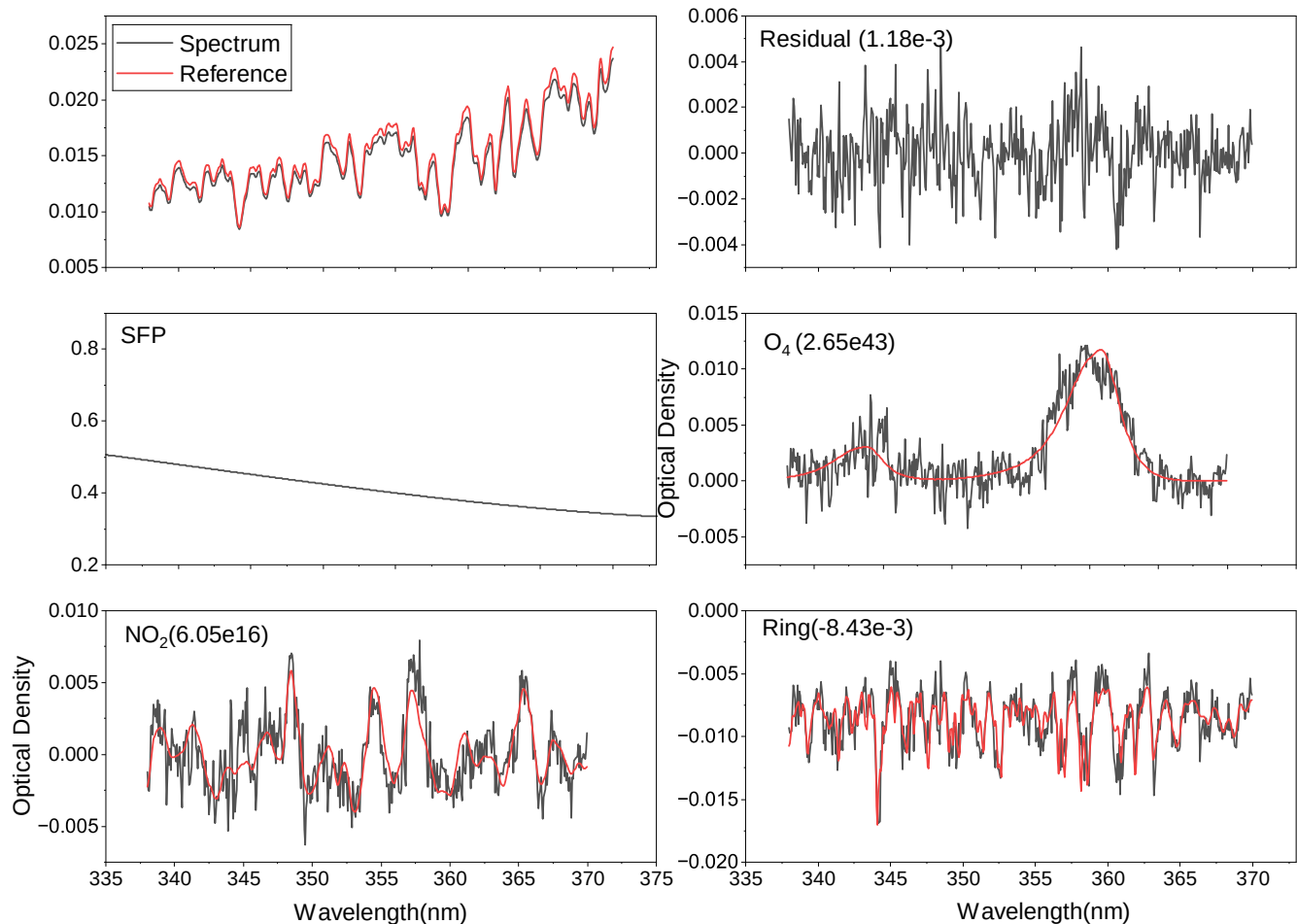


Figure S2. Example of QDOAS spectral fitting in the 338–370 nm fitting window, including the measured and reference spectra, fitting residual, SFP, and fitted O_4 , NO_2 , and Ring structures.

unchanged. Smoothing was applied only when the repeating instrumental pattern was identified, while cloud-related fluctuations without this pattern were retained. This treatment reduced the installation-related structure while preserving cloud-related temporal variability. The TSI classification threshold remained unchanged.

S3.2 Pre-processing of the 30° O_4 dSCD

The 30° O_4 dSCD is used to distinguish low-cloud from high-cloud conditions relative to the classification threshold. Low clouds generally shorten the effective light path and reduce this quantity, whereas cloud-free and high-cloud observations usually remain above the threshold. Geometrically induced local decreases can therefore cause cloud-free or high-cloud observations to be misclassified as low cloud.

As the solar position changes, the relative geometry between the Sun and the viewing direction changes the effective light path. This can produce approximately arch-shaped variations in the 30° O_4 dSCD, with local decreases in observations acquired at particular azimuth settings. The affected settings depend on the solar position rather than on a fixed azimuth-dependent

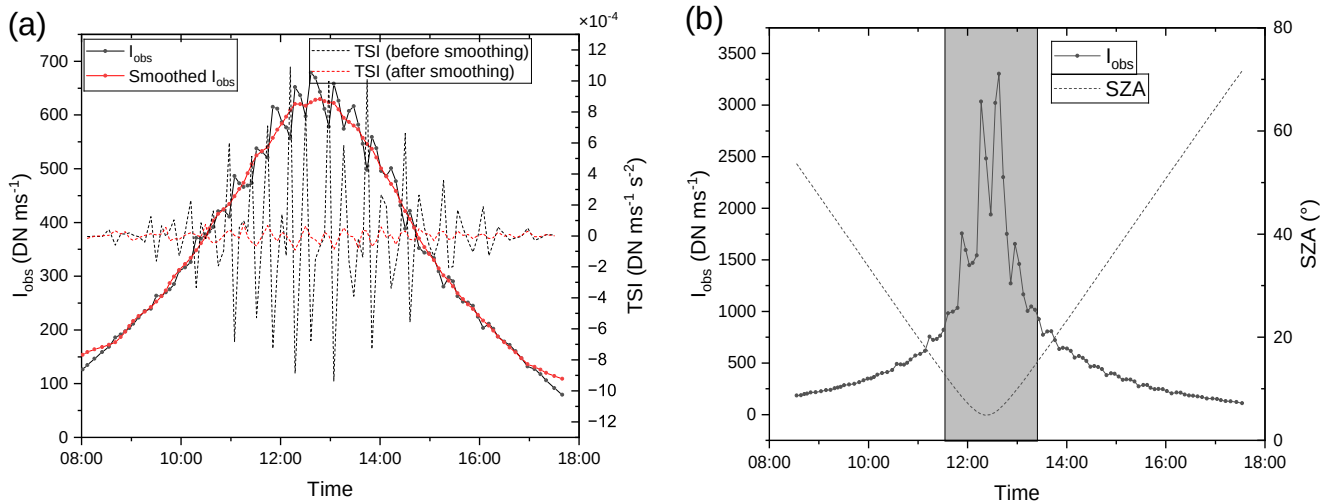


Figure S3. Zenith-signal smoothing before TSI calculation. (a) Zenith signal I_{obs} at 400 nm (DN ms^{-1}) before and after smoothing, together with the corresponding TSI values ($\text{DN ms}^{-1} \text{s}^{-2}$). (b) Zenith signal and SZA, showing the repeating eight-azimuth signal pattern around noon.

instrumental offset. Individual local values can fall below the classification threshold even when the surrounding observations remain above it.

Preprocessing targeted local low values identified as part of this solar-viewing-geometry-related pattern, rather than all observations below the threshold. These selected values were treated locally using neighbouring higher O_4 dSCD values as a reference. The purpose was to reduce false low-cloud identification caused by shortened observation paths in cloud-free or high-cloud conditions, rather than to remove the sustained O_4 response associated with low clouds. The O_4 classification threshold was unchanged.

S4 Supplementary overview of year-round MAX-DOAS aerosol retrievals

Fig. S5 shows the year-round MAX-DOAS-retrieved aerosol extinction profiles, illustrating the temporal coverage of the retrieval dataset. The profiles include observations under different sky conditions and therefore contain both aerosol variability and cloud-induced retrieval responses.

65 S5 Supplementary overview of year-round lidar extinction profiles and cloud-parameter statistics

S5.1 Year-round lidar extinction profiles

Fig. S6 shows the year-round 355 nm lidar extinction coefficient profiles, covering both daytime and nighttime observations. The profiles illustrate the temporal evolution and vertical structure of aerosol and cloud layers. Near-surface extinction variations indicate changes in lower-layer aerosols, while enhanced extinction aloft may be associated with aerosol or cloud layers.

70 S5.2 Cloud parameters identified using the threshold method

Fig. S7 shows the distributions of CBH, COD, and cloud geometrical thickness for cloud samples identified from the lidar extinction profiles using the threshold method. CBH corresponds to the base of the lowest detected cloud layer; COD and

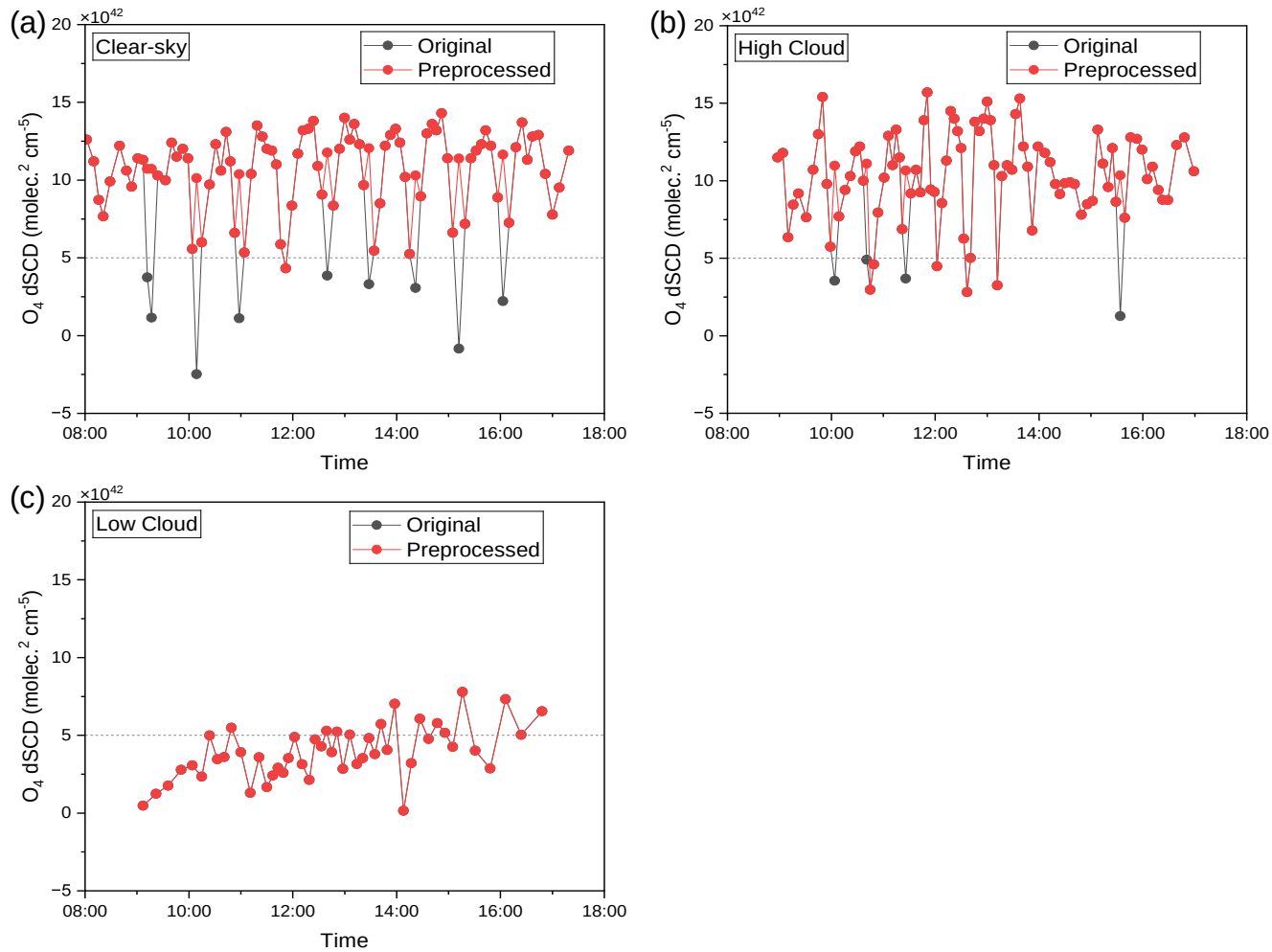


Figure S4. Example of preprocessing for the 30° O_4 dSCD under clear-sky, high-cloud, and low-cloud conditions. Black points indicate the original values, and red points indicate the preprocessed values. The dashed line marks the classification threshold.

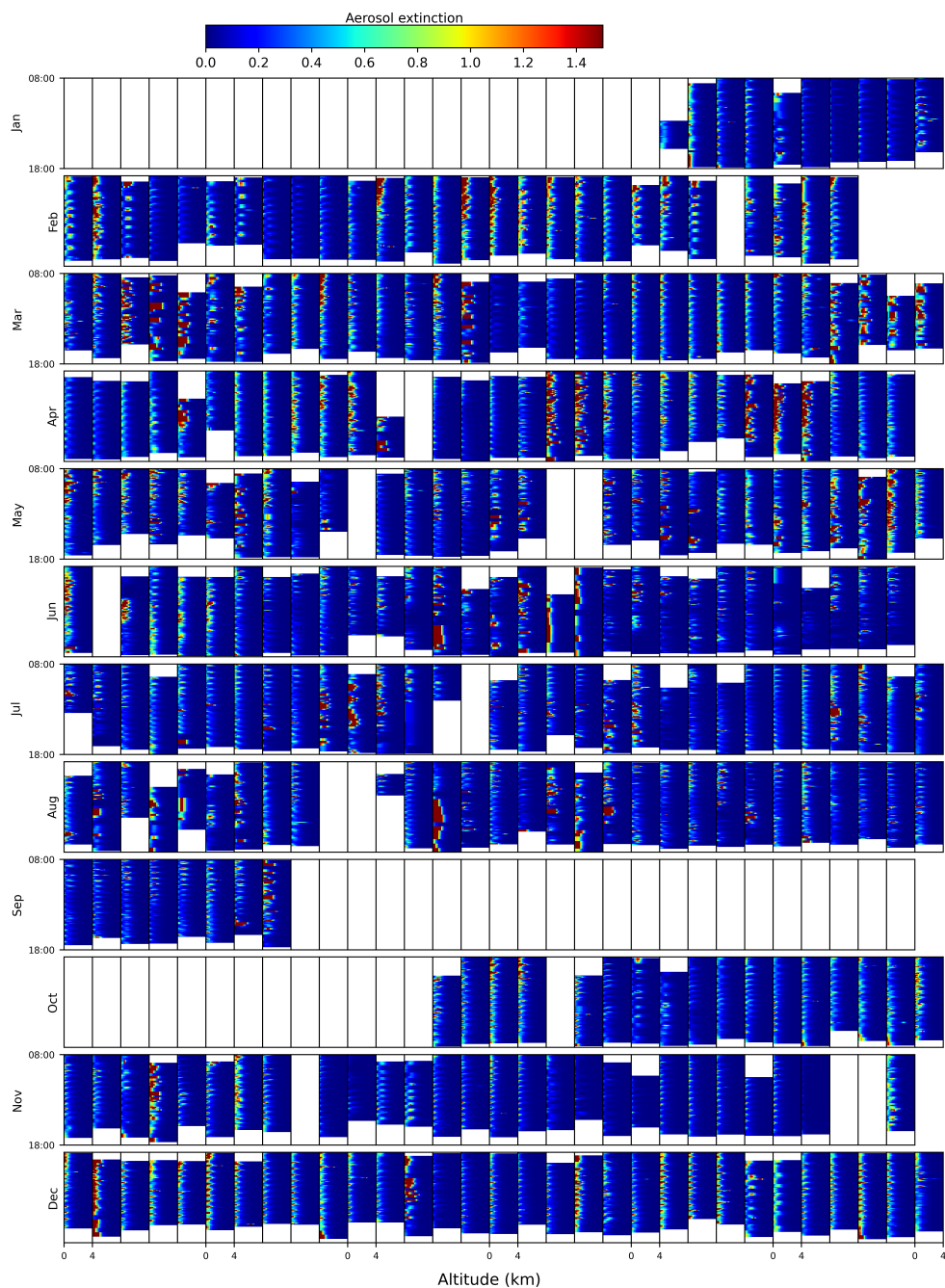


Figure S5. Year-round MAX-DOAS-retrieved aerosol extinction profiles. Colours indicate aerosol extinction, and white areas denote periods with missing or unavailable retrievals. The profiles include observations under different sky conditions and show both aerosol variability and cloud-induced retrieval responses.

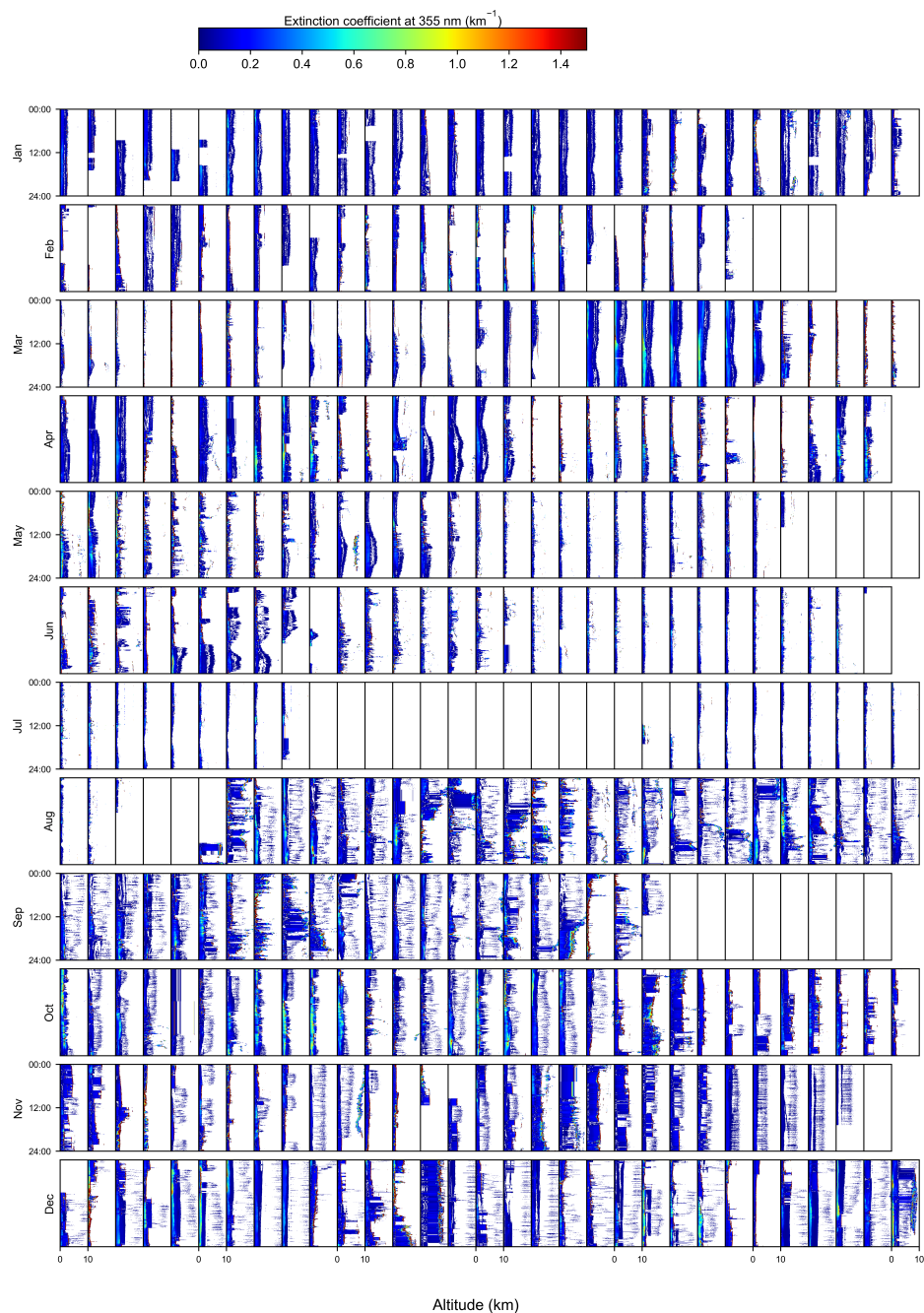


Figure S6. Year-round 355 nm lidar extinction coefficient profiles. Colours indicate the extinction coefficient (km⁻¹). White areas denote missing or unavailable data and extinction values below 0.01 km⁻¹, which were masked for display.

geometrical thickness were derived from the identified cloud-layer ranges. These distributions characterize the detected cloud samples during the observation period.

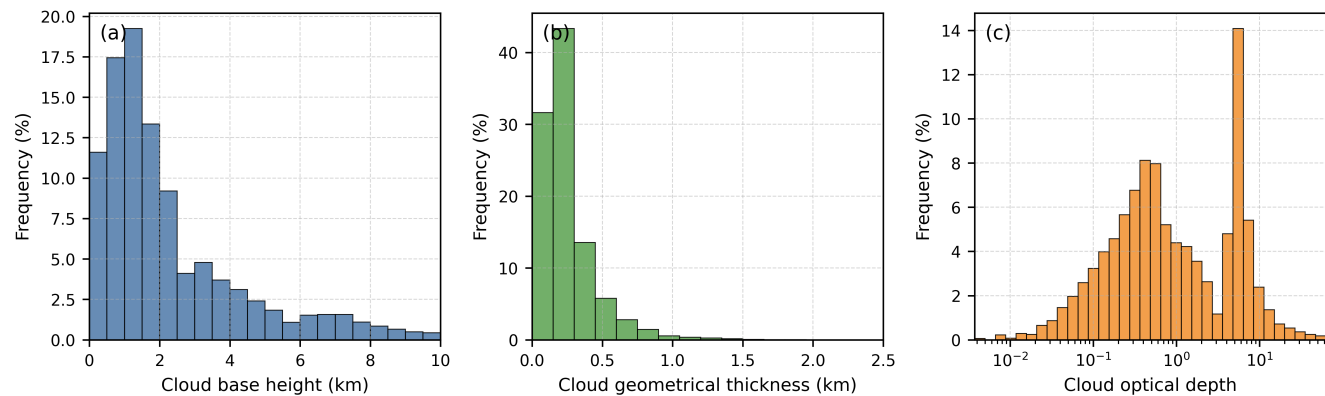


Figure S7. Distributions of (a) CBH, (b) cloud geometrical thickness, and (c) COD for lidar cloud samples identified using the threshold method.