Supplementary Information for: Multi-year ground-based multi-wavelength Raman lidar aerosol dataset over the interior Tibetan Plateau (Yangbajing, 4284 m a.s.l.; 2021-2024)

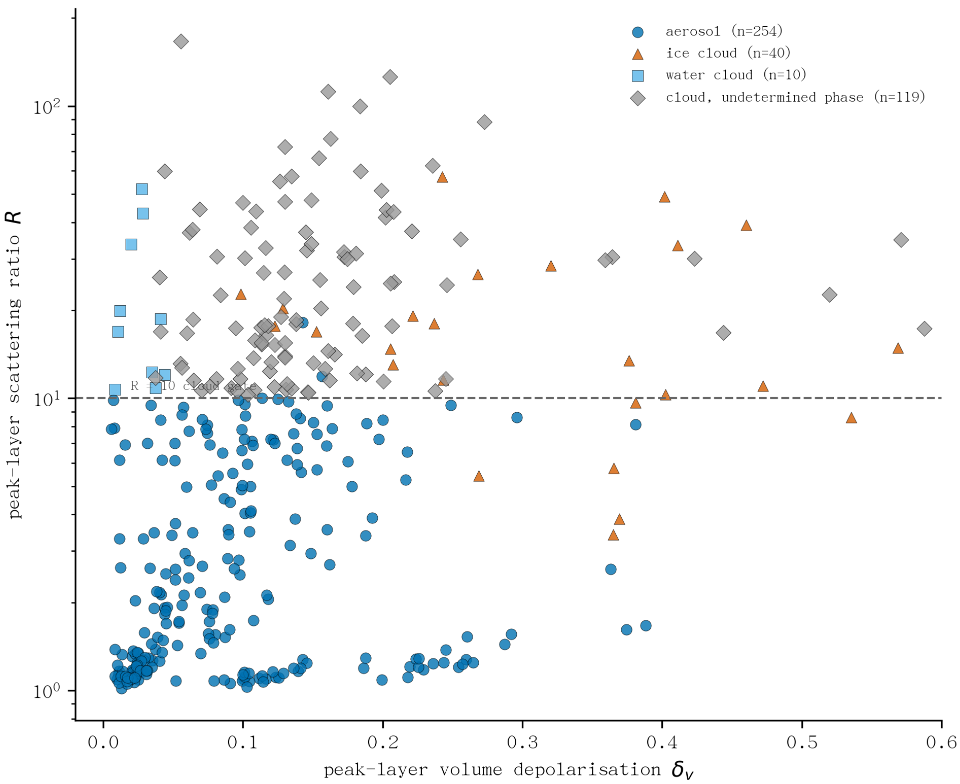

Figure S1. Cloud–aerosol discrimination in scattering-ratio–depolarisation feature space (Sect. 4.1). Each point is the peak-scattering layer of one night: its 532 nm peak-layer scattering ratio R (logarithmic, 3-bin median-filtered so the peak reflects a sustained layer) versus its volume depolarisation ratio δ_v, coloured by the archived classification (aerosol, ice cloud, water cloud, and cloud of undetermined phase; sample sizes in the legend, over the calibrated-δ_v nights). Aerosol layers, including the densest dust and background-aerosol nights, stay at or below the R = 10 dense-cloud gate (dashed; 95th percentile 8.9, maximum 10.0), whereas dense clouds lie above it. The cloud points below R = 10 are ice, flagged either by the cold-cirrus criterion (localised R peak, $\delta_v > 0.2$, T < −40 °C) or, for warmer mixed-phase ice, by the particle-depolarisation gate ($\delta_p > 0.45$; Sect. 4.1). Water clouds cluster at low δ_v (spherical droplets) and ice at higher δ_v, consistent with the phase assignment.

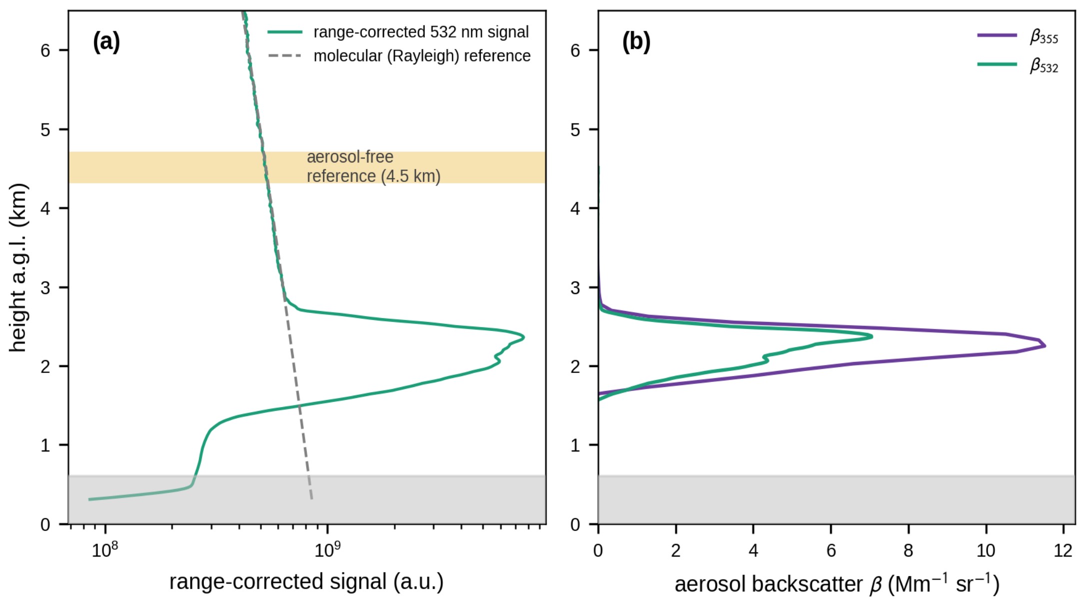

Figure S2. Elastic backscatter retrieval (Sect. 3.2) for a representative clean night (18 October 2023). (a) Range-corrected 532 nm signal and the molecular (Rayleigh) reference; the aerosol-free reference layer (~4.5 km, where the total signal matches the molecular profile) sets the Fernald boundary condition. (b) The retrieved aerosol backscatter coefficients β_{355} and β_{532}, showing a single low-tropospheric aerosol layer. The grey band marks the incomplete-overlap zone (<0.6 km a.g.l.).

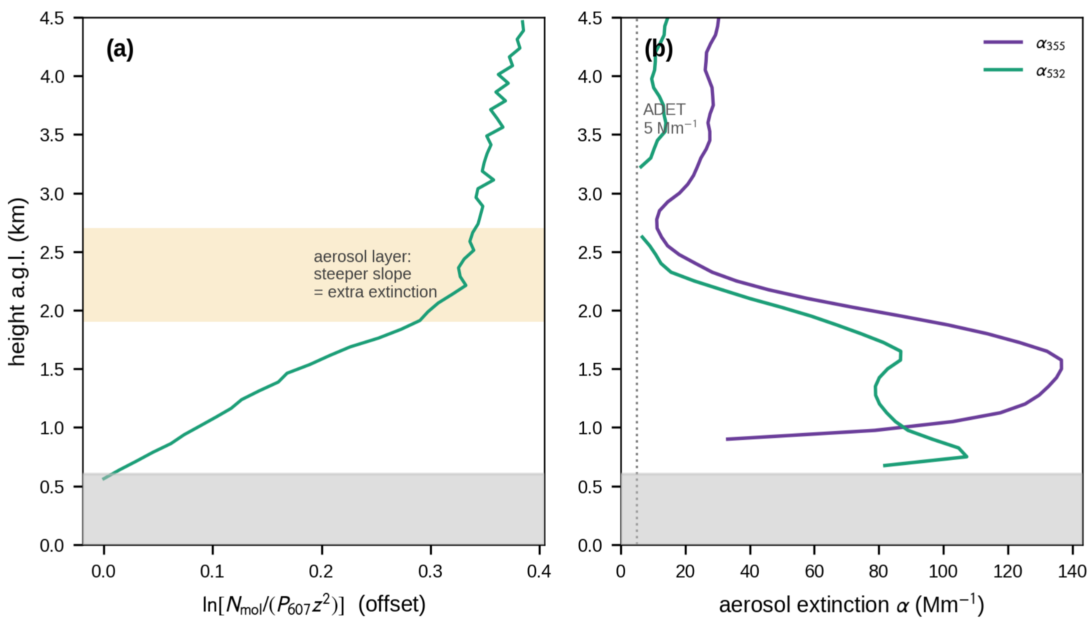

Figure S3. Nitrogen-Raman extinction retrieval (Sect. 3.4) for the 8 December 2024 dust event of Fig. 3. (a) The log signal ratio ln[N_mol/(P_{607} z^2)]; its height-derivative is the total extinction, and the aerosol layer (shaded) steepens the slope. (b) The retrieved aerosol extinction α_{355} and α_{532} with the 5 Mm^{-1} absolute detection threshold (ADET, dotted); bins below threshold are withheld. The extinction reaches ~100 Mm^{-1} in the dust layer. The grey band marks the incomplete-overlap zone.

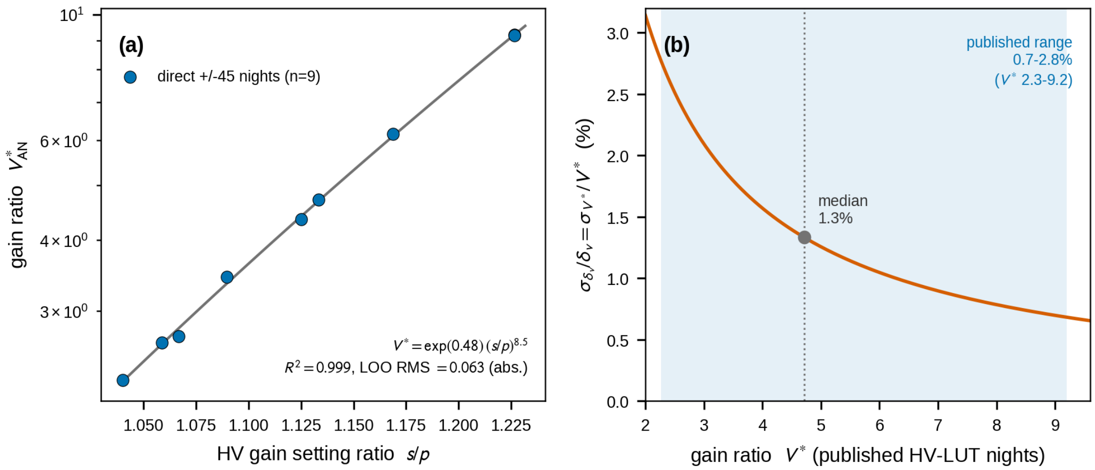

Figure S4. Depolarisation-calibration uncertainty (Sect. 3.5). (a) The PMT high-voltage gain law $V^* = \exp(a)(s/p)^b$ fitted to the nine direct +/-45 degree calibration nights ($R^2 = 0.999$); leave-one-out cross-validation gives an absolute RMS error of 0.063 in the predicted gain ratio V*. (b) Because δ_v scales inversely with V*, the resulting relative depolarisation-calibration uncertainty σ_{V^*}/V *decreases with V* from ~2.8 % to ~0.7 % (median ~1.3 %) across the published HV-LUT nights (V* = 2.3 to 9.2, shaded), rather than being a single fixed percentage.

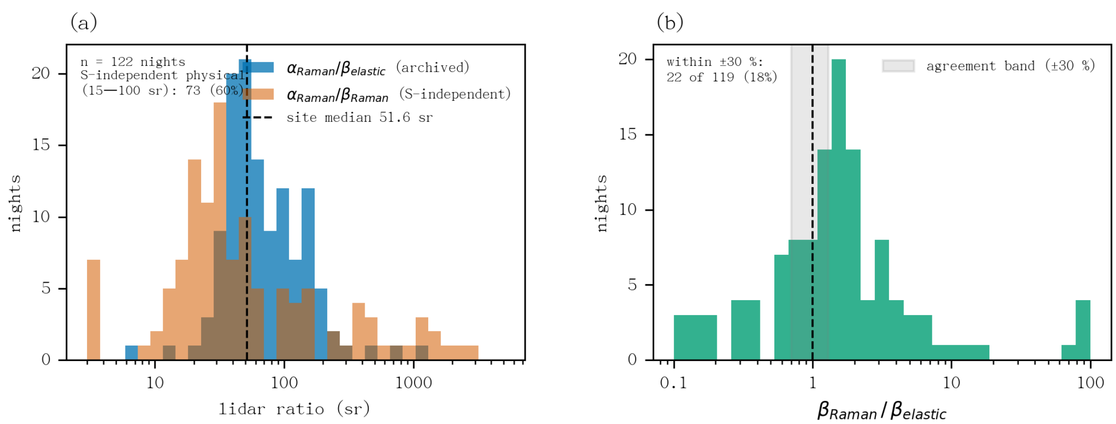

Figure S5. Feasibility of an S-independent Raman (Ansmann) backscatter at this clean site (Sect. 10.1), from the highest-quality nights (n = 122). (a) Distribution of the 1 to 5 km column lidar ratio computed as Raman extinction over the archived elastic backscatter ($\alpha_{Raman}/\beta_{elastic}$; blue) and over the self-calibrated Raman backscatter ($\alpha_{Raman}/\beta_{Raman}$; orange); the elastic-based ratio is compact around the site median (51.6 sr, dashed line), whereas the S-independent ratio spans two orders of magnitude and is physically plausible (15 to 100 sr) on only about 60 % of nights. (b) Distribution of the ratio of the Raman to the elastic backscatter; the two agree within 30 % (shaded band) on 18 % of nights. The instability reflects the low aerosol signal at this site, not an instrumental fault (Sect. 10.1).

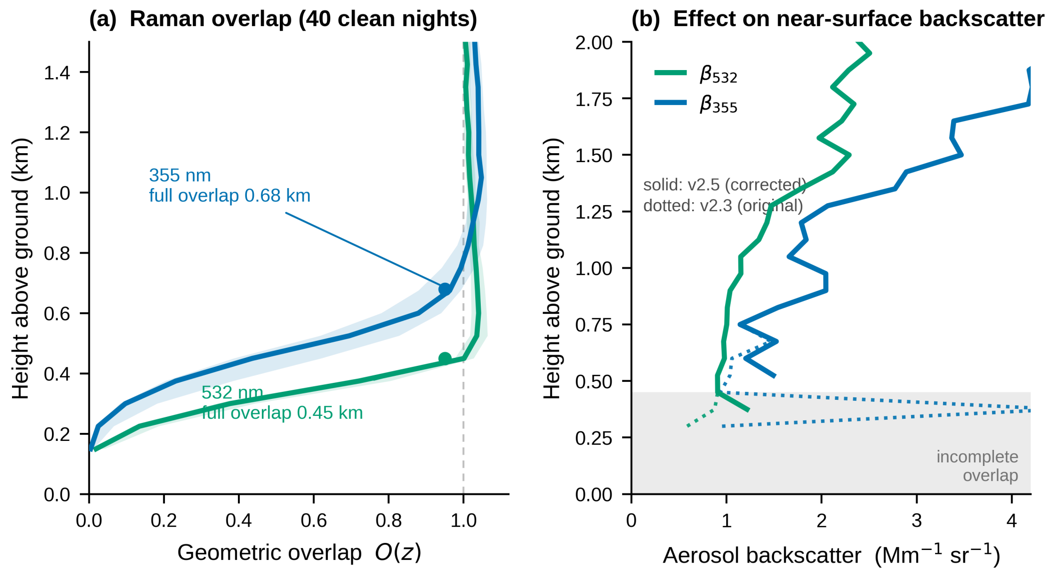

Figure S6. Experimental geometric-overlap correction (Sect. 4.2). (a) Overlap function O(z) at 355 and 532 nm, retrieved by the nitrogen-Raman method (Wandinger and Ansmann, 2002; Comeron et al., 2023) from an ensemble of 40 clean nights (column AOD at 532 nm below 0.05) and normalised in the aerosol-free 3 to 5 km layer; solid lines are the ensemble medians and shading the interquartile range. The overlap reaches unity (markers) by 0.45 km a.g.l. at 532 nm and 0.68 km at 355 nm. (b) All-night median near-surface aerosol backscatter before (v2.3, dotted) and after (v2.5, solid) the overlap correction; the correction lifts only the lowest range bins and leaves the free-tropospheric profile unchanged. Bins where O < 0.5 are set to _FillValue; the shaded band marks the near-field zone below the 532 nm full-overlap height.

Table S1. Raman-derived geometric overlap function O(z) (ensemble median of 40 clean nights) and its robustness to the Raman signal-to-noise ratio (SNR). The full-overlap height (O >= 0.95) is 0.45 km at 532 nm and 0.68 km at 355 nm; restricting the ensemble to the higher-SNR half of the nights leaves these values essentially unchanged (532 nm: 0.45 km; 355 nm: 0.75 km), confirming that the channel-to-channel difference is not a signal-to-noise artefact.

Height (km a.g.l.)	**O(z), 532 nm**	**O(z), 355 nm**
0.45	1.00	0.44
0.525	1.04	0.70
0.60	1.04	0.88
0.675	1.04	0.96
0.75	1.03	0.99
Full overlap (O>=0.95)	0.45 km	0.68 km
high-SNR half	0.45 km	0.75 km
low-SNR half	0.45 km	0.68 km