

Seasonality, not monthly stability anomalies, organizes aerosol vertical regimes over Central Asia

Abduxoliq Ashuraliyev

Jack00040008@outlook.com

Independent researcher <https://orcid.org/0009-0003-5482-5526>

Article

Keywords: Central Asia, AERONET, IGRA, ERA5, MERRA-2, CALIOP, aerosol optical depth, boundary-layer stability, dust

Posted Date: August 12th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10389296/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Seasonality, not monthly stability anomalies, organizes aerosol vertical regimes over Central Asia

Abduxoliq Ashuraliyev
Independent researcher
Tashkent, Uzbekistan
Jack00040008@outlook.com
ORCID: 0009-0003-5482-5526

July 16, 2026

Abstract

Central Asia connects desert dust sources, polluted valleys and mountain snow, yet vertically resolved aerosol–meteorology constraints are sparse. We build a station-screened radiosonde lower-tropospheric stability record and join it with AERONET, ERA5, MERRA-2, CALIOP and OpenAQ. Aerosol size and vertical structure form strong seasonal regimes, but monthly stability anomalies do not independently explain fine-mode variability once the seasonal cycle is removed: the raw stability–fine-fraction correlation of $r = 0.81$ falls to $r = -0.15$. CALIOP places dust higher in coarse dust-transport months than in near-surface stagnation months (2.3 versus 1.6 km), a separation that persists when event classes are defined from meteorology and column aerosol alone and that a GEOS-Chem experiment supports. HYSPLIT back-trajectories and a GEOS-Chem tagged-tracer run produce similar transport-connectivity partitions, and a native-emission-weighted experiment identifies the Aral–Caspian and Karakum–Kyzylkum sectors as the dominant source regions. The result is an observation-constrained, bounded baseline for Central Asian aerosol regimes.

Keywords: Central Asia; AERONET; IGRA; ERA5; MERRA-2; CALIOP; aerosol optical depth; boundary-layer stability; dust

1 Introduction

Atmospheric aerosols over Central Asia connect processes that are usually studied separately: mineral dust emission from arid basins, combustion aerosol accumulation in populated valleys, long-range transport across the Eurasian interior, and deposition onto mountain snow and ice. The Aralkum, Kyzylkum, Karakum, Taklamakan and surrounding drylands are active dust source regions whose emissions perturb regional radiation and transport [1, 2, 8, 18, 21, 23]. Mineral dust also affects climate and snow albedo beyond its source region, including documented Taklamakan dust-storm darkening over surrounding mountains [20, 35, 40]. Central Asia is therefore scientifically important for climate, air quality and cryosphere studies, yet it remains sparse in the vertically resolved observations needed to distinguish surface trapping from elevated transport. The problem is not that open archives are absent. Rather, the challenge is that radiosonde, sun-photometer, reanalysis and lidar products have rarely been joined with enough station-level scrutiny to support a defensible aerosol-regime interpretation.

The vertical structure matters because aerosol optical depth alone cannot identify the meteorological mechanism that produced a high-loading month. Fine-mode pollution can accumulate when stable layers suppress turbulent exchange, whereas coarse dust can dominate during mechanically mixed or transported conditions. Central Asia is a natural test bed for this contrast. Passive satellite studies show regional aerosol optical-depth seasonality; Dushanbe AERONET

and ground-lidar studies have characterised mixtures of dust, pollution and radiative effects over Tajikistan [14, 15, 33, 38]; and CALIOP has already been used to map the three-dimensional dust climatology and transport pathways of Central Asia [13, 22]. Our contribution is therefore not the regional dust vertical climatology itself, which these studies establish, but a specific and previously untested question: whether a station-screened lower-tropospheric stability index organises aerosol size and vertical structure into distinct regimes *after* seasonal confounding is removed, and whether those regimes correspond to distinct, model-corroborated dust source pathways. This combined stability–size–vertical–source–attribution test, anchored by an explicit seasonal-confounding null, has not to our knowledge been reported for the region.

Here we ask a narrower question that the observations can answer reproducibly: whether a station-screened lower-tropospheric stability climatology organizes Central Asian aerosol observations into defensible seasonal regimes after obvious seasonal-cycle confounding is removed, and whether the resulting regimes add vertical or surface-air-quality information beyond column AOD. To avoid station misidentification, every sounding file is checked against the official IGRA station inventory before analysis, and non-regional or pressure-level-incompatible records are excluded. We then use the nine screened regional station files, of which six provide valid 700–925 hPa stability months. We combine those soundings with AERONET spectral deconvolution at the available regional sites, ERA5 monthly pressure-level and single-level diagnostics, MERRA-2 aerosol extinction by species, CALIOP Level 3 aerosol typing, OpenAQ PM_{2.5} observations and bounded trajectory and residence diagnostics. The result is an observational baseline with an explicit null result: stability and fine-mode aerosol covary strongly as seasonal regimes, but the independent deseasonalized monthly relationship is weak.

2 Results

2.1 Station screening creates a defensible regional record

The official IGRA station inventory identified nine regional candidate stations within 35–50° N and 55–90° E. All nine period-of-record files were downloaded and parsed; six contributed valid strict 700–925 hPa stability months (Urumqi, Karaganda, Zhezkazgan, Almaty, Kyzylorda and Taraz, 176–581 months each), while Kashi, Hotan and Mashhad lacked the required pressure-level pair under the strict quality rule (high terrain and coverage constraints) and were catalogued but excluded. This station verification is central because it prevents non-regional or topographically incompatible records from entering the stability index.

The resulting regional series spans July 2010 through October 2025, with 184 monthly values and 53,017 contributing soundings across seasons. The stability metric is the regional monthly median potential-temperature difference between 700 and 925 hPa, after first aggregating by station to prevent high-frequency stations from dominating the result. This is not a direct boundary-layer-height measurement and it is not a stratospheric trend metric. It is a reproducible lower-tropospheric stability proxy that can be compared with gridded meteorological and aerosol products.

2.2 Seasonality explains the apparent stability–fine-mode relationship

The monthly station record resolves a strong seasonal contrast in lower-tropospheric stability and aerosol size partitioning (Table 1). DJF has a mean stability jump of 16.1 K, nearly three times the JJA mean of 5.8 K. Dushanbe AERONET fine-mode aerosol optical depth (AOD) is highest in DJF, while coarse-mode AOD is highest in JJA. This seasonality is consistent with winter stagnation favouring fine-mode pollution accumulation and warm-season mixing or transport favouring coarse dust, but it also means that raw correlations must not be interpreted as independent monthly coupling.

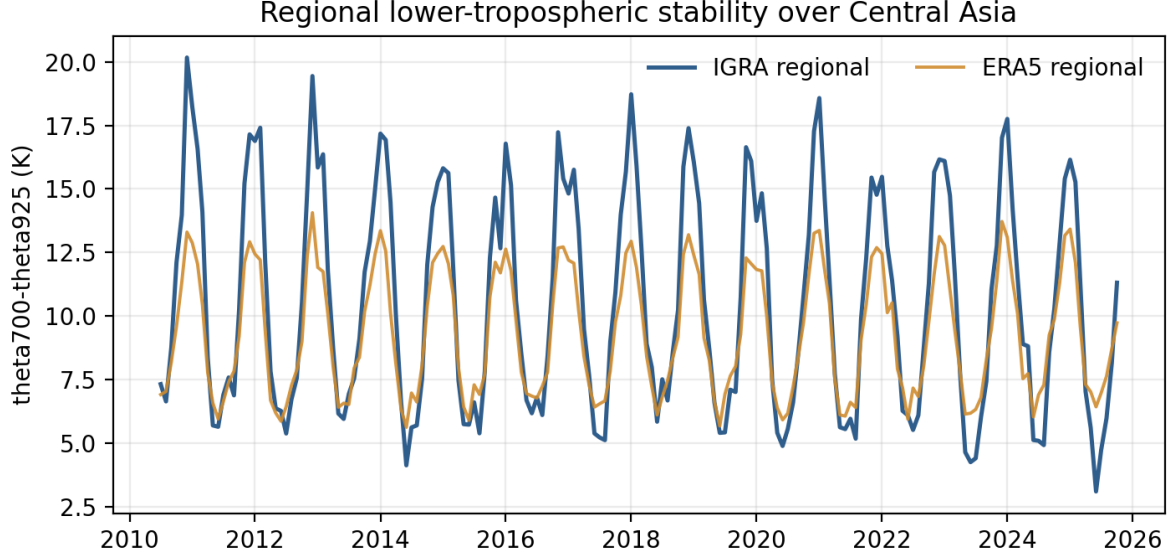


Figure 1: **Regional lower-tropospheric stability across Central Asia.** Monthly median 700–925 hPa potential-temperature difference from screened IGRA stations, with ERA5 shown as a gridded counterpart. The figure is regional, not a single-city station series.

Table 1: **Seasonal stability and AERONET aerosol-size statistics.** Stability is the regional monthly median 700–925 hPa potential-temperature difference. Fine- and coarse-mode AOD are Dushanbe AERONET 500 nm spectral-deconvolution products.

Season	Stability (K)	Fine AOD	Coarse AOD	Soundings
DJF	16.081	0.169	0.040	13,100
MAM	8.977	0.102	0.106	13,201
JJA	5.827	0.111	0.199	13,406
SON	11.304	0.151	0.145	13,310

The AERONET archive also bounds the regionality of the size result. Over the 94 months with at least two valid regional sites (Dushanbe, Issyk-Kul; Kashi has one valid month and is catalogued only), the two-site median fine fraction tracks Dushanbe closely ($r = 0.888$), so Dushanbe is not an isolated outlier. The two-site sensitivity reinforces the main caution: raw stability and fine fraction correlate strongly ($r = 0.789$) but the deseasonalized two-site anomaly relation is essentially zero ($r = -0.033$, AR(1)-adjusted $P = 0.753$).

Pairwise monthly correlations confirm the seasonal contrast but expose its limitation (Table 2). Stability correlates strongly with raw Dushanbe fine-mode fraction ($r = 0.810$, $n = 121$) and negatively with raw coarse-mode AOD ($r = -0.634$), but after removing the calendar-month climatology the fine-fraction relation reverses sign and is not significant ($r = -0.146$), and partial regressions controlling for month, precipitation, wind and boundary-layer height give no independent effect (partial $r = -0.140$ to -0.100). The robust conclusion is a seasonal aerosol-regime contrast, not a universal monthly predictor.

Table 2: **Seasonality-control tests for the stability–fine-fraction relationship.** AR(1)-adjusted significance uses an effective sample size from lag-1 autocorrelation. Partial regressions include calendar-month fixed effects and ERA5 meteorological controls.

Analysis	r or partial r	95% CI	Adjusted P
Raw regional stability vs fine fraction	0.810	[0.678, 0.892]	1.7×10^{-11}
Raw regional stability vs two-site fine fraction	0.789	[0.645, 0.879]	1.2×10^{-10}
Deseasonalized regional anomalies	-0.146	[-0.326, 0.045]	0.133
Deseasonalized two-site anomalies	-0.033	[-0.237, 0.173]	0.753
Partial: month, precipitation, wind	-0.140	[-0.322, 0.052]	0.151
Partial: month, precipitation, wind, BLH	-0.100	[-0.282, 0.089]	0.299
Deseasonalized Dushanbe ERA5 anomalies	0.105	[-0.077, 0.280]	0.256

Within-season tests show the same caution: stability is not a significant fine-fraction predictor inside DJF or JJA and is only moderately positive in MAM, with SON the one strong season-specific exception ($r = 0.796$, $n = 34$, AR(1)-adjusted $P = 4.1 \times 10^{-7}$). Lag tests (-3 to $+3$ months) show no coherent lead–lag sequence after deseasonalization, and the 850–925 and 700–850 hPa layer variants retain the raw seasonal correlation but weaken after deseasonalization (full within-season statistics in the Supplementary Information). These tests move the interpretation from a causal statement to a seasonal-regime statement.

2.3 Regime climatology and a secondary fixed-station stability decline

We next converted the multi-platform evidence into an explicit monthly aerosol-regime classifier. Across 184 months, the classifier identifies 81 coarse dust-transport months, 51 transition months, 32 mixed stagnation–dust months, 19 fine-mode stagnation months and one weak-background month. The seasonal structure is sharp: MAM and JJA contain 34 and 33 coarse dust-transport months, respectively, whereas DJF is dominated by mixed stagnation–dust (25 months) and fine-mode stagnation (11 months). SON is mainly transitional but includes eight fine-mode stagnation months, consistent with the strong SON within-season stability–fine-fraction relationship.

The screened 184-month record also supports a secondary change test. A station fixed-effects panel regression on all 957 station-months (station and calendar-month fixed effects) gives an IGRA 700–925 hPa stability decline of -0.78 K decade $^{-1}$, essentially the regional-median monthly value (-0.73 K decade $^{-1}$), and the sign and significance survive inference allowing for both serial and cross-sectional dependence in this small panel (wild-cluster bootstrap by station $P = 5 \times 10^{-4}$; HAC(12) 95% CI $[-1.13, -0.43]$). Five of the six stations decline, a common-station panel restricted to the three $\geq 95\%$ -coverage stations is stronger

($-1.09 \text{ K decade}^{-1}$), and the trend is stable across AR(1) adjustment, residual-block bootstrap, endpoint removal, Theil–Sen slope, the CALIOP- and AERONET-overlap periods and all leave-one-station-out aggregates, so it is a property of the fixed stations rather than of changing station composition (SI Tables S1, S3 and S15). Because these are raw IGRA temperatures from a single consistently processed archive without instrument- or radiation-correction homogenisation (RHARM/RAOBCORE-class), and the panel has only six stations, we report a consistent observational decline whose precise magnitude may retain uncorrected radiosonde discontinuities, not a homogenised regional climate trend. Coincident ERA5 boundary-layer-height ($+10.6 \text{ m decade}^{-1}$) and MERRA-2 dust-extinction ($+0.0135 \text{ decade}^{-1}$) increases have consistent HAC signs but weaker anomaly-level support and are treated as supporting context. Classifier-derived regime-frequency trends are partly conditioned on the evolving observing system (CALIOP ends June 2023; AERONET and OpenAQ coverage is site- and period-dependent) and are shown descriptively (SI Fig. S1), not advanced as independent change claims. Full trend, breakpoint, station-composition and product-period detail is in the Supplementary Information (SI Tables S1–S4, S15 and Supplementary Trend Detail).

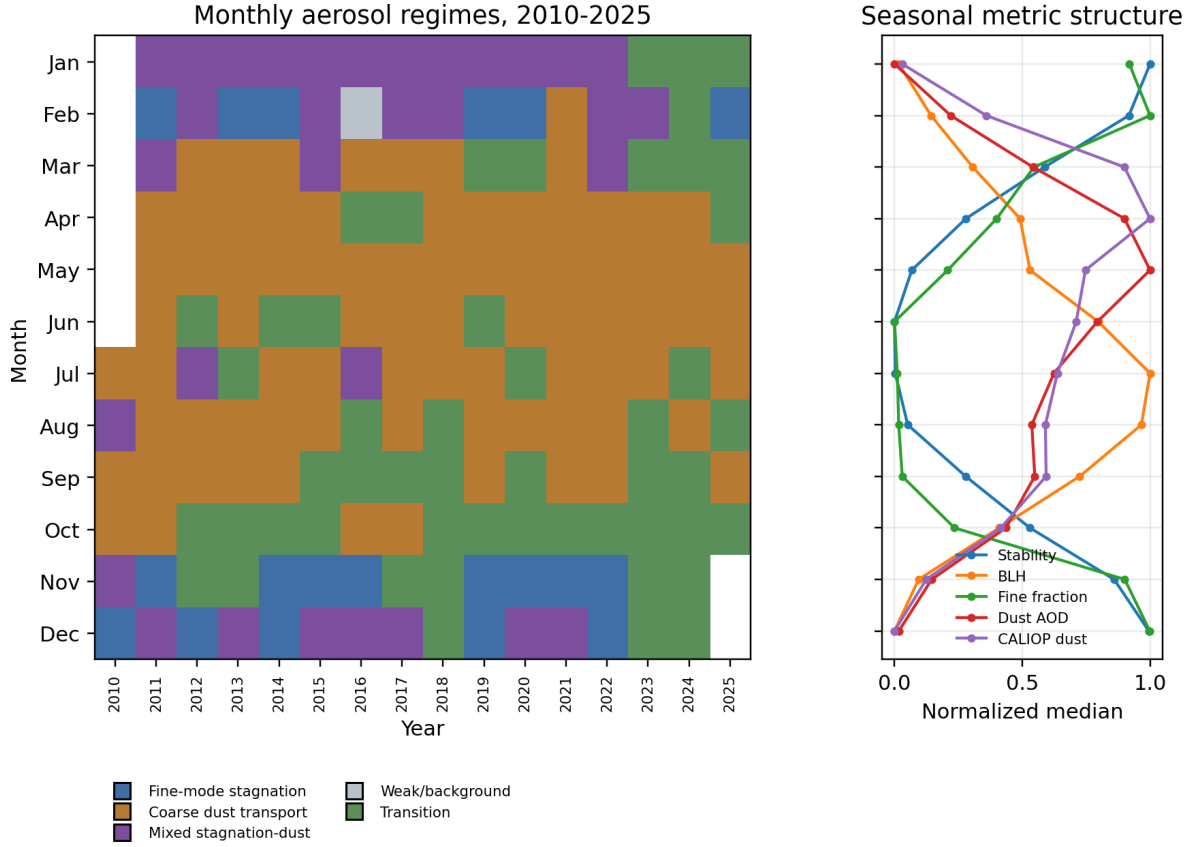


Figure 2: **Monthly aerosol-regime climatology.** The matrix classifies each month using the joined IGRA, ERA5, AERONET, MERRA-2 and CALIOP evidence chain. Side curves show normalized median seasonal structure for stability, Dushanbe boundary-layer height, AERONET fine fraction, MERRA-2 dust extinction and CALIOP dust optical depth.

2.4 Vertical lidar structure and PM prediction add mechanism tests

CALIOP vertical subtype occurrence provides the clearest mechanism check beyond column AOD. Because the monthly Level 3 product used here gives aerosol-subtype detection fractions in 0–2, 2–5, 5–10 and 10–15 km bins, we interpret the following as occurrence-weighted vertical

structure rather than extinction-weighted mass. Dust-plus-polluted-dust occurrence is more elevated during coarse dust-transport months than during mixed stagnation–dust months: mean occurrence-weighted layer height is 2.26 km (95% bootstrap CI 2.21–2.31 km, $n = 72$) for coarse dust transport, 1.80 km (1.70–1.90 km, $n = 32$) for mixed stagnation–dust and 1.69 km (1.60–1.79 km, $n = 17$) for fine-mode stagnation. The corresponding 0–2 km occurrence fraction is 0.59 for coarse dust transport but 0.73 for mixed stagnation–dust and 0.76 for fine-mode stagnation (SI Table S5). This is the vertical distinction that column AOD alone cannot provide.

Monthly-anomaly regressions show that this vertical placement is *not* skilfully predicted at the anomaly level by either aerosol loading or meteorology. For the two placement metrics that are defined independently of any predictor—dust-plus-polluted-dust center altitude and near-surface (0–2 km) fraction—AOD-only, meteorology-only and combined HAC regressions all explain essentially no month-to-month anomaly variance (adjusted $R^2 \leq 0.014$ in every case, $n = 154$). A third metric, the fraction of occurrence above the ERA5 Dushanbe boundary-layer height, at first appears highly predictable by meteorology (adjusted $R^2 = 0.726$), but this is a circularity rather than a mechanism: ERA5 boundary-layer height enters both the response (which is defined relative to it) and the predictor set, and because the lowest CALIOP bin spans 0–2 km a lower boundary layer mechanically raises the computed above-boundary-layer fraction. When boundary-layer height is dropped from the predictors, the meteorological skill for this same outcome collapses to adjusted $R^2 = 0.075$ —barely above the AOD-only value of 0.039. We therefore do not interpret the above-boundary-layer regression as evidence of meteorological control; the defensible, non-circular result is that monthly vertical-placement anomalies are weakly constrained by both aerosol loading and meteorology, consistent with the seasonal-regime rather than monthly-anomaly character of the signal.

The MERRA-2 3-D dust-profile extraction adds a first model vertical-placement check for the CALIOP overlap period, comparing normalized MERRA-2 0–15 km dust-burden fractions with normalized CALIOP dust-plus-polluted-dust occurrence (a profile-shape diagnostic, not an extinction-retrieval validation). Histogram-intersection overlap is high in every regime (0.90–0.93; SI Table S6), and the pooled matched-month value (0.915, $n = 154$) lies well above the trivial all-near-surface floor (0.652). A seasonality-preserving null is essential because both products share a strong seasonal cycle: permuting MERRA-2 only *within* each calendar month gives 0.910 (95% range 0.909–0.912), so although the matched overlap is statistically above it ($P < 10^{-3}$) the excess is only 0.005—almost all of the correspondence is the shared seasonal cycle, not month-specific skill. Indeed MERRA-2 matches individual CALIOP months (0.915) *less* well than a leave-one-year-out CALIOP calendar-month climatology does (0.962; SI Table S14). MERRA-2 therefore reproduces the seasonally-varying climatological dust-layer shape but shows no demonstrated month-specific skill beyond climatology, with only small regime-dependent center-height and near-surface biases (SI Table S6) that remain relevant for surface-exposure and forcing interpretation—a bounded model-evaluation result, not a month-resolved profile validation.

The same July 2019 GEOS-Chem Classic 14.7.1 case experiment (fullchem, $4^\circ \times 5^\circ$, MERRA-2 meteorology) adds an out-of-sample vertical check from a model whose dust fields come from an entirely different emission and optics chain than the assimilated MERRA-2 aerosol (though its transport shares MERRA-2 meteorology). For this CALIOP-classified coarse dust-transport month GEOS-Chem places the dust mass centre at 2.43 km, within 0.08 km of the CALIOP dust-plus-polluted-dust centre (2.35 km) and 0.51 km above the MERRA-2 3D centre (1.92 km), with a near-surface 0–2 km fraction (0.45) also closer to CALIOP (0.55) than to MERRA-2 (0.63). A model with independent aerosol physics therefore reproduces the elevated CALIOP placement; it underestimates the dust column (a known coarse-grid source-function bias), so we use it as a shape-and-placement check, not a burden validation (SI Fig. S2).

An unsupervised Gaussian-mixture model fit to calendar-month anomalies of the core meteorological and aerosol variables selects three components by BIC and independently supports the

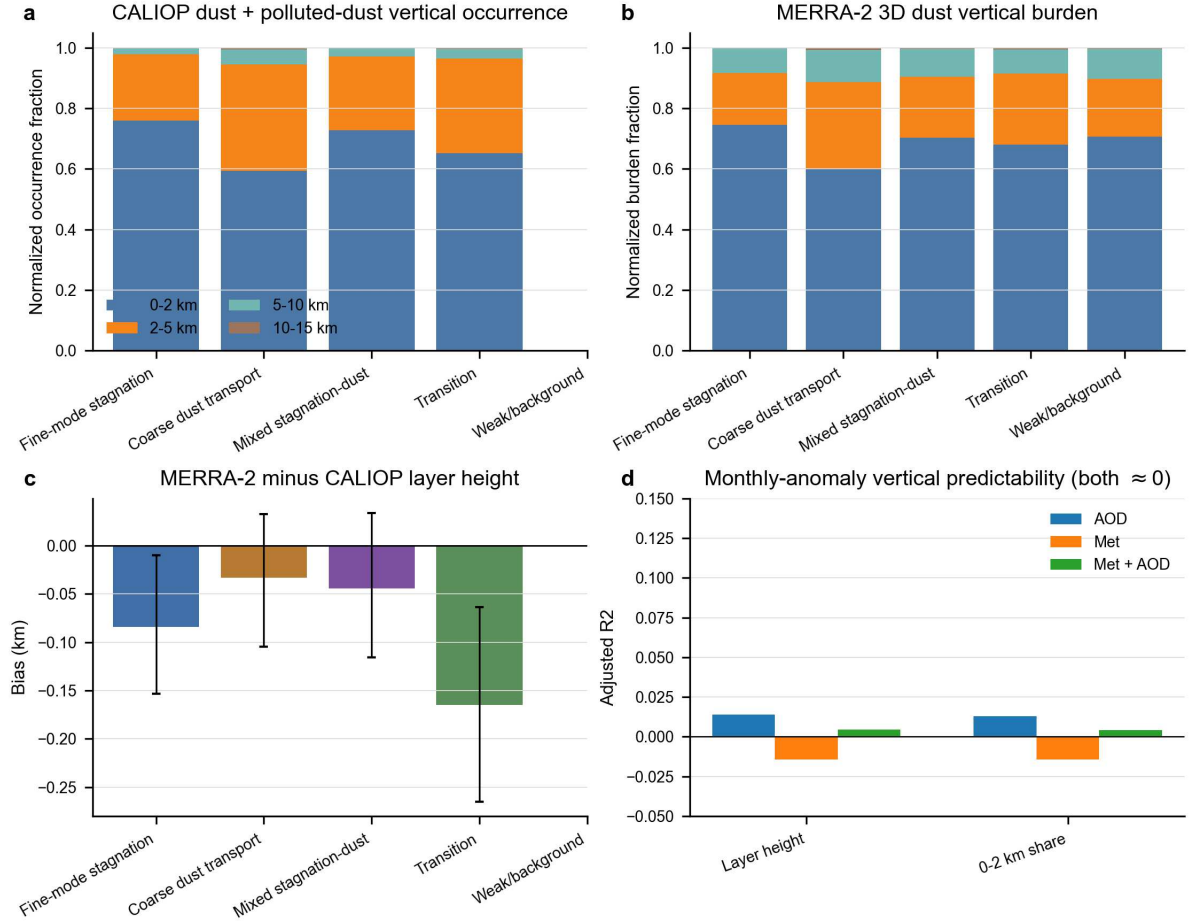


Figure 3: **CALIOP and MERRA-2 vertical aerosol structure by regime.** CALIOP dust plus polluted-dust subtype detections are normalized across 0–2, 2–5, 5–10 and 10–15 km bins. MERRA-2 3D dust burden is normalized over the same 0–15 km layer. Error bars are bootstrap 95% confidence intervals for regime means. The regression panel (d) shows the two predictor-independent monthly-anomaly outcomes (layer height, 0–2 km share), which are essentially unpredictable ($R^2 \approx 0$) by AOD or meteorology; the boundary-layer-relative outcome is a circular specification and is documented in SI Table S11, not shown here.

dust side of the hand classifier: 58 of 81 coarse dust-transport months fall in a single ventilated-dust anomaly component, while fine-mode and mixed regimes are more heterogeneous after deseasonalization. Full posteriors, transition statistics and BIC selection are reported in SI Table S9.

Surface PM_{2.5} observations from four cities give a bounded consequence test. To prevent outcome-normalization leakage the PM anomaly is recomputed strictly inside every cross-validation fold (year-blocked folds use training-year climatology; city-held-out folds, where a new city has no training climatology, predict absolute PM_{2.5}). Under this leakage-free protocol adding meteorology, CALIOP vertical metrics and regime variables to aerosol columns improves out-of-city absolute-PM RMSE by 6.8% (24.4 to 22.7 $\mu\text{g m}^{-3}$; observation- and city-block bootstrap 95% CIs both exclude zero) and gives a smaller, less certain year-blocked anomaly improvement of 3.3%; the monthly-mean 15 $\mu\text{g m}^{-3}$ severity-screen classification is not robust to city holdout (SI Tables S10, S13). This is a small, genuinely out-of-city improvement that rests on only four cities, so it remains a secondary result and is not an epidemiological attribution.

2.5 Reanalysis and satellite records test spatial coherence and aerosol type

The stability index is not an isolated radiosonde artefact, but ERA5 should be treated as a gridded meteorological counterpart rather than a fully independent validation because radiosondes and other observations enter the reanalysis system. ERA5 monthly pressure-level data closely match the IGRA 700–925 hPa stability series with $r = 0.976$ across 184 months (AR(1)-adjusted $P = 9.3 \times 10^{-24}$). ERA5 boundary-layer height is strongly anticorrelated with the IGRA stability index ($r = -0.942$, AR(1)-adjusted $P = 3.4 \times 10^{-16}$), supporting the shallow-mixing interpretation. The spatial mismatch with single-site Dushanbe AERONET is also reduced by local ERA5 checks: regional ERA5 stability correlates with Dushanbe-grid ERA5 stability ($r = 0.873$), and the IGRA regional index correlates with Dushanbe-grid ERA5 stability ($r = 0.880$). MERRA-2 aerosol speciation provides a separate aerosol check: the MERRA-2 fine fraction correlates with the raw stability index ($r = 0.679$, $n = 184$), and with the AERONET fine-mode fraction ($r = 0.614$, $n = 121$).

The clearest satellite aerosol-type validation is for dust. CALIOP V5.00 Level 3 aerosol typing was downloaded from January 2010 through June 2023, yielding 322 successfully processed day/night granules and 161 monthly regional records. MERRA-2 dust extinction correlates with CALIOP regional dust AOD ($r = 0.788$, $n = 154$), and AERONET coarse-mode AOD correlates with CALIOP dust AOD ($r = 0.419$, $n = 119$). By contrast, AERONET fine-mode AOD does not correlate significantly with CALIOP elevated-smoke AOD ($r = -0.070$, $n = 119$). We therefore interpret the coarse-mode regime as independently dust-linked, while treating the fine-mode regime as a stability-associated accumulation signal rather than a fully source-resolved combustion attribution.

The largest joint fine-mode month in the final rule set occurred in November 2019. The regional radiosonde stability jump reached 16.64 K, AERONET total AOD was 0.500, fine-mode AOD was 0.480 and the fine-mode fraction was 0.947. ERA5 placed the boundary layer at only 165 m, while MERRA-2 and CALIOP both indicated mixed aerosol rather than pure pollution: MERRA-2 dust and fine fractions were 0.530 and 0.464, and CALIOP dust plus polluted-dust AOD represented 0.109 of regional total AOD. A compact top-10 event ensemble generalizes this example under a deliberately non-circular design: the three event classes are ranked using *only* meteorology and column-aerosol variables (regional dust extinction and AOD, ERA5 boundary-layer height and 10 m wind, lower-tropospheric stability and AERONET fine-mode fraction), so CALIOP vertical structure, surface PM and trajectory residence are never used in selection and any class difference in them is an independent consequence. Classified this way, the classes still separate cleanly in the withheld vertical and surface metrics: elevated dust-transport months have a mean CALIOP dust-plus-polluted-dust center altitude of 2.34 km and 0–2 km fraction of 0.57, near-surface stagnation months a much lower center altitude of 1.63 km

and 0–2 km fraction of 0.78, and mixed dust-pollution months an intermediate 1.76 km (0.74). The between-class differences in these withheld outcomes are well resolved and not expected by chance: the elevated-minus-stagnation center-altitude difference is 0.71 km (month-resampling bootstrap 95% CI [0.54, 0.88]; permutation $P < 10^{-4}$ against random equal-size partitions of the 154-month candidate pool) and the near-surface-fraction difference is -0.21 ($[-0.26, -0.16]$, $P < 10^{-4}$; SI Tables S8 and S12). The surface-PM contrast points the same way: near-surface stagnation months carry a mean OpenAQ PM_{2.5} anomaly of $+26.3 \mu\text{g m}^{-3}$ against $-0.8 \mu\text{g m}^{-3}$ for elevated dust-transport months (the mixed subset is negative on average because many high-ranking months pre-date the city sensor coverage). Because the classifier never uses these vertical or PM outcomes, this is a strong independent withheld-outcome association: it supports—though, being observational, does not prove—that the vertical and surface-PM separation is a physical regime distinction rather than an artefact of the selection. The independence is also instrument-clean: when the classes are re-ranked using strictly non-CALIOP variables (MERRA-2 dust extinction and AERONET coarse-mode AOD replacing the CALIOP-typed dust AOD, with ERA5 meteorology and IGRA stability), the withheld CALIOP center-altitude separation is essentially unchanged (2.30 versus 1.59 km, permutation $P < 10^{-4}$; SI Table S18), so it does not arise from selection and validation drawing on the same CALIOP product family. A 2430-member HYSPLIT back-trajectory ensemble driven by NOAA GDAS1 meteorology (30 event months, three release heights, nine receptor perturbations and three back-durations) resolves these regimes into distinct Lagrangian source pathways. For the near-surface 500 m release air that carries dust to the receptor, the fraction of trajectory time spent below 1500 m over the arid dust-source basins rises monotonically across regimes—15.6% (95% block-bootstrap confidence interval 6.1–26.4) for near-surface stagnation months, 21.9% (10.6–34.9) for mixed dust-pollution months and 33.8% (30.3–37.3) for coarse dust-transport months—with the Karakum–Kyzylkum lowlands the dominant basin in every regime. The contrast is sharpest for the near-surface release and weakens with release height, as elevated westerly transport becomes less regime-diagnostic (Fig. 4). This ensemble back-trajectory residence and connectivity analysis corroborates the vertical and surface-air-quality regime separation with independent kinematic transport physics, a Lagrangian result that an independent Eulerian chemical-transport experiment confirms below.

A source-tagged extension of the same July 2019 GEOS-Chem case provides an independent Eulerian analysis that corroborates the transport geometry and then resolves the source emission. We first released four inert passive tracers—one from each of the same four arid dust-source boxes—with a uniform per-area surface flux; because the emission is uniform, the normalised near-surface arrival measures *transport connectivity* between each region and the receptor, assigning 64.7% to the Karakum–Kyzylkum lowlands and 32.6% to the Aral–Caspian corridor, with the Iran–Afghanistan plateau (2.0%) and Tarim/Taklamakan sector (0.8%) weakly connected. This connectivity is similar to the Lagrangian HYSPLIT near-surface residence fractions renormalised over the same four regions (Karakum–Kyzylkum 66.5%, Aral–Caspian 33.4%, Iran–Afghanistan 0.0%, Tarim/Taklamakan 0.1%)—within two percentage points in every region (Fig. 5)—so the two structurally independent transport formulations produce similar regional connectivity partitions despite their different sampling periods. We then ran a native-emission-weighted experiment in which each regional passive tracer is emitted with GEOS-Chem’s own spatially resolved dust emission flux rather than a uniform flux (the tracer is passive, so it carries the native emission pattern but omits dust size-dependent deposition). This materially changes the result: because dust is emitted from specific cells—most intensely a single coarse grid cell containing the Aralkum region—and those cells are better connected to Dushanbe than the region average, the native-emission-weighted arrival assigns 77% to the Aral–Caspian corridor and 23% to Karakum–Kyzylkum (SI Table S16). The robust conclusion common to both weightings is that the two arid basins immediately north-west of Dushanbe account for more than 99% of the native-emission-weighted tagged arrival from the four predefined source regions. A closure experiment—a fifth tracer emitting the model’s dust flux from the study domain (15–58°N,

HYSPLIT GDAS1 ensemble (2430 back-trajectories): Central Asian aerosol source pathways

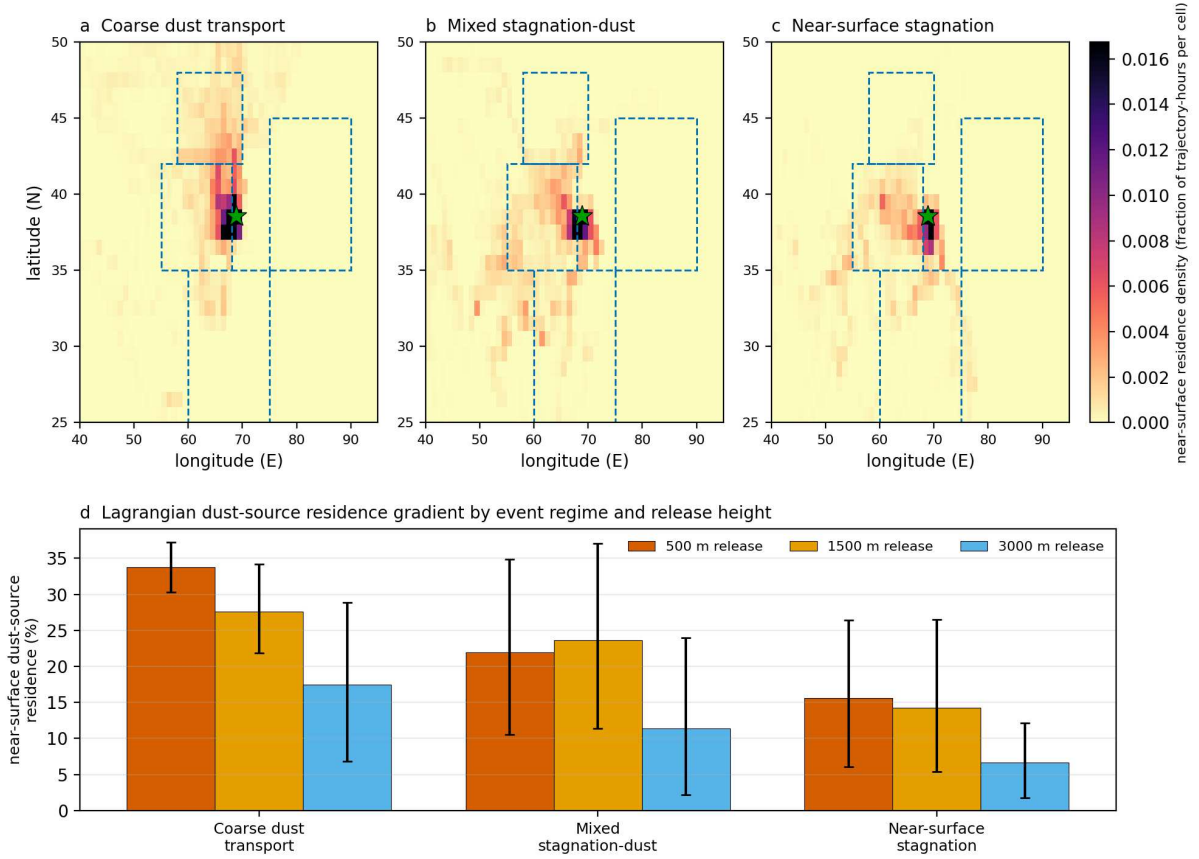


Figure 4: **HYSPLIT source-region residence and transport connectivity.** Panels (a)–(c) show the near-surface (< 1500 m AGL) Lagrangian residence-time density over Central Asia for each aerosol event type, computed from the 2430-member NOAA GDAS1 back-trajectory ensemble, with the dust-source basins (dashed) and the Dushanbe receptor (star) overlaid. Panel (d) shows the near-surface dust-source residence fraction by event regime and release height with 95% block-bootstrap confidence intervals, revealing a monotonic regime gradient that is sharpest for the near-surface 500 m release.

20–110°E) outside the four boxes—shows those four regions capture 95.5% of the near-surface tagged-tracer arrival at Dushanbe in this five-region system (only 4.5% from the rest of the domain), so the two north-western sectors account for roughly 95% of the native-emission-weighted passive-tracer arrival. This is a geographical closure of the passive tags rather than a closure against the full depositing dust field. Within the four regions the Iran–Afghanistan plateau and Tarim/Taklamakan sector are negligible—even though Iran–Afghanistan is a comparable *total* dust producer (26% of the four-region emission) that is simply too poorly connected to reach the receptor. Which of the two north-western basins leads depends on whether one weights by transport geometry (favouring Karakum–Kyzylkum, in agreement with HYSPLIT) or by the model’s actual emission field (favouring the Aral–Caspian corridor), and at $4^\circ \times 5^\circ$ resolution the cell-level split between them is not tightly resolved. A direct boundary-sensitivity test makes this explicit: the two 42°N grid cells that carry most of the Aral box’s emission straddle the Aral–Karakum boundary, and reassigning them from the Aral–Caspian box to Karakum–Kyzylkum swings the native-emission-weighted split from Aral 77%/Karakum 23% to Aral 2%/Karakum 98%—the leading basin flips entirely on the classification of two coarse cells. Crucially, the two-basin *combined* fraction remains above 99.9% and the Iran–Afghanistan and Tarim sectors remain negligible ($\leq 0.2\%$) under both assignments. We therefore advance the two-basin dominance as the robust result and treat the internal Aral-versus-Karakum split as unresolved at this resolution rather than as a determined attribution. That the emission weighting changes the leading basin is itself the key methodological point: transport connectivity alone, whether Lagrangian or Eulerian, does not determine source attribution without the actual emission field.

3 Discussion

This analysis presents a multi-platform observational study of Central Asian aerosol regimes anchored by station-screened radiosonde stability. The central result is deliberately narrower than the raw correlations alone might suggest. The large fine-fraction relationship is dominated by the seasonal cycle, and deseasonalized, within-season and partial-regression tests do not support a universal month-to-month stability control on aerosol size. What remains robust is still scientifically useful: screened radiosondes and ERA5 agree on the stability cycle, ERA5 boundary-layer height supports the stagnation interpretation, the regional stability index tracks Dushanbe-grid ERA5 stability, the two-site AERONET sensitivity supports the seasonal size-partitioning signal, CALIOP confirms that the coarse aerosol signal is dust-related, and CALIOP vertical occurrence shows that coarse transport and mixed stagnation–dust regimes differ in layer height.

The broader significance is that Central Asia appears to act as a transition corridor between aerosol states with different climate and air-quality implications. Stable winter months coincide with fine-mode aerosol and reduced dispersion, a pattern relevant for human exposure and valley air quality. Warmer, less stable periods shift the column toward elevated or transported coarse dust, a pattern relevant for radiation, cloud microphysics and cryosphere deposition [8, 20, 21, 35, 40]. Recent SBDART, COSMO–MUSCAT and MODIS–SNICAR studies already quantify regional dust forcing, Aralkum radiative-dynamical effects and Taklamakan snow darkening [2, 8, 40]; our contribution is instead the station-screened lower-tropospheric stability and aerosol-regime constraint that helps interpret when those dust and stagnation pathways appear in observations. The PM prediction tests give this constraint a practical check: vertical and regime information improves city-held-out absolute-PM prediction relative to aerosol columns alone, but classification of the monthly-mean $15 \mu\text{g m}^{-3}$ severity screen remains sensitive to city holdout and is not a health-impact model. The November 2019 monthly case illustrates how these regimes can overlap: fine-mode AERONET loading was extreme, but CALIOP and MERRA-2 both detected a substantial dust or polluted-dust component. This mixed structure is exactly why column AOD alone is insufficient for Central Asian aerosol attribution.

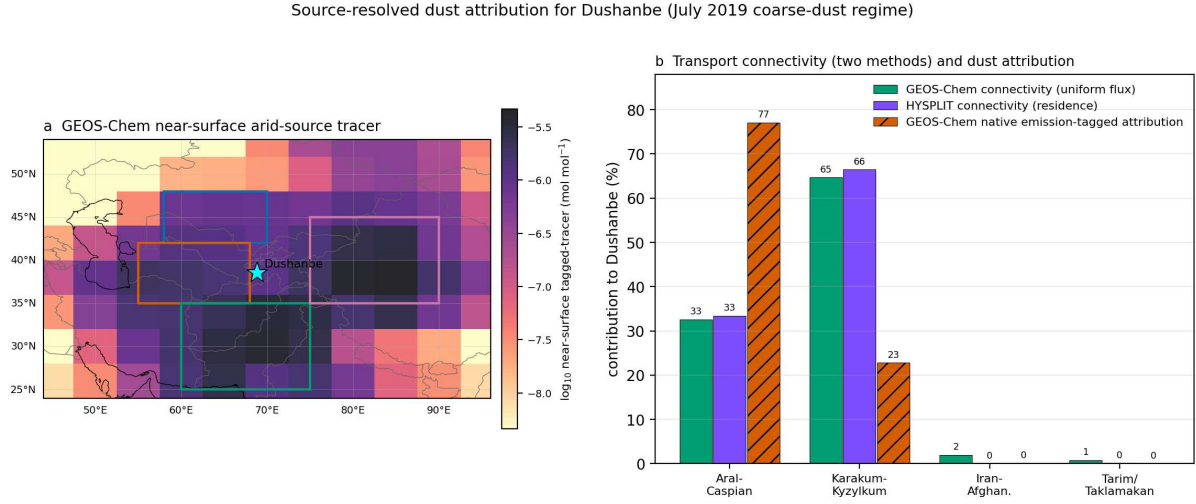


Figure 5: **Source-resolved GEOS-Chem tagged-tracer dust source partition versus HYSPLIT.** Panel (a) maps the near-surface (0–2 km) total tagged dust-source tracer concentration from the GEOS-Chem transport run, with the four source-region emission boxes outlined in colour and the Dushanbe receptor (star). Panel (b) shows, for each region, the transport connectivity to Dushanbe from the independent Eulerian (GEOS-Chem uniform-flux tracer) and Lagrangian (HYSPLIT 2430-member back-trajectory ensemble) methods—which produce similar regional connectivity partitions (within two percentage points, despite their different sampling periods) and favour Karakum–Kyzylkum—together with the native-emission-weighted passive-tracer arrival (each tracer emitted with GEOS-Chem’s own spatially resolved dust emission flux), which shifts the leading contribution to the Aral–Caspian sector because the coarse boundary cell containing the Aralkum region is better connected than the region average. Both weightings agree the two north-western arid sectors together account for more than 99% of the tagged arrival from the four predefined regions; the internal split is not resolved at $4^\circ \times 5^\circ$ —reassigning the two 42°N boundary cells that carry most of the Aral-box emission to Karakum–Kyzylkum swings the native-emission split from Aral 77%/Karakum 23% to 2%/98% while the two-basin total stays above 99.9% (SI Table S17).

The study also clarifies what is still required for full attribution. We include population-weighted PM_{2.5} attributable-mortality scenarios, MODIS–CERES radiative closure, a same-AOD vertical-profile heating-placement diagnostic (SI Table S7), an empirical snow-impurity forcing emulator and all-month ERA5 residence climatology, but these remain bounded observational products. They do not replace age-structured health-impact modelling, full aerosol radiative transfer, SNICAR-class snow radiative transfer, turbulence-resolved FLEXPART dispersion climatologies or emission-constrained chemical-transport climatologies. The HYSPLIT GDAS1 back-trajectory ensemble does, however, provide a quantitative single-particle residence and transport-connectivity constraint, and a GEOS-Chem Classic simulation with independent aerosol physics (though MERRA-2-driven meteorology) reproduces the CALIOP-observed elevated dust layer for the July 2019 coarse-dust month—placing dust higher than MERRA-2, as the lidar does—which corroborates the vertical-placement result with mechanistic aerosol physics rather than a second reanalysis diagnostic. Run as source-tagged passive-tracer experiments, the same Eulerian model produces a regional transport-connectivity partition similar to the Lagrangian HYSPLIT ensemble (both within two percentage points, though the two sample different periods), and a native-emission-weighted version identifies the two arid basins immediately north-west of Dushanbe—the Aral–Caspian corridor and Karakum–Kyzylkum lowlands—as the dominant tagged source regions (together more than 99% of the tagged arrival from the four predefined regions). Which of the two leads depends on whether one weights by transport geometry (favouring Karakum–Kyzylkum, in agreement with HYSPLIT) or by the model’s spatially resolved emission field (favouring the Aral–Caspian sector and the coarse cell containing the Aralkum region), a distinction only the native-emission-weighted run resolves and one that the coarse grid does not yet pin down. The CALIOP Level 3 aerosol product does not cover the full radiosonde and reanalysis period, and the vertical comparison uses CALIOP detection occurrence against MERRA-2 dust burden rather than matched aerosol-extinction profiles. These limitations are scientifically useful because they identify the next decisive additions: CALIOP curtain- or profile-level extinction constraints, full radiative-transfer experiments driven by observed aerosol profiles, age-structured city exposure surfaces, physically prescribed snow impurity optical properties, and ensemble dispersion or chemical-transport modelling driven by resolved emissions.

4 Methods

4.1 IGRA radiosonde screening and stability

Radiosonde profiles were obtained from the NOAA Integrated Global Radiosonde Archive version 2.2 [7]. The workflow first downloads the official station list and checks station identifier, latitude and longitude before processing profile data. Files are accepted only if the station metadata fall inside 35–50° N and 55–90° E. The six stations retained in the stability aggregate are CHM00051463 (Urumqi), KZM00035394 (Karaganda), KZM00035671 (Zhezkazgan), KZM00036872 (Almaty), KZM00038064 (Kyzylorda) and KZM00038341 (Taraz); CHM00051709 (Kashi), CHM00051828 (Hotan) and IRM00040745 (Mashhad) were downloaded but excluded from the strict 700–925 hPa metric. The final station inventory is written to `outputs/tables/ingest/igra_station_inventory.csv`.

Potential temperature was computed as

$$\theta = T \left(\frac{1000}{p} \right)^{0.2854}, \quad (1)$$

where T is temperature in kelvin and p is pressure in hPa. The primary stability metric is $\theta_{700} - \theta_{925}$ for soundings containing both pressure levels. Monthly medians and interquartile ranges were calculated per station, then aggregated regionally as a median across available stations. This two-stage aggregation gives each station equal monthly weight.

4.2 AERONET aerosol size partitioning

Aerosol size observations use NASA AERONET Version 3 Level 2.0 monthly AOD and spectral-deconvolution products [10, 17, 31]. We used 500 nm total AOD, 440–870 nm Angstrom exponent, and spectral-deconvolution fine- and coarse-mode AOD. Dushanbe remains the primary site in the joined monthly table because it is the closest long-running AERONET site to the radiosonde-domain interpretation and has 121 valid size-deconvolution months from July 2010 through December 2022. To test the single-site critique, we also extracted all paired local AOD/SDA files present in the downloaded AERONET monthly archives. After 2010, Issyk-Kul provides 148 valid Level 2.0 SDA months from January 2010 through October 2024, while Kashi provides only one valid month and is retained only in the coverage inventory. The regional AERONET sensitivity uses the monthly median across valid sites and masks correlations to months with at least two sites.

4.3 ERA5 and trajectory analysis

ERA5 monthly pressure-level and single-level fields were downloaded for 2010–2025 over 35–50° N, 55–90° E [12]. Regional quantities are area means; Dushanbe-grid values are nearest-grid extractions at 38.56° N, 68.86° E. We computed ERA5 analogues of the IGRA stability metric from 700 and 925 hPa temperatures (with 850–925 and 700–850 hPa sensitivity layers) and used boundary-layer height, surface pressure, 10 m and 700 hPa winds and precipitation as covariates, noting that ERA5 assimilates observations and is not a fully independent system. Kinematic ERA5 and NOAA READY/HYSPLIT [34] back trajectories for individual case months, and ERA5 monthly-mean ranked-target and full-period residence ensembles, are used only as screening and case-study checks (perturbation design and case configurations in the Supplementary Information). A separate top-10 event ensemble ranks elevated-dust, near-surface-stagnation and mixed months using *only* meteorology and column aerosol (regional MERRA-2 dust extinction and CALIOP-typed dust AOD, ERA5 boundary-layer height and 10 m wind, stability and AERONET fine-mode fraction), withholding CALIOP vertical structure, surface PM and residence so that class differences in those are independent consequences. The quantitative transport result uses a local HYSPLIT v5.4.2 [36] ensemble of 2430 back trajectories driven by NOAA GDAS1 (1°) meteorology, released at 500, 1500 and 3000 m AGL from nine receptor points within $\pm 0.5^\circ$ of Dushanbe for 120, 168 and 240 h back-durations over 30 event months (ten per class). Each hourly endpoint was assigned to a fixed Central Asian source sector; per-trajectory residence-time fractions (all altitudes and restricted to < 1500 m AGL over the arid basins) were averaged within event type with a whole-event-month block bootstrap. Trajectory residence is not a class-selection variable, so the residence gradient is non-circular and stable to reclassification (500 m fractions 33.8/21.9/15.6%). These remain single-particle kinematic calculations rather than turbulence-resolved FLEXPART dispersion or chemical-transport simulation.

4.4 MERRA-2 and CALIOP aerosol typing

MERRA-2 monthly aerosol diagnostics (M2TMNXAER, January 2010–October 2025) were regionally averaged over the same domain [5, 9, 32]: total, dust, sea-salt, black-carbon, organic-carbon and sulfate extinction, with fine-species extinction the sum of the carbonaceous and sulfate components. Deposition (M2TMNXADG) dry and wet bins were summed by species and converted to $\text{mg m}^{-2} \text{ month}^{-1}$. For the CALIOP-overlap period we also derived 72-layer 3-D dust burden (M2I3NVAER dust bins DU001–DU005 with layer pressure thickness and air density) integrated into 0–2, 2–5, 5–10 and 10–15 km bins.

CALIOP Level 3 tropospheric aerosol profile products were queried through NASA CMR and downloaded from January 2010 through June 2023 [19, 37, 39]. The product collection is CAL_LID_L3_Tropospheric_APro_CloudFree-Standard-V5-00. The workflow processed 322

day/night granules without failures and produced 161 monthly regional records. We extracted `AOD_Mean`, `AOD_Mean_Dust`, `AOD_Mean_Elevated_Smoke`, `AOD_Mean_Polluted_Dust`, and subtype accepted-sample counts. Aerosol subtypes were aggregated into total, dust, elevated-smoke and polluted-dust optical-depth metrics and into vertical detection fractions for 0–2, 2–5, 5–10 and 10–15 km layers. For the vertical-regime analysis, dust and polluted-dust detection fractions were summed by layer, normalized across the four bins, and used to compute occurrence-weighted center altitude, 0–2 km fraction, 2–5 km fraction, fraction above 5 km and fraction above ERA5 Dushanbe boundary-layer height. The above-BLH fraction accounts for partial bin overlap with the monthly BLH. These are detection-occurrence diagnostics, not CALIOP extinction-profile mass integrals.

4.5 Independent GEOS-Chem Classic dust simulation

To obtain an out-of-sample vertical check from a mechanistic chemical-transport model rather than another reanalysis diagnostic, we compiled and ran GEOS-Chem Classic 14.7.1 [3, 4] in its full-chemistry configuration at $4^\circ \times 5^\circ$ resolution on 72 hybrid-sigma levels, driven by MERRA-2 meteorology and HEMCO default emissions including the online mineral-dust source. The model was initialised from the standard July restart and integrated over the opening days of July 2019, the CALIOP-classified coarse dust-transport month, writing daily `SpeciesConc` (seven mineral-dust size bins `SpeciesConcVV_DSTbin1–DSTbin7`), `Aerosols` (`AODDust`) and `AerosolMass` ($\text{PM}_{2.5}$, PM_{10}) collections. The model was built against a stable HDF5 1.14/netCDF toolchain, and every HEMCO input file was content-verified before use. In post-processing, total dust volume mixing ratio was summed over the size bins and reduced to a cosine-latitude-weighted mean over the study domain ($35\text{--}50^\circ$ N, $55\text{--}90^\circ$ E). Mid-level pressures were reconstructed from the hybrid coefficients (`hyam+hybm`×surface pressure) and converted to geometric altitude with the US Standard Atmosphere; the dust mass-weighted centre of gravity and 0–2 km near-surface fraction were then computed and compared with the July 2019 CALIOP dust-plus-polluted-dust and MERRA-2 3D metrics. This is a short simulation on a coarse grid and is reported as an independent-physics shape-and-placement check, not a burden validation or climatology.

We additionally used the same GEOS-Chem Classic configuration for a source-resolved (tagged-tracer) dust-attribution experiment that is directly comparable to the HYSPLIT back-trajectory ensemble. We defined four inert passive tracers (`Is_Tracer`, no chemistry, deposition or sinks) and emitted each from one of the four arid dust-source region boxes used by the HYSPLIT source assignment—the Aral–Caspian corridor, the Karakum–Kyzylkum lowlands, the Iran–Afghanistan plateau and the Tarim/Taklamakan sector—with a uniform per-area surface flux applied through HEMCO region masks (grid cells assigned to at most one region by the same priority order as the trajectory analysis). The model was run in transport mode (advection, convection and non-local boundary-layer mixing active; gas-phase chemistry, photolysis and deposition disabled so the tracers are exactly conserved) for 1–5 July 2019, and the near-surface (0–2 km) and column tracer concentrations reaching the Dushanbe grid cell were normalised across the four regions to give each region’s fractional contribution. Because the emission is uniform per unit area, the normalised arrival is each region’s transport connectivity to the receptor (the source-to-receptor transport geometry), which we compare with the HYSPLIT near-surface residence fractions renormalised over the same four regions. To obtain a native-emission-weighted source estimate we then ran a second, native-emission-tagged experiment: we diagnosed GEOS-Chem’s own dust emission flux field (the summed `EmisDSTbin1–4` HEMCO diagnostic) and emitted each regional tracer with that spatially resolved emission pattern masked to its region, rather than a uniform flux. The near-surface arrival from this run is the model’s native-emission-weighted passive-tracer arrival, without any scalar-weighting approximation; because the tracer is passive it carries the native emission pattern but omits dust size-dependent deposition. A fifth tracer emitting the model’s dust from the study domain

outside the four boxes provides a closure check: the four source regions account for 95.5% of the total near-surface tagged-dust arrival at Dushanbe. Emission masks and attribution outputs are produced by the `scripts/impacts/make_geoschem_*` and `geoschem_*_attribution` scripts. Because GEOS-Chem covers early July 2019 whereas the HYSPLIT ensemble spans thirty coarse-, mixed- and stagnation-regime event months, the connectivity agreement is a comparison of stable regional transport geometry rather than a date-matched experiment.

4.6 Regime classification, latent regimes and surface PM

Monthly aerosol regimes were assigned from quantile-based flags in the joined table. Fine-mode stagnation requires high fine-mode fraction or MERRA-2 fine-species fraction together with high lower-tropospheric stability and either shallow ERA5 boundary-layer height or weak 10 m wind. Coarse dust transport requires high AERONET coarse AOD, MERRA-2 dust extinction, CALIOP dust AOD or CALIOP polluted-dust AOD. Mixed stagnation–dust is assigned when dust evidence coincides with fine-mode or stagnation evidence. Weak-background months require low total AOD in AERONET and MERRA-2 and no strong fine or dust signal; all remaining months are transition months. To test whether these hand-built regimes were reducible to arbitrary thresholds, we also fit Gaussian mixture models with two through seven components to calendar-month anomalies of IGRA stability, ERA5 Dushanbe boundary-layer height, ERA5 10 m wind, ERA5 precipitation, MERRA-2 total extinction, MERRA-2 dust extinction and MERRA-2 dust/fine fractions. Missing core values were median-imputed and standardized before model fitting; the component count was selected by BIC.

Surface PM_{2.5} observations were obtained from the OpenAQ API [28] for Dushanbe, Tashkent, Almaty and Bishkek. We used monthly `days/monthly` sensor aggregates for PM_{2.5} sensors with records overlapping the regime table, retained city-month values with at least 50% coverage, averaged multiple sensors within a city-month, and computed PM_{2.5} anomalies by subtracting each city’s calendar-month climatology. For exposure screening, positive anomalies were multiplied by days per month and summed by regime. As a severity screen we also flagged city-months whose monthly mean exceeds $15 \mu\text{g m}^{-3}$; this is the numerical value of the WHO 2021 24-hour guideline [29] applied to a monthly mean, and is therefore a severity level rather than a 24-hour-exceedance metric (which the monthly OpenAQ aggregates cannot resolve). The timescale-matched WHO 2021 annual guideline ($5 \mu\text{g m}^{-3}$) is exceeded in nearly every populated city-month here and so does not discriminate regimes. We then built three cross-validated PM models: aerosol-column predictors only, meteorology predictors only, and a combined model using aerosol columns, ERA5 meteorology, CALIOP vertical occurrence metrics, the hand regime and latent-regime posterior features. Ridge regression predicted fold-local PM_{2.5} anomalies under year-held-out validation and absolute PM_{2.5} under city-held-out validation; logistic regression predicted the monthly-mean $15 \mu\text{g m}^{-3}$ severity screen. Validation used leave-one-year-out and leave-one-city-out splits, with all preprocessing (median imputation, standardisation) fitted inside each training fold via a scikit-learn pipeline. To avoid outcome-normalization leakage, the regression anomaly target is also recomputed fold-locally: year-blocked folds subtract each city’s calendar-month climatology estimated from the training years only, and city-held-out folds—for which a new city has no training climatology—predict absolute PM_{2.5}, a fully out-of-city test. Paired observation-level and city-block bootstrap intervals are reported for the improvement over aerosol columns. The bounded population-weighted PM relative-risk and attributable-mortality screens (GHSL 2025 population in 10–50 km city buffers [16], World Bank crude death rates [30], and log-linear relative risks of 1.04–1.12 per $10 \mu\text{g m}^{-3}$ [6]), together with the MODIS pixel-scale snow-darkening and CERES/MODIS clear-sky closure screens, are locator diagnostics rather than attribution and are not age- or cause-structured; their essential assumptions are summarised in the Supplementary Methods and their full parameterisation is regenerated by the repository scripts in `scripts/impacts/`.

4.7 Impact-screen methods

The MODIS snow/albedo and pixel snow-darkening screens (MCD43A3 and MOD10A1/MYD10A1 via Google Earth Engine [11, 25–27]), the CERES EBAF clear-sky flux benchmark and first-order closure [24], the fixed-optics same-AOD heating-placement diagnostic, the empirical snow-impurity forcing emulator, and the population-weighted health screens are bounded locator diagnostics that do not enter the central inference. Their essential assumptions are given in the Supplementary Methods; full parameter choices and values are documented in and regenerated by the repository scripts in `scripts/impacts/`, with outputs in `outputs/tables/impacts/`. CERES is used only as an observed flux benchmark; none of these screens constitutes a radiative-transfer, snow-radiative-transfer or epidemiological attribution.

4.8 Statistics and reproducibility

Pearson correlations were computed pairwise (Fisher- z intervals), with autocorrelation-adjusted significance from a lag-1 effective sample size. To separate seasonal from monthly coupling we deseasonalized by calendar-month climatology and repeated correlations within season, ran partial regressions controlling for calendar month, precipitation, 10 m wind and boundary-layer height, and computed -3 to $+3$ month lags and 850–925/700–850 hPa layer variants. Change tests used monthly fixed-effect regressions (continuous) and binomial generalized linear models (occurrence) with HAC(12) standard errors; the full trend-robustness family (FDR correction, AR(1) adjustment, moving-block bootstrap, Theil-Sen/Kendall, endpoint removal, station-level and leave-one-station-out aggregates, product-period subsets) and classifier sensitivity (70/30 thresholds, leave-one-product-out, period subsets, MERRA-2-only proxy, agreement and Cohen’s κ) are tabulated in the Supplementary Information. Single-break dates maximize the two-sample standardized mean difference (minimum 36-month segments), with significance from 2,000 calendar-preserving 12-month block permutations. All reported values are regenerated by the scripts documented in `scripts/README.md`; the exact ordered commands are in the Supplementary Information.

5 Data Availability

The processed tables, figures and manuscript data products are contained in this repository and archived at Zenodo (<https://doi.org/10.5281/zenodo.17792122>). The principal joined monthly table is `outputs/tables/validation/reanalysis_validation_monthly.csv`; all trend, classifier, vertical-regime, overlap-null, event-ensemble, PM-prediction, station-panel, trajectory and GEOS-Chem tagged-tracer outputs are under `outputs/tables/`, with a complete machine-readable file manifest in `scripts/README.md`. The raw data sources are public archives from NOAA IGRA, NASA AERONET, ECMWF/CDS ERA5, NASA GES DISC MERRA-2, NASA ASDC CALIOP and CERES, NASA/USGS/NSIDC MODIS products hosted in Google Earth Engine, EC JRC GHSL, the World Bank Indicators API, NOAA READY/HYSPLIT, GEOS-Chem Classic driven by MERRA-2, and OpenAQ.

6 Code availability

The custom scripts used to generate the processed tables, figures and manuscript results are available in this repository and archived with the data products at <https://doi.org/10.5281/zenodo.17792122> and at <https://github.com/ProgrmerJack/Stratospheric-science-Analysis>. The ordered reproduction commands are provided in the Supplementary Information, and script provenance is documented in `scripts/README.md`.

Acknowledgements

We acknowledge NOAA NCEI for IGRA data provision, NASA AERONET for Dushanbe aerosol observations, ECMWF and the Copernicus Climate Data Store for ERA5, NASA GES DISC for MERRA-2, NASA ASDC for CALIOP and CERES data access, NASA LP DAAC/NSIDC and Google Earth Engine for MODIS access, the European Commission Joint Research Centre for GHSL population surfaces, the World Bank for crude death rates, the World Health Organization for air-quality guidelines, and NOAA Air Resources Laboratory for READY/HYSPLIT. This study received no funding.

Author contributions

A.A. conceived the study, curated the data, wrote the code, performed the analyses, produced the figures and wrote the manuscript.

Competing interests

The author declares no financial or non-financial competing interests.

References

- [1] Banks, J. R., Heinold, B. & Schepanski, K. Impacts of the desiccation of the Aral Sea on the Central Asian dust life-cycle. *J. Geophys. Res. Atmos.* **127**, e2022JD036618 (2022).
- [2] Banks, J. R., Heinold, B. & Schepanski, K. Dust aerosol from the Aralkum Desert influences the radiation budget and atmospheric dynamics of Central Asia. *Atmos. Chem. Phys.* **24**, 11451–11475 (2024).
- [3] Bey, I. et al. Global modeling of tropospheric chemistry with assimilated meteorology: model description and evaluation. *J. Geophys. Res. Atmos.* **106**, 23073–23095 (2001).
- [4] The International GEOS-Chem User Community. geoschem/GCClassic: GEOS-Chem Classic 14.7.1. *Zenodo* <https://doi.org/10.5281/zenodo.1343546> (2025).
- [5] Buchard, V. et al. The MERRA-2 aerosol reanalysis, 1980 onward. Part II: Evaluation and case studies. *J. Clim.* **30**, 6851–6872 (2017).
- [6] Chen, J. & Hoek, G. Long-term exposure to PM and all-cause and cause-specific mortality: a systematic review and meta-analysis. *Environ. Int.* **143**, 105974 (2020).
- [7] Durre, I., Yin, X., Vose, R. S., Applequist, S. & Arnfield, J. Enhancing the data coverage in the Integrated Global Radiosonde Archive. *J. Atmos. Oceanic Technol.* **35**, 1753–1770 (2018).
- [8] Gan, Y., Zhang, Z., Chu, W., Ding, J. & Ren, Y. Assessment and prediction of dust emissions, deposition and radiation forcing in Central Asia. *Atmos. Chem. Phys.* **26**, 3881–3900 (2026).
- [9] Gelaro, R. et al. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Clim.* **30**, 5419–5454 (2017).
- [10] Giles, D. M. et al. Advancements in the Aerosol Robotic Network (AERONET) Version 3 database: automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth measurements. *Atmos. Meas. Tech.* **12**, 169–209 (2019).

- [11] Gorelick, N. et al. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).
- [12] Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
- [13] Han, Y. et al. CALIOP-based quantification of Central Asian dust transport. *Remote Sens.* **14**, 1416 (2022).
- [14] Hofer, J. et al. Long-term profiling of mineral dust and pollution aerosol with multiwavelength polarization/Raman lidar at the Central Asian site of Dushanbe, Tajikistan: case studies. *Atmos. Chem. Phys.* **17**, 14559–14577 (2017).
- [15] Hofer, J. et al. Long-term profiling of aerosol light extinction, particle mass, cloud condensation nuclei, and ice-nucleating particle concentration over Dushanbe, Tajikistan, in Central Asia. *Atmos. Chem. Phys.* **20**, 4695–4711 (2020).
- [16] European Commission Joint Research Centre. GHSL: global population surfaces 1975–2030 (P2023A), GHS-POP. https://developers.google.com/earth-engine/datasets/catalog/JRC_GHSL_P2023A_GHS_POP (accessed 3 July 2026).
- [17] Holben, B. N. et al. AERONET—a federated instrument network and data archive for aerosol characterization. *Remote Sens. Environ.* **66**, 1–16 (1998).
- [18] Indoitu, R., Orlovsky, L. & Orlovsky, N. Dust storms in Central Asia: spatial and temporal variations. *J. Arid Environ.* **85**, 62–70 (2012).
- [19] Kim, M.-H. et al. The CALIPSO version 4 automated aerosol classification and lidar ratio selection algorithm. *Atmos. Meas. Tech.* **11**, 6107–6135 (2018).
- [20] Kok, J. F. et al. Mineral dust aerosol impacts on global climate and climate change. *Nat. Rev. Earth Environ.* **4**, 71–86 (2023).
- [21] Li, L. & Sokolik, I. N. The dust direct radiative impact and its sensitivity to the land surface state and key minerals in the WRF-Chem-DuMo model: A case study of dust storms in Central Asia. *J. Geophys. Res. Atmos.* **123**, 4564–4582 (2018).
- [22] Li, J. et al. Three-dimensional structure and transport properties of dust aerosols in Central Asia—new insights from CALIOP observations, 2007–2022. *Remote Sens.* **16**, 2049 (2024).
- [23] Nobakht, M., Shahgedanova, M. & White, K. New inventory of dust emission sources in Central Asia and northwestern China derived from MODIS imagery using dust enhancement technique. *J. Geophys. Res. Atmos.* **126**, e2020JD033382 (2021).
- [24] NASA/LARC/SD/ASDC. CERES Energy Balanced and Filled (EBAF) TOA and Surface monthly means data in NetCDF Edition 4.2.1. https://doi.org/10.5067/TERRA-AQUA-NOAA20/CERES/EBAF_L3B004.2.1 (accessed 1 July 2026).
- [25] NASA LP DAAC. MODIS/Terra+Aqua BRDF/Albedo Albedo Daily L3 Global 500m SIN Grid V061. <https://doi.org/10.5067/MODIS/MCD43A3.061> (accessed 1 July 2026).
- [26] NASA NSIDC DAAC. MODIS/Terra Snow Cover Daily L3 Global 500m SIN Grid V061. <https://doi.org/10.5067/MODIS/MOD10A1.061> (accessed 1 July 2026).
- [27] NASA NSIDC DAAC. MODIS/Aqua Snow Cover Daily L3 Global 500m SIN Grid V061. <https://doi.org/10.5067/MODIS/MYD10A1.061> (accessed 1 July 2026).

- [28] OpenAQ. OpenAQ API, PM_{2.5} sensor monthly aggregates. <https://api.openaq.org/> (accessed 1 July 2026).
- [29] World Health Organization. WHO global air quality guidelines: particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. <https://www.who.int/publications/i/item/9789240034228> (2021).
- [30] World Bank. Death rate, crude (per 1,000 people), indicator SP.DYN.CDRT.IN. <https://data.worldbank.org/indicator/SP.DYN.CDRT.IN> (accessed 6 July 2026).
- [31] O’Neill, N. T., Eck, T. F., Smirnov, A., Holben, B. N. & Thulasiraman, S. Spectral discrimination of coarse and fine mode optical depth. *J. Geophys. Res. Atmos.* **108**, 4559 (2003).
- [32] Randles, C. A. et al. The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation. *J. Clim.* **30**, 6823–6850 (2017).
- [33] Rupakheti, D. et al. Columnar aerosol properties and radiative effects over Dushanbe, Tajikistan in Central Asia. *Environ. Pollut.* **265**, 114872 (2020).
- [34] Rolph, G., Stein, A. & Stunder, B. Real-time Environmental Applications and Display sYstem: READY. *Environ. Model. Softw.* **95**, 210–228 (2017).
- [35] Skiles, S. M., Flanner, M., Cook, J. M., Dumont, M. & Painter, T. H. Radiative forcing by light-absorbing particles in snow. *Nat. Clim. Change* **8**, 964–971 (2018).
- [36] Stein, A. F. et al. NOAA’s HYSPLIT atmospheric transport and dispersion modeling system. *Bull. Am. Meteorol. Soc.* **96**, 2059–2077 (2015).
- [37] Tackett, J. L. et al. CALIPSO lidar level 3 aerosol profile product: version 3 algorithm design. *Atmos. Meas. Tech.* **11**, 4129–4152 (2018).
- [38] Wang, D., Zhang, F., Yang, S., Xia, N. & Ariken, M. Exploring the spatial-temporal characteristics of the aerosol optical depth (AOD) in Central Asia based on the moderate resolution imaging spectroradiometer (MODIS). *Environ. Monit. Assess.* **192**, 383 (2020).
- [39] Winker, D. M. et al. The CALIPSO mission: A global 3D view of aerosols and clouds. *Bull. Am. Meteorol. Soc.* **91**, 1211–1229 (2010).
- [40] Xing, Y. et al. Dust storms from the Taklamakan Desert significantly darken snow surface on surrounding mountains. *Atmos. Chem. Phys.* **24**, 5199–5219 (2024).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementaryinformation.pdf](#)