PREPRINT Not peer reviewed. Posted on air.nonarkara.org, 4 November 2026. Version 1

Nobody owns the haze

An open, end-to-end warning system for Thailand, and its first 110 days

Pun Thiengburanathum[1,2] **Non Arkaraprasertkul**[3,*]

[1] Faculty of Engineering, Chiang Mai University, Thailand [2] Research and Innovation Acceleration Agency for Competitiveness and Area Development (RCAD), Thailand
[3] Digital Economy Promotion Agency of Thailand (depa), Bangkok [*] Correspondence: via air.nonarkara.org

ABSTRACT

Smoke has no passport, and neither does the failure to manage it. In Thailand the agencies that measure the air, forecast it, map the fires and warn the public each hold one piece, and the regional agreement that should bind the neighbours binds states but not plumes. We asked whether a small team could close three gaps (coordination between agencies, data-driven response, and a chain that runs from the burning field to the phone in a pocket) without waiting for any institution to change. We built AirDash, an open system that reads 16 public data feeds from more than ten organisations, joins them, and pushes a warning to a phone. Over the 110 days before the 2026–27 burning season it made 58,298 scheduled pulls, 98.7% of them logged as successful (two feeds that need a free token skipped quietly all season), a ground $PM_{2.5}$ reading reached the system a median of 28 minutes after the hour it described, and an alert was handed to Telegram a median of 7 seconds after it was raised. The same record audits the data Thailand already has. 45% of the national $PM_{2.5}$ stations stand in Bangkok and three provinces have none; two official sources put a province in different air-quality bands on 23% of province-days, nearly always between the two cleanest bands; 86% of the fire cell-days on the mainland lay outside Thailand. Our own system also fell short. Only 3 of 13 forecast warnings reached the Thai 24-hour standard at any station, carrying the last observation forward beat our bias-corrected CAMS forecast, and cameras and airport visibility added little. We publish the failures, the code, and a pre-specified test plan (Table 3) for the season that begins on 1 December.

Keywords: haze; $PM_{2.5}$; Thailand; early warning; transboundary pollution; open data; biomass burning; forecast verification

1 Introduction

Every dry season the north of Thailand disappears under smoke. The cost is counted in clinics. In the upper north, same-day PM_{10} was associated with more outpatient visits for chronic lower respiratory disease (incidence rate ratio 1.020, 95% CI 1.012–1.028) [1]; all-cause mortality rose 0.9% for every 10 $\mu g\,m^{-3}$ of vegetation-fire $PM_{2.5}$ [2]; and in haze periods the fatality and re-admission rates for COPD rose significantly [3]. Globally, smoke from landscape fires is estimated to kill about 339,000 people a year, 110,000 of them in Southeast Asia [4].

The smoke is also not only Thai. A source-apportionment simulation of 2019 for the upper north puts Myanmar as the main contributor in the west and centre and domestic burning in the east at the peak of the season [5]. Burma alone accounted for 37% of the hotspots in Southeast Asia in March in one analysis [6]. Across Southeast Asia, an estimated 23% of the roughly 900 thousand premature deaths a year from air pollution come from transboundary air, with impacts especially adverse in border regions [7]. Origin and destination are different places, run by different governments.

Southeast Asia has a treaty for this. The ASEAN Agreement on Transboundary Haze Pollution was signed in 2002 [8]; Indonesia, the last of the ten members, deposited its instrument in January 2015 [9, 10]. Scholars describe it as soft by design, a non-binding agreement built on sovereignty and non-intervention [11, 12], and two ASEAN roadmaps have followed, for 2016–2020 and for 2023–2030 [13, 14]. Inside Thailand, policy reviews describe many agencies with different mandates and little integration across sectors [15], and a structural problem of silo culture [16]; one analysis argues that fragmentation and protected economic interests have compromised the transboundary governance of the air [17]. We do not try to settle the diplomacy here. We take it as given that neither the treaty nor the agencies will change before the next burning season, and ask what can be built in the meantime.

We name three gaps. The first is *coordination*. The data that matter are public, but they sit in separate systems: ground stations at one agency, a satellite–ground fusion at another, burn scars at a research institute, weather at a third, and fires at NASA. The second is *responsiveness*. A number on a dashboard is not a response. Alerts can change behaviour but they also fatigue, the evidence on health outcomes is mixed, and localised, actionable information is likely to help [18–20]. The third is the *end-to-end chain*. Nobody owns the path from the field where the fire starts, through the air that carries it, to the person it reaches.

Research question. Can a small team, bypassing bureaucracy, build one system that closes these three gaps and that people can use right away; and what does its own first season of data say about whether the gaps are real?

Contributions. (i) An open design and implementation of the chain (Section 3). (ii) A quantitative audit of the integrated record: 110 days, 175 ground stations and 58,298 pulls (Section 5). (iii) The system's failures, reported with the same weight as its successes. (iv) A set of tests, fixed before the burning season, that will decide which parts of the system deserve to stay (Section 6).

Scope. We have not obtained a multi-year, machine-readable ground record. The ground record here begins on 17 July 2026 and ends on 3 November 2026, before the dry, burning season, which we take as 1 December to 30 April. The multi-year argument rests on the literature and on three seasons of burn-scar maps (2024–2026). What follows is therefore a systems paper with an early-season audit, not a trend analysis. Where the data cannot carry a claim, we say so.

2 Background

Where the smoke comes from. Fire aerosols explain nearly all very-low-visibility events in Southeast Asia, and burning on the mainland accounts for 99% of fire-produced $PM_{2.5}$ in Bangkok [21]. The region's aerosol system is among the most complex in the world and its satellite aerosol products diverge widely [22]; field campaigns document the biomass-burning aerosol near its sources and far downwind [23]. El Niño strongly modulates burning in the Maritime Continent [24, 25], but Cambodia, Myanmar and Thailand respond far less: hotspot severity rose by under 25% in El Niño, positive-dipole or combined years, against more than 72% in Kalimantan in the same comparison [26]. One global chemistry model underestimated the aerosol burden, most strongly early in the burning season [27].

What can be measured. Regulatory stations remain the reference. Low-cost optical sensors fill the gaps but read high without calibration, by about 40% in one large US evaluation [28], and their fitness for purpose depends on the calibration and on the aerosol [29–31]. More than 1,500 such sensors are installed in Thailand [32]; DustBoy, run from Chiang Mai University, is one such network [33]. From space, VIIRS detects active fires at 375 m [34] through NASA's FIRMS service [35, 36], and AERONET gives the column aerosol optical depth and the Ångström exponent, a qualitative indicator of particle size [37–40]. Thai agencies publish a fire layer [41] and a burn-scar monitoring site [42], and academic work maps burned area from Sentinel-2 imagery [43].

What can be forecast. CAMS, the Copernicus atmosphere service, is the free workhorse; its reanalysis assimilates satellite aerosol retrievals [44]. Its $PM_{2.5}$ forecast skill is modest and uneven: weaker over Malaysia than elsewhere [45], strongly lead-dependent in China [46], and biased in different directions in different regions [47]. Model output statistics, regressing the model onto local observations, is the classical remedy [48]. Trajectory models such as HYSPLIT trace where an air parcel came from [49].

What the eye and the camera can see. Horizontal visibility follows the Koschmieder relation between extinction and range [50], and visibility-derived extinction tracks $PM_{2.5}$ well in polluted regions but poorly in the cleanest [51]. Haze can be estimated from a single image [52], and camera-derived extinction has correlated with $PM_{2.5}$ at $R^2 = 0.7$ over a limited concentration range [53]. A hybrid CNN–LSTM on Bangkok photographs improved on plain CNNs [54], and a model that joins surveillance images to numerical data forecast $PM_{2.5}$ in Shanghai well [55]. In Chiang Mai, however, trained staff estimating the visual range reached 79% accuracy for PM_{10} but only 26% for $PM_{2.5}$ [56].

Warning people. Warnings change what people do, and they fatigue. A Sydney study found alerts cut cycling by 14–35% and also found evidence of alert fatigue [18]; English data show avoidance behaviour for children's asthma [57]; a Cochrane review found the evidence on health-care use conflicting [19]; and a review of 37 communication studies found actionable, localised information likely to help [20]. A review of citizen-science monitoring found little research on how its data feed policy [58].

3 The system

3.1 Design rules We wrote five rules and have tried to keep them; Table 2 and Section 5 show where we did not. **Public before permitted.** Use only what agencies already publish, and ask nobody. No memorandum of understanding, no data-sharing agreement, no meeting. Where a stream needs a free token, the system is meant to say so rather than fail quietly. **Provenance on every number.** Each value carries its source, its observation time and its age; a number without a time is not shown. **Corroborate, never substitute.** Cameras, airport visibility and satellite aerosol may raise a flag or add a reason to the verdict; they never produce a $\mu g\,m^{-3}$ figure. **Fail closed, and say so.** Inbound webhooks are refused until their secrets are configured. "Unknown" is never reported as "fine". **One laptop.** The server is Node.js with 0 package dependencies and one SQLite file; anyone can run it. The code is 30,009 lines of server, 29,372 of browser JavaScript and 10,064 of tests in 97 files, with 85 API routes, over 349 commits since 2026-07-16, written with AI coding assistants under the direction of one author.

3.2 The chain Figure 1 shows the path from a fire to a phone. Table 1 lists what the system reads.

Ground and fusion. The Pollution Control Department (PCD) publishes hourly readings from 175 $PM_{2.5}$ stations; the Geo-Informatics and Space Technology Development Agency (GISTDA) publishes a satellite–ground fusion for all 77 provinces. Both are polled and stored with their own observation times.

Forecast and its correction. CAMS $PM_{2.5}$ forecasts arrive through Open-Meteo for each province. They read low against Thai ground stations, so the system applies a province-specific ratio learned over the trailing 14 days, clamped to 0.5–3, and falls back to the national ratio when a province has fewer than five paired days. This is a simple form of model output statistics [48]. A ratio learned on clean days overshoots on dirty ones, so it applies only up to the mean CAMS level it was learned on; beyond that, the calibrated absolute offset is added instead.

Where the smoke is coming from. VIIRS detections from three satellites (Suomi-NPP, NOAA-20, NOAA-21) are gridded to 0.25° cells. For each province the system sums the fire radiative power of cells within 500 km and within ±45° of the forecast wind-from direction, weighted by $1/(1 + d/150)$ for distance d in km, and grades the result low, moderate or high at 20, 150 and 600 MW. This is a transport indicator, not a concentration forecast: it says where smoke can come from, not how much will arrive.

The verdict. A province's Air Watch Score is 0.40 ground $PM_{2.5}$ + 0.10 other pollutants + 0.15 trend + 0.20 forecast + 0.15 stagnation, anchored to the Thai air quality index breakpoints, whose 24-hour $PM_{2.5}$ values are 15, 25, 37.5 and 75 $\mu g\,m^{-3}$ [59]. Thailand's 24-hour standard has been 37.5 $\mu g\,m^{-3}$ since 1 June 2023 [60], which equals the WHO interim target 3; the WHO guideline is 15 [61].

Corroboration. Public cameras (about 1,000 in the catalogue, drawn from GISTDA, iTIC, the Department of Water Resources and a volunteer aggregator) are judged only against their own clear-day baseline at a similar hour. A province whose cameras look hazier than usual gets a reason added to its verdict; if fog is likelier (relative humidity ≥ 90% or an airport report of mist) the reason says so. Airport visibility from METAR reports is read through the Koschmieder relation [50]. In exactly one case do cameras move a band: no ground sensor is reporting and haze is widespread without fog, which raises the province to at least WATCH.

The push. Alerts are raised by rule (level crossed, forecast above the line, all-clear, rain washout) and fanned out to Telegram and LINE subscribers. Telegram is the main channel because the free LINE tier is capped at 300 messages a month, which one haze episode would exhaust. A gap of 30 minutes between broadcasts and of three hours per subscriber keeps a volatile hour from becoming a flood. Both webhooks verify who is speaking.

3.3 Built to survive a season The system runs on one laptop behind a named tunnel and a static front end. Table 2 lists what broke between July and October and what was changed. Each row is a rule the system now enforces rather than a lesson learned.

4 Data and methods

Window. All analyses are frozen at 00:00 on 4 November 2026 (Bangkok time). Ground and reliability records cover 17 July to 3 November (110 days, 2,640 hours). Fire detections cover 21 October to 3 November (14 days); camera frames 28 September to 3 November; airport reports 11 October to 3 November; the alert and Telegram log 28 September to 3 November; the PCD reading lag from 7 August; burn scars for 16 months between January 2024 and April 2026 (62 provinces); forecast pairs 7 August to 3 November.

Ground record. Hourly $PM_{2.5}$ values come from the PCD feed. A station-day needs at least 18 of 24 hourly values; a province-day is the mean of its station-days. Values of 500 µg m^{-3} or more are treated as instrument faults (11 values at 2 stations; the raw maximum was 950). An exceedance of 37.5 is called *isolated* when at least two stations within 50 km reported that day and their median was under half its value, the signature of a local source or a fault. "Upper north" means the eight provinces of Chiang Mai, Chiang Rai, Lampang, Lamphun, Mae Hong Son, Nan, Phayao and Phrae.

Intervals. Where the unit of analysis repeats (frames per camera, hours per airport, province-days per day) we use a 95% percentile interval from a bootstrap that resamples whole clusters [62], with 400 to 2,000 resamples depending on cost. Simple proportions over station-days, detections, alerts and warned province-days carry Wilson 95% intervals [63, 64] that treat the observations as independent; these are optimistic where events cluster in time, and we say so where it matters. Rank correlations are Spearman's ρ [65]; agreement between category assignments is Cohen's κ.

Forecast verification. For lead k days, the target is the province's mean station-day value on day D and the forecast is the first one issued on local day $D - k$. The bias correction is replayed *out of sample*: the ratio for a forecast issued on day I uses observations only through $I - 1$, with the production rule above, applied to station-days with at least 18 hourly values. The baseline is persistence, the last observed daily province mean available at issue time. Events are scored by probability of detection (POD), false-alarm ratio (FAR) and critical success index (CSI) [66, 67].

Smoke transport replay. The upwind indicator is re-implemented from the production code (500 km radius, ±45° sector, distance weights, 0.25° cells, a two-day look-back, cells within 15 km counted whatever the wind) and restricted to the box the production code uses, 4–29°N and 92–110°E. It is evaluated with the wind forecast issued on the morning of each day (today's, where production prefers tomorrow's). Each province's next-day $PM_{2.5}$ and log upwind fire power are centred on the province's own mean, to remove geography, and then on the day's mean across provinces, to remove everything that changes for all provinces together, such as the fortnight-long fall in fire counts. Provinces share the weather on a given day, so the interval resamples days, of which only 9 have a next-day record.

Reproducibility. Every number in the text is filled from a JSON file written by the analysis scripts at build time; a number that does not exist stops the build. The extraction step reads the live database and the long-term archive read-only; those are not released. The extracts it writes (station-day means, forecast pairs, the ONI file) are in paper/data/inputs, and every other analysis and figure rebuilds from them. Data are public feeds; no personal data were analysed, and subscriber records were not used.

5 Results

5.1 The network we inherited The Pollution Control Department reports $PM_{2.5}$ from 175 stations in 74 of 77 provinces (Fig. 2a). Bangkok holds 79 of them, 45% of the network. Three provinces (Phetchaburi, Phang Nga and Ranong) have none, and 60 have exactly one. The eight upper-north provinces, where the dust season does its worst, share 14 stations. Only 67 stations also report PM_{10}, in 36 provinces, so the question "is this smoke or dust?" can be asked of the ground record in fewer than half the country. Of the 161 stations that produced at least one full station-day, 132 reported at least 90% of their hours, and 17 reported fewer than half.

The record itself is quiet (Fig. 3). Across 14,357 station-days, 38% exceeded the WHO 24-hour guideline of 15 µg m^{-3}, but only 31 (0.22%; 95% CI 0.15%–0.31%) exceeded the Thai standard of 37.5, and 2 exceeded 75. The monthly mean of station-days was 14.2 in July (from the 17th), 11.4 in August, 13.0 in September and 17.3 µg m^{-3} in October, with a station mix that changes from month to month. The worst day was 9 October, when the highest station in each of 6 provinces crossed 37.5. Of the 31 exceeding station-days, 12 were isolated spikes, 13 were not isolated, and 6 had too few neighbours to tell. The largest value in the record, 151.8 µg m^{-3} at one station in Chon Buri on 7 September, stood alone: the median of its 7 neighbours within 50 km was 12.9. That station supplied 8 of the 11 fault values, and the highest hourly value that survived the fault filter was 461. We read this as a baseline. It is the clean air against which the burning season will be judged, and it is not evidence that the north is safe.

5.2 Does the machinery hold? The collector made 58,298 scheduled pulls across 23 sources, and 747 failed (1.3%). It was up 94.6% of the window. Seven gaps longer than six hours add up to 122 hours (Fig. 4a). Two match entries in Table 2: 19.8 hours on 26–27 September (the database incident) and 30.9 hours on 24–25 October (the disk failure). The other five, the longest 34.3 hours on 23–25 August, have no entry in our operating record and we cannot explain them. The newest successful PCD pull was no older than two hours 95.3% of the time; the feed was stale for 124 hours in 19 episodes.

Three clocks matter to a person waiting for a warning (Fig. 4b–d). A PCD reading was first seen by the system a median of 28 minutes after the hour on its stamp (95th percentile 67; 99.9% within two hours). An alert was handed to Telegram a median of 7 seconds after it was raised, 85% within five minutes and all 485 within 30. This covers the 485 severe alerts logged from 28 September, each matched to the next broadcast line in the server log; it measures the server's send to Telegram, not arrival on a device, to one or two chats, and the tail is bounded by the 30-minute gap between broadcasts. Combining the two hops by random pairing gives a median of 32 minutes (95th percentile 71) from the hour on a reading's stamp to the Telegram hand-off; most readings raise no alert at all. The slow hop is the fire: a VIIRS detection was first seen a median of 392 minutes (6.5 hours) after the overpass, and a tenth took over 23.6 hours. Much of that is our own six-hour polling of a rolling file (the fifth percentile is 172 minutes), so shortening the interval is the first fix. Until then the warning should say how old the fire picture is.

Success is not delivery. Five sources wrote no new readings in 110 days. Three of them (the weather relay, camera health and the flood sister system) relay into a key–value store or other tables by design. The other two, satellite rain (IMERG; 5,077

Table 1. What the system reads. Seventeen streams, their owners, interface and measured behaviour over the 110-day window. "Interval" is the configured polling interval, not a measured one. "Pulls" counts scheduled attempts; "success" is the share the adapter logged as successful, which is not the same as delivering data (see text).

Stream	Owner	Interface	Interval (set)	Pulls	Success
PM2.5, PM10, gases, 175 ground stations	Pollution Control Dept. (PCD)	JSON	15 min	3,865	99.3%
PM2.5 fused with satellite, 77 provinces	GISTDA	JSON	1 h	2,549	98.1%
Noise level, 27 stations	PCD (Noise4Thai)	HTML page, parsed	30 min	5,038	99.4%
Rain gauges, about 4,200	HII, for several agencies	JSON	10 min	14,701	99.1%
Weather forecast and stations	Thai Meteorological Dept.	JSON via sibling relay	15 min	4,235	98.2%
Wind, rain and PM2.5 model forecast	Open-Meteo (CAMS, weather models)	JSON	6 h	1,657	92.2%
Satellite rain (IMERG)	NASA GES DISC	OPeNDAP	30 min	5,077	100.0%
El Niño index (ONI)	NOAA CPC	text file	12 h	342	94.2%
Agricultural burn scars by crop	HII and Kasetsart Univ.	JSON	daily	105	97.1%
Drought risk, crop water use	GISTDA	JSON (GeoJSON)	6 h	230	98.3%
Fire detections, three VIIRS satellites	NASA FIRMS	CSV	6 h	196	92.9%
Aerosol optical depth, 5 sites	NASA AERONET	CSV	daily	46	97.8%
Airport visibility and weather codes	NOAA Aviation Weather Center	JSON	30 min	1,578	97.9%
Low-cost PM2.5 sensors, northern Thailand	Chiang Mai Univ. (DustBoy)	JSON	1 h	831	100.0%
About 1,000 public cameras	GISTDA, iTIC, DWR, volunteer aggregator	HLS, MJPEG, JPEG	30 min	3,370	99.1%
Haze news and citizen reports	News outlets (RSS)	RSS	20-30 min	7,474	99.1%
Flood verdicts (sibling system)	FloodDash	JSON, local	10 min	6,398	98.5%

Table 2. What broke, and the rule that now prevents it. Dates are 2026. Entries are taken from the project's operating record; the last was found while preparing this paper.

When	What happened	What is enforced now
19 Sep	Another local service bound the API's port and the tunnel served it; the public API was dead for about 20 hours.	The port moved; every response carries a service header; the edge and the watchdog treat a response without it as "backend down".
26–27 Sep	The live database was placed on a USB hard disk; synchronous queries froze the site for 4.5 hours (the collector log shows a 19.8-hour gap on 26–27 September).	The live database stays on the SSD as a bounded hot tier; the permanent record is archived to the external drive.
27 Sep	One syntax error in a browser module left every visitor on the boot splash while server health stayed green.	Deployment parses every browser module first and refuses to ship on error.
27 Sep	Thirteen restarts in one day spent the shared Open-Meteo quota.	A source that succeeded within its interval waits for its normal turn after a restart.
28 Sep	A bias ratio of 2.9 (Rayong) turned a raw forecast of 14 into "41", crossed the warning line and was broadcast.	The ratio applies only up to the CAMS level it was learned on; beyond that, an absolute offset.
24 Oct	The internal disk filled; log writes failed and the machine died, 30.9 hours of downtime.	A guard measures free space and sheds optional writes (camera frames) below 3 GB; readings keep running.
Nov	Found while preparing this paper, after the airport data analysed here were collected: 10 of 28 airport codes in the visibility table were dead and 3 were mislabelled.	The table was rebuilt from the aviation weather API; a verifier script checks it against the live service.

pulls) and DustBoy (831 pulls), are gated by a free token that nobody had registered; their adapters skip quietly when it is absent, and both logged 100% success. Together they are 10% of all pulls, so the headline success rate flatters the system. The bureaucracy we set out to bypass came back at the smallest scale, as a registration form, and the logs reported it as success.

5.3 Two official answers to one question GISTDA's satellite–ground fusion and PCD's stations should agree about a province. Over 5,748 paired province-days in 68 provinces they correlate ($\rho = 0.76$; 95% CI 0.72–0.79) but differ in level: GISTDA reads 2.0 µg m^{-3} higher on average, with 95% limits of agreement from −4.8 to 8.7 (Fig. 2b). Sorted into Thai air-quality bands, they place a province in the same band on 77% of days ($\kappa = 0.49$; weighted $\kappa = 0.52$), and GISTDA is the higher on 18% (Fig. 2c). Of the 1,314 disagreements, 88% lie between the two cleanest bands (15 or below, and 15–25), which GISTDA's offset largely explains. On the 5,748 province-days compared, PCD alone put a province above 25 on 51, GISTDA alone on 102, and both on 42. Above 37.5, GISTDA never did so; PCD did on 7 province-days. A citizen looking up the same province on two official sites will sometimes get two colours, usually adjacent ones. The two sources measure different things at different scales; we found no published rule for choosing between them.

5.4 The forecast CAMS read low against the ground. At the same-day lead its mean bias was −4.9 µg m^{-3} over 4,199 province-days (mean absolute error, MAE, 6.2), and the pooled ratio of observation to model was 1.62. The production correction cuts the error: MAE falls to 3.8, a change of −2.4 (95% CI −2.8–−2.0) (Fig. 5a–c). It is a real improvement, and it does not beat the simplest competitor. Carrying the last observed daily value forward has MAE 2.4 at the same-day lead, so the corrected model is 1.45 worse (CI 1.24–1.68). The gap narrows with lead (to 0.63 tomorrow and 0.50 the day after) but does not close. The correction leaves a residual bias of −1.2. The ratios behind it are large and uneven: the median province ratio was 1.9, and 14% of province-days sat at the clamp of 3.

For events the picture is thin because the season was quiet. At the same-day lead the corrected forecast detected 27 of 89 days above 25 (POD 0.30, FAR 0.72, CSI 0.17), against 0.29, 0.67 and 0.18 for persistence, and none of the 5 days above 37.5.

The system also sent warnings. Of 21 forecast warnings in 7 provinces, 13 distinct province-days could be checked against ground data. The median warned value was 41 µg m^{-3}; the median observed province mean was 25. Only 3 of 13 reached 37.5 at any station (Wilson 95% CI 8%–50%) and 1 in the province mean (Fig. 5d). 10 of 13 over-called, though not wildly: 6 reached 25 in the province mean. All 21 warnings

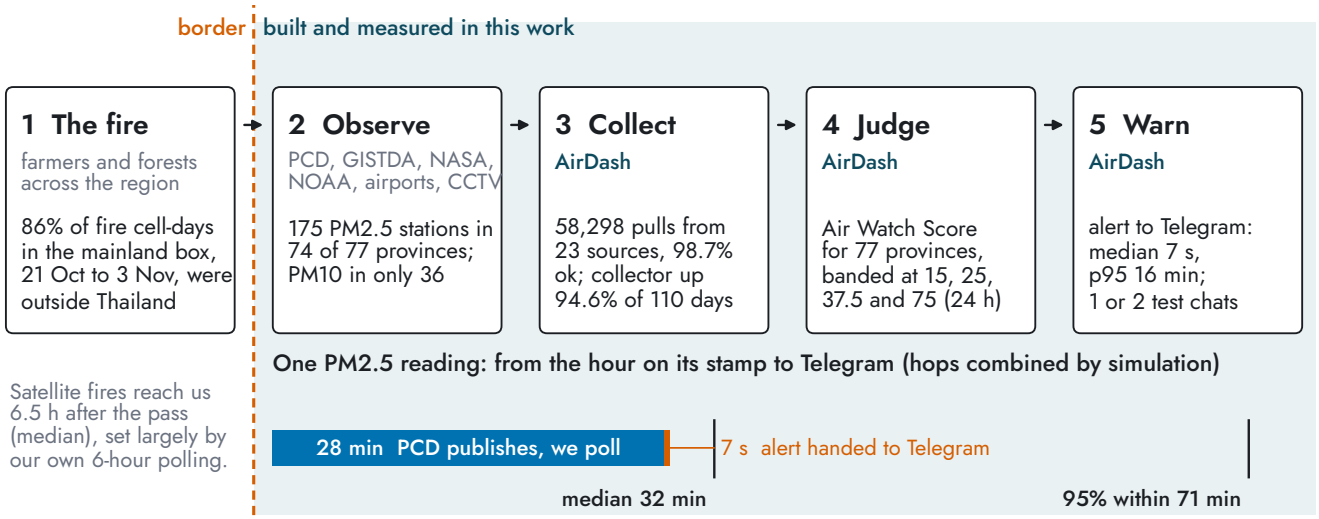


Fig. 1. From a burning field to a phone. The five stages of the system and the clock measured at each hop. Medians are from stored timestamps over the window; the satellite hop is the lag between a VIIRS overpass and the first time the system could see the detection.

were sent between 30 September and 30 October, after the bounded correction of Table 2, and the 13 province-days fall on only 6 target dates, so the interval above, which treats them as independent, is optimistic. We have not tested why they over-called.

5.5 Whose fire is it? In the 14 days from 21 October the detection feed held 196,385 detections, but 93% of them fell outside the box the production smoke indicator looks at, mostly in Indonesia, which the feed has covered since 19 October. In the production box (4–29°N, 92–110°E: mainland Southeast Asia and its neighbours) there were 13,779 detections, 1,348 of them inside Thailand. Three satellites often see the same fire, so we count 0.25° cells per day: 86% of the 5,159 fire cell-days lay outside Thailand (95% CI over days 84%–89%), and 91% of the fire radiative power (89%–93%) (Fig. 6b). The box also holds Laos, Cambodia, Vietnam, southern China, Peninsular Malaysia and Myanmar, so "outside" is not "Myanmar". On 73% of 847 province-days some upwind fire power (20 MW or more, distance-weighted) was present, and on 47% of the 617 active ones (CI 43%–50%, treating province-days as independent) most of it lay outside Thailand. Phrae, Nan and Chiang Rai led, narrowly: the fifth, Yasothon, stood at 170 MW against 187 (Fig. 6c). Of the upwind fire power for Nan, 73% lay across the border, and for Chiang Rai 68%.

Does the indicator predict anything? Across provinces and days, next-day $PM_{2.5}$ rose with the log of upwind fire power after removing each province's mean and each day's mean ($\rho = 0.28$; 95% CI over days 0.07–0.49; 527 province-days in 59 provinces, but only 9 days). In the eight upper-north provinces, where the claim matters most, it is $\rho = 0.53$ (CI 0.10–0.73); elsewhere $\rho = 0.24$ (CI 0.01–0.46). The result is not stable. Centring on the province alone gives $\rho = 0.07$ in the upper north, and eight provinces that share one wind are not eight independent tests. Nine days cannot settle it, and the window sits before the season in which it matters. We therefore do not claim that the indicator forecasts transported smoke; it describes where smoke could come from.

The domestic picture is steadier, for one kind of burning. The burn-scar product covers agricultural fires (rice, sugar cane and maize) in 62 provinces over three seasons; it does not map forest or grass fires. Agricultural burning is concentrated: ten provinces hold 52%, 60% and 53% of the burned area in 2024, 2025 and 2026, and the ranking barely moves between years ($\rho = 0.96$ for 2024 against 2026; CI 0.92–0.98); 8 of the ten are the same in all three years (Fig. 6a). The upper north holds only 6%–7% of the agricultural burned area of January to April, and there the crop is mostly maize (57%), where elsewhere it is paddy and sugar cane. How much of the upper north's smoke comes from forest fire instead, these maps cannot say. Agricultural burned area in January to April 2024, a season that followed an El Niño (the index stood at +1.84 in December–February, NOAA), was 23,454 km^2, 1.7 times that of 2025 (14,004 km^2); with three seasons, that is a hint and not a test.

5.6 What cameras and airports can see Public cameras are cheap and numerous, and of little use for this question so far. Of 114,803 frames from 993 cameras, 38% were taken in low light and 38% could be judged against a camera's own clear-day baseline. Across 21,535 frames from 104 cameras with a ground station nearby, contrast loss rose with $PM_{2.5}$ (Fig. 7a) but the correlation is weak: $\rho = 0.11$ (CI 0.07–0.15), and within a single camera over time the median ρ is 0.04. As an alarm, scored against the hourly $PM_{2.5}$ at the nearest station, the "possible haze" flag has specificity of 97% and sensitivity of 4% for frames at 25 or above, and precision of 11% (Fig. 7c). Of the 757 flagged frames with a humidity reading, 39% were at 90% relative humidity or above, where fog is the likelier cause; a 1-in-25 sample of 1,205 unflagged frames gave 38%, so humidity does not separate them.

Airport visibility fares no better. METAR caps visibility at 10 km, so 81% of hours carry no information about extinction. Below the ceiling, visibility and $PM_{2.5}$ are weakly related, and with the wrong sign for extinction, which should fall as $PM_{2.5}$ rises (ρ =+0.18; CI −0.07–0.24; 14 airports). In the 25–37.5 band, 64% of hours had visibility under 10 km against about 20% in the cleaner bands, but that rests on 270 hours at 5 airports (cluster CI 25%–72%), 76% of them from a single airport; without it the share is 33% (Fig. 7b). In 6,543 airport-hours the observers' weather codes never once attributed reduced visibility to haze or smoke, but those codes are normally reported only at low visibility and just 2.2% of hours were below 5 km. The test needs the burning season.

5.7 Alerts people would see The alert log holds 497 $PM_{2.5}$ level alerts, of which 91 were raised from PCD stations and 406 from

the GISTDA fusion. Alerts are hourly; people live by the day. Of 85 evaluable PCD alerts, 23 (27%; CI 19%–37%) were followed by a station-day mean above 37.5, although 74% stayed above the line for at least three of the next six hours. Of 399 evaluable GISTDA alerts, 0 were followed by a daily mean above 37.5, and 42% persisted for three hours. The PCD alerts were mostly right about the neighbourhood: 90% had a neighbouring station at 25 or above at the alert hour (CI 81%–95%). But 35% came from five stations, the first being the field-crop research centre in Rayong, one of the stations whose spikes stood alone. The alerts say something true about the next few hours and much less about the day, and fatigue is a known cost [18].

5.8 The year it arrives in The Oceanic Niño Index for August to October 2026 is +2.46, the highest of the 77 August–October seasons in NOAA's record, above 1997 and 2015. Mainland Southeast Asia responds less to El Niño than the Maritime Continent does [26], and we make no forecast of the coming season from it. We note it because the pre-specified tests below will be read in a year that is unusual on one dimension.

6 Tests fixed before the season

An early-season audit can tell what to doubt, not what to believe. Table 3 fixes, before the data exist, the tests that will decide what stays in the system. The evaluation window is 1 December 2026 to 30 April 2027; the analysis code is the code in the repository at the commit that releases this preprint, and the same scripts will be re-run unchanged. Two thresholds (T2 and T4) were set after we had seen this audit's results and are not independent of it. A test that fails changes the system: the feature it supports is removed or demoted, and the result is published with the rest.

7 Discussion

The three gaps are real, and the audit shows what each looks like in numbers.

Coordination. The data exist, and they do not agree. Two official sources assign a province to different air-quality bands on 23% of province-days; the ground network puts 45% of its stations in one city and 14 in the eight provinces the dust season hits hardest; seventeen streams (sixteen public feeds and one sister system) arrive through eight kinds of interface from more than ten organisations, two of them gated by a free account. None of this is a failure of any one agency. It is what a system looks like when nobody is assigned to the join. A small team could do the join, because every input is public. That is the argument for the approach, and it is also its limit: a join can show the disagreement but cannot say which source is right.

Responsiveness. Speed is not the bottleneck; judgement is. A reading reaches the system in about half an hour and an alert reaches Telegram's servers in seconds. But the alert rule fires on hours, people respond to days, and 10 of 13 forecast warnings never reached the standard at any station. Alert fatigue is a documented cost [18], and an alert that is quick and often unconfirmed invites it. The remedy in Table 3 is to push less.

The chain. The slowest measured link is the first, the fire detection (6.5 hours at the median), and much of that delay is our own six-hour polling, so the first fix is ours. The upwind indicator turned out to be what we said it was, a map of where smoke can come from, and not yet a forecast. The most useful thing in the chain may be the least glamorous: a dated, source-tagged statement that 86% of the fire cell-days on the mainland in a fortnight lay outside Thailand.

Bypassing, honestly. What did bypassing the bureaucracy buy? It bought speed (a working chain in one season, by one author with AI assistants and one laptop), cost (zero package dependencies; no budget line) and a record of its own failures that nobody had to approve. It did not buy the sensors that are missing, the agreement between sources, or the registration form for a free token. Nor does it touch the diplomatic layer. Satellites can say that a fire burns in a particular place; they cannot say who lit it, and we make no claim of attribution. What an open record can do for diplomacy is modest: it offers a shared set of numbers on which an argument could start. The treaty rests on each party's own policies for its own territory [8], and whether such a record changes anything is not something this paper tests.

What we would ask of others. These are proposals that follow from the audit, not findings. (i) Agree one published value per province, or publish the difference between sources with the reason. (ii) Fill the upper-north gap with calibrated low-cost sensors [28, 32] whose record is published beside the regulatory one. (iii) Pair every $PM_{2.5}$ station with a PM_{10} reading. (iv) Warn on persistence, not on the hour. (v) Publish the age of every fire detection beside it. (vi) Say what happens to a free-token source that has no token. None needs a treaty.

8 Limitations

The ground record is 110 days long and ends before the dust season, so we cannot say how the system behaves in the conditions it was built for. Most conclusions about burning rest on a 14-day window of fire detections and on three seasons of burn scars; the one forecast we could verify, CAMS, is the only model we tested, and we did not run trajectory models [49]. The correction, the thresholds and the weights of the score were set by the authors and not fitted to health outcomes, which we did not study. Station-day means rest on the PCD feed as served, and while we removed faults of 500 $\mu g\,m^{-3}$ or more, we did not audit calibration. The verification sample for high events is small: 5 province-days above 37.5. The push channel was exercised with one or two subscribed chats, and the latency we report is the server's hand-off to Telegram, so we say nothing about delivery to people at the scale of a population. Many of the intervals use Wilson bounds that treat events as independent when they cluster in time, and the fire and transport results rest on 14 days, of which nine have a next-day record. The airport data were collected before the visibility station table was rebuilt, when 3 of 28 codes were mislabelled. Camera and visibility analyses are pilots; their confidence intervals are wide and one airport dominates the visibility result. The system runs on one machine and its record contains its own gaps; five gaps of over six hours are unexplained. The code was largely written with AI assistants, which is a risk we reduced with tests (97 test files) and did not remove. We have not shown that the system changes anyone's behaviour or health, or that people can use it right away beyond a handful of test users. Finally, the audit was prepared by the system's developers; the tests above, and the public code, exist so that others can check it.

Declarations

Data and code. The system and every analysis in this paper are open at https://github.com/Nonarkara/airdash, under paper/; the derived data behind every figure and every number in the text are in paper/data, and the preprint is rebuilt from them with paper/build.py. The live database and archive that the extraction step reads are not released. The live system is at https://air.nonarkara.org. Upstream data are the public feeds of the organisations in Table 1.

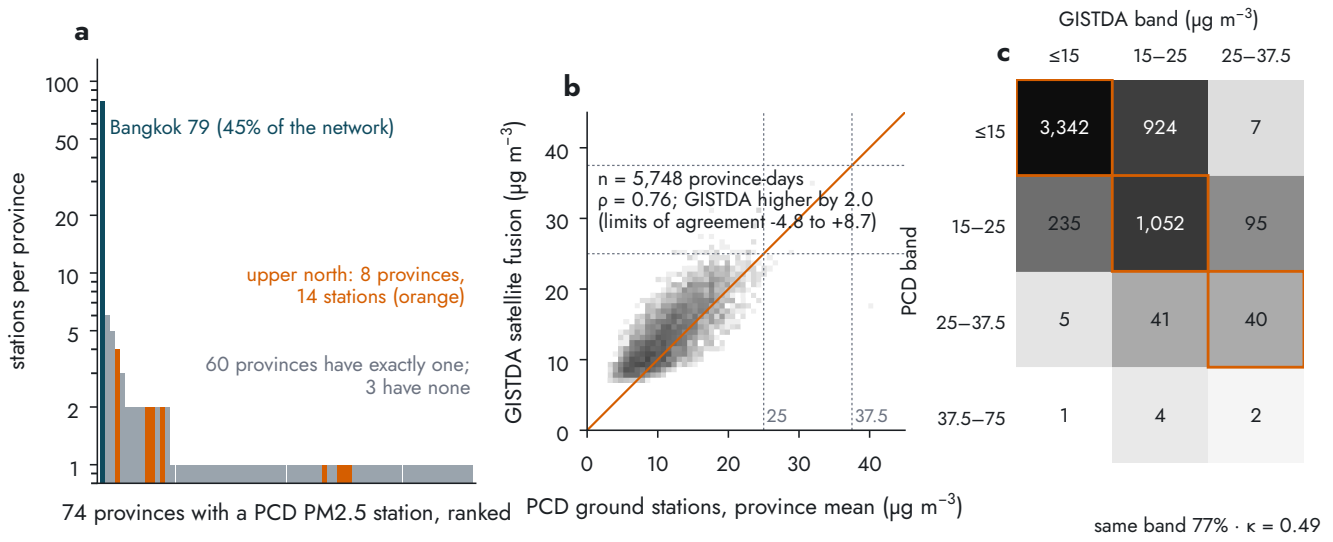


Fig. 2. Where the sensors are, and two official answers. a, Stations per province among the 74 provinces with a PCD $PM_{2.5}$ station (log scale); upper-north provinces in orange. **b**, Province-day means from PCD stations against the GISTDA satellite–ground fusion; the line is equality, dotted lines mark 25 and 37.5. **c**, The same province-days sorted into Thai air-quality bands; outlined cells are agreement.

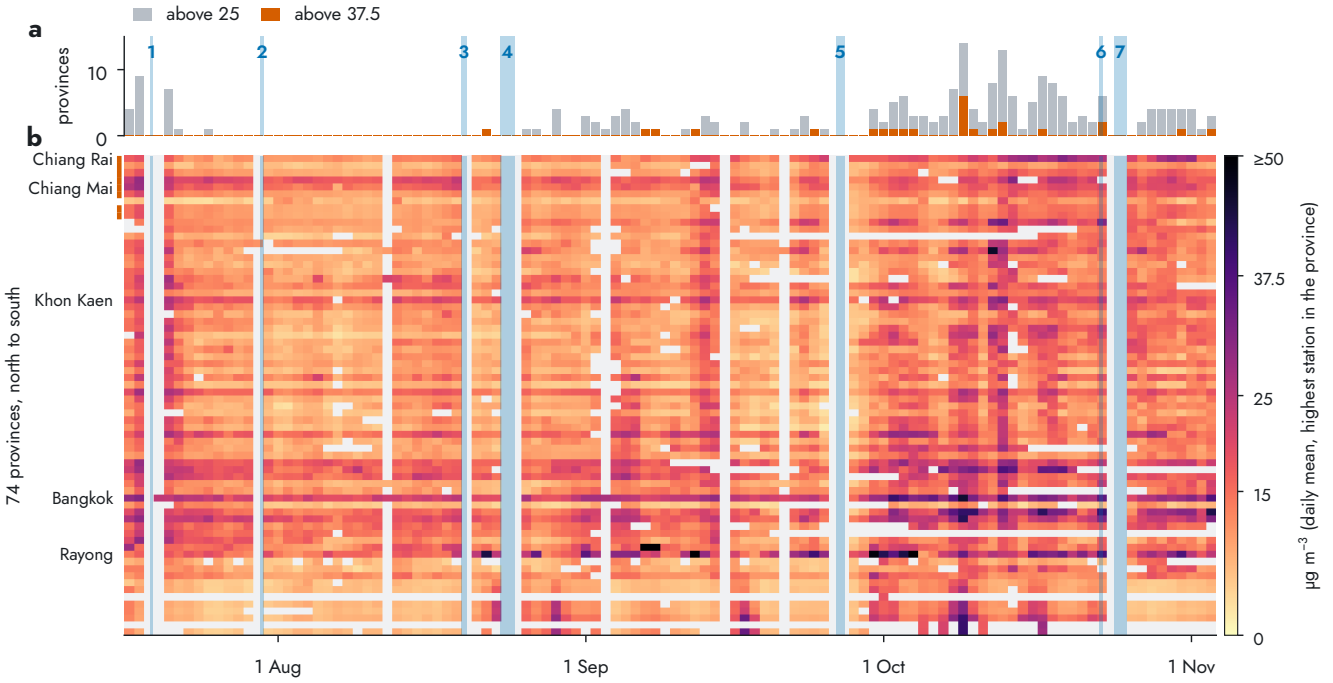


Fig. 3. The 110-day record. a, Provinces whose highest station-day mean exceeded 25 (grey) or 37.5 (orange) each day. **b**, Highest station-day mean per province and day, provinces ordered north to south. Pale blue bands are collector outages of over six hours (numbered as in Fig. 4a); grey cells are missing data. Orange ticks at left mark upper-north provinces.

Use of AI. The analysis code, the figures and the first draft of this text were produced with Claude (Anthropic) under the authors' direction, and the system's code was written with AI coding assistants. Every number is computed from the stored record by the code in the repository. References were checked against Crossref or the source's own page; for some, only the bibliographic record could be verified. An independent automated review found errors in the first version, which were corrected. The authors are responsible for the content.

Competing interests. N.A. developed and operates AirDash, which is free to use and carries no advertising; depa, N.A.'s affiliation, is named as a steward of the project.

Author contributions and funding. To be completed by the authors.

Ethics. Only public, aggregate environmental data were analysed. No personal data or subscriber records were used.

Acknowledgements. We thank the Pollution Control Department, GISTDA, the Hydro-Informatics Institute and Kasetsart University (Tam Roy Pao), the Thai Meteorological Department, NASA (FIRMS, AERONET, GES DISC), NOAA, Open-Meteo and Chiang Mai University for publishing data openly.

References

[1] Mueller W, Loh M, Vardoulakis S, Johnston HJ, Steinle S, Precha N, et al. (2020) Ambient particulate matter and biomass burning: an ecological time series study of respiratory and cardiovascular hospital visits in northern Thailand. *Environmental Health* **19**(1), 77. doi:10.1186/s12940-020-00629-3

[2] Uttajug A, Seposo X, Phosri A, Phung VLH, Tajudin MABA, Ueda K (2024) Effects of Coexposure to Air Pollution from Vegetation Fires and Extreme Heat on Mortality in Upper Northern Thailand. *Environmental Science & Technology* **58**(23), 9945-9953. doi:10.1021/acs.est.3c08074

[3] Sapbamrer P, Assavanopakun P, Panumasvivat J (2024) Decadal Trends in Ambient Air Pollutants and Their Association with COPD and Lung Cancer in Upper Northern Thailand: 2013–2022. *Toxics* **12**(5), 321. doi:10.3390/toxics12050321

[4] Johnston FH, Henderson SB, Chen Y, Randerson JT, Marlier M, DeFries RS, et al. (2012) Estimated global mortality attributable to smoke from landscape fires. *Environmental Health Perspectives* **120**(5), 695-701. doi:10.1289/ehp.1104422

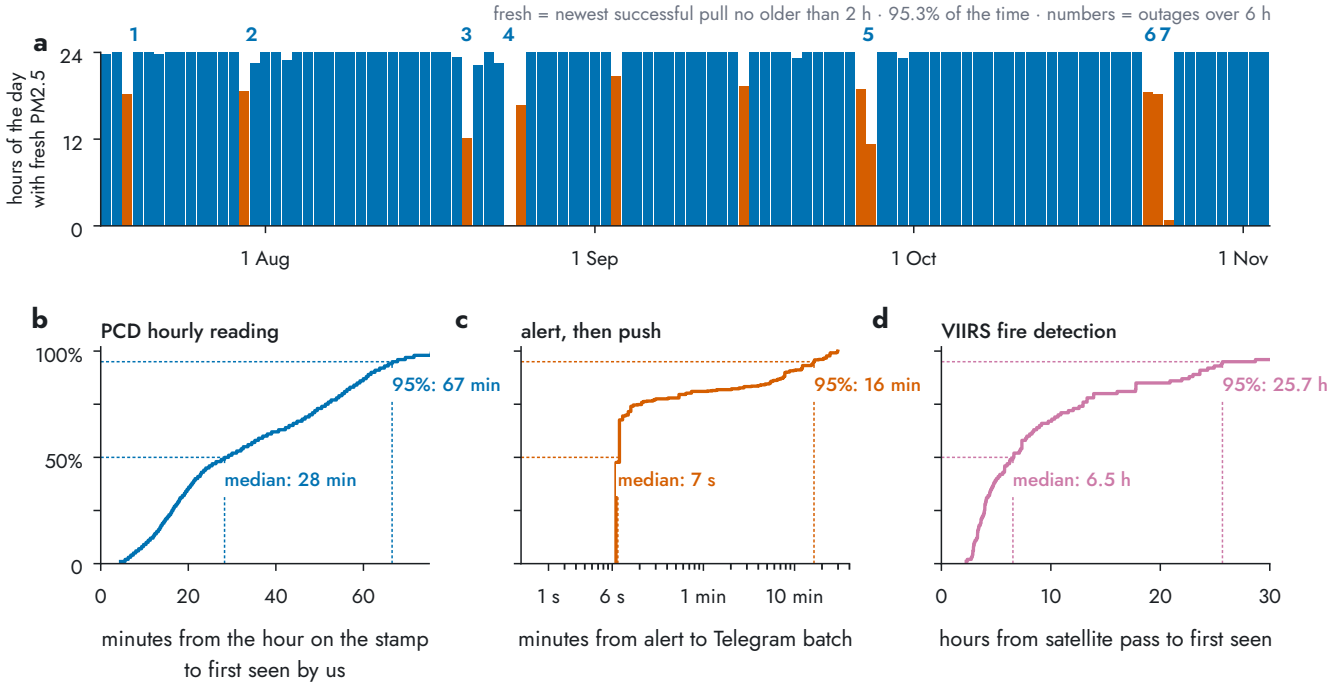


Fig. 4. Does it stay up, and how long is the wait? **a**, Hours of each day with a fresh PM2.5 feed (newest successful pull no older than two hours); orange days had a fresh feed for under 90% of the day and numbers mark the seven outages over six hours. **b–d**, Empirical distributions of the three clocks: the lag from the hour on a PCD reading to the first time the system saw it, the time from an alert to its Telegram batch, and the lag from a VIIRS overpass to first sight of the detection.

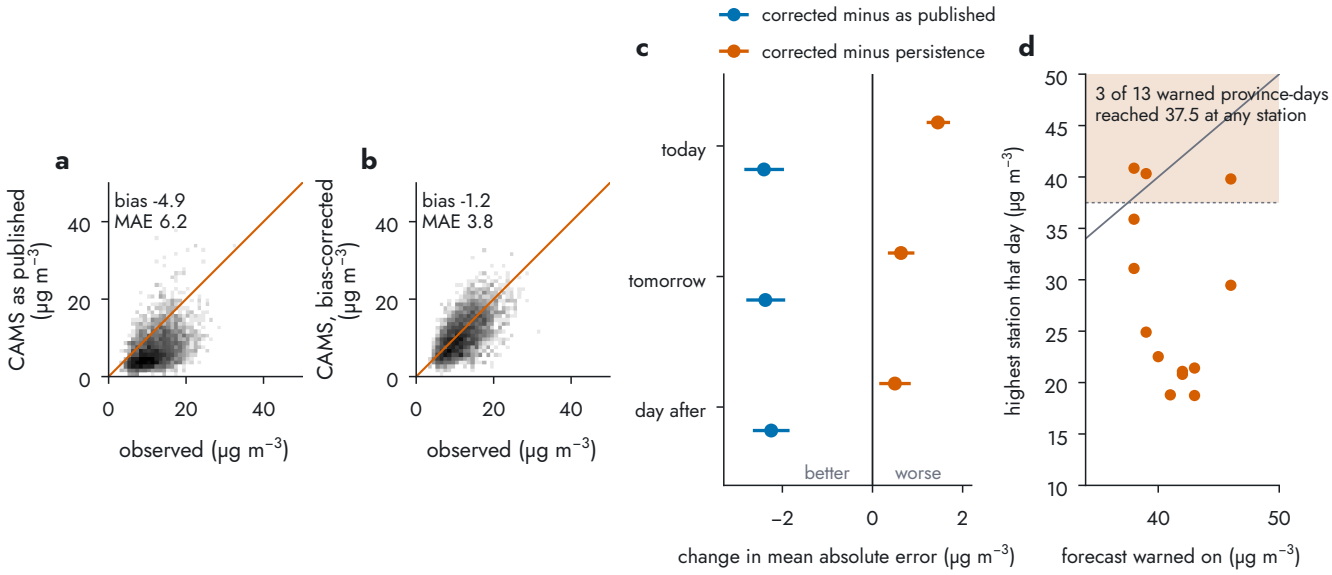


Fig. 5. The correction helps, persistence still wins, and the warnings over-called. **a, b**, Same-day CAMS forecast against province-day observations, as published (a) and after the out-of-sample bias correction (b); line is equality. **c**, Change in mean absolute error with 95% cluster-bootstrap intervals, by lead; left of zero is better. **d**, The 13 distinct province-days on which a forecast warning was sent: the highest value warned for the province-day against the highest station that day; the shaded region is where the Thai standard was reached.

[5] Chantaraprachoom N, Shimadera H, Uranishi K, Mui LV, Matsuo T, Kondo A (2024) A Nation-by-Nation Assessment of the Contribution of Southeast Asian Open Biomass Burning to PM2.5 in Thailand Using the Community Multi-scale Air Quality-Integrated Source Apportionment Method Model. *Atmosphere* **15**(11), 1358. doi:10.3390/atmos15111358

[6] Amnuaylojaroen T, Inkom J, Janta R, Surapipith V (2020) Long Range Transport of Southeast Asian PM2.5 Pollution to Northern Thailand during High Biomass Burning Episodes. *Sustainability* **12**(23), 10049. doi:10.3390/su122310049

[7] Gu Y, Fang T, Yim SHL (2024) Source emission contributions to particulate matter and ozone, and their health impacts in Southeast Asia. *Environment International* **186**, 108578. doi:10.1016/j.envint.2024.108578

[8] ASEAN (2002) ASEAN Agreement on Transboundary Haze Pollution (signed Kuala Lumpur, 10 June 2002; in force 25 November 2003). ASEAN Secretariat (agreement.asean.org). https://agreement.asean.org/media/download/20220330063051.pdf (accessed 2026-11-04)

[9] ASEAN Secretariat (2015) Indonesia Deposits Instrument of Ratification of the ASEAN Agreement on Transboundary Haze Pollution. ASEAN Environment Knowledge Hub (media release, 21 January 2015; Internet Archive copy). https://web.archive.org/web/20230810064427/https://environment.asean.org/news/detail/indonesia-deposits-instrument-of-ratification-of-the-asean-agreement-on-transboundary-haze-pollution (accessed 2026-11-04)

[10] ASEAN Haze Portal / ASEAN Secretariat (2022) ASEAN Agreement on Transboundary Haze Pollution (status and ratification table). hazeportal.asean.org (Internet Archive snapshot 2022-12-01). https://web.archive.org/web/20221201110219/https://hazeportal.asean.org/action/asean-agreement-on-transboundary-haze-pollution (accessed 2026-11-04)

[11] Heilmann D (2015) After Indonesia's Ratification: The ASEAN Agreement on Transboundary Haze Pollution and its Effectiveness as a Regional Environmental Governance Tool. *Journal of Current Southeast Asian Affairs* **34**(3), 95-121. doi:10.1177/186810341503400304

[12] Hurley A, Lee T (2020) Delayed ratification in environmental regimes: Indonesia's ratification of the ASEAN agreement on transboundary haze pollution. *The Pacific Review* **34**(6), 1108-1137. doi:10.1080/09512748.2020.1801816

[13] ASEAN Haze Portal / ASEAN Secretariat (2016) Roadmap on ASEAN Cooper-

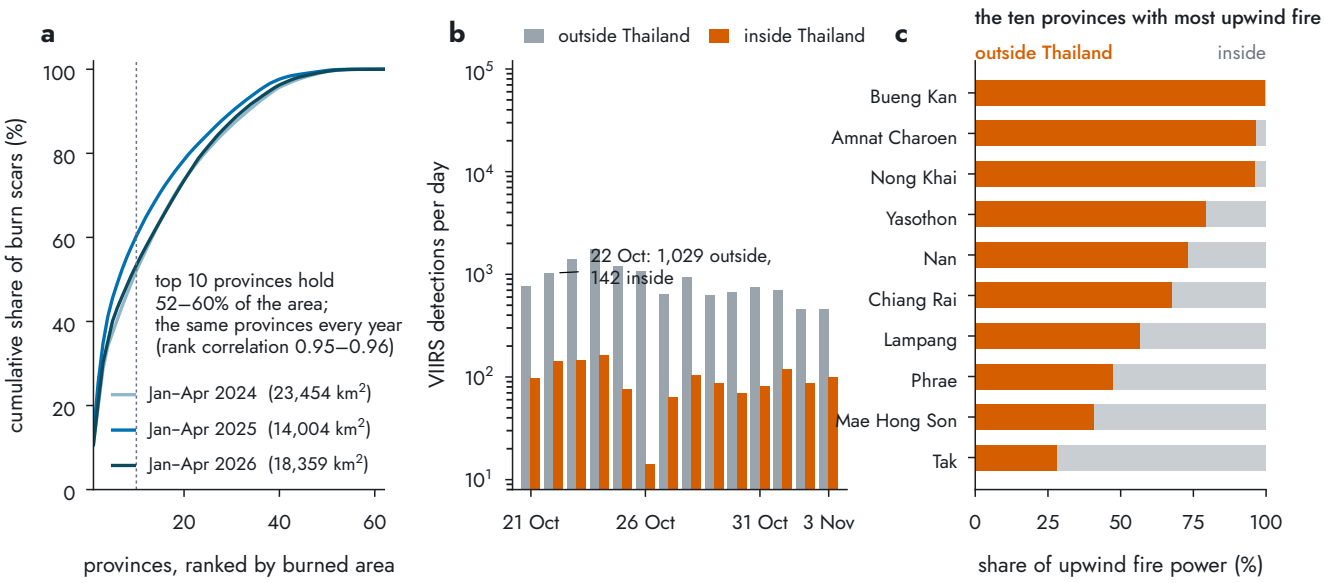


Fig. 6. Where the fire is. a, Cumulative share of burn-scar area by province, ranked, for January–April of three seasons. **b**, VIIRS detections per day (log scale) inside and outside Thailand in the production box (4–29°N, 92–110°E), 21 October to 3 November. **c**, For the ten provinces with the most upwind fire power, the share of that power from outside Thailand.

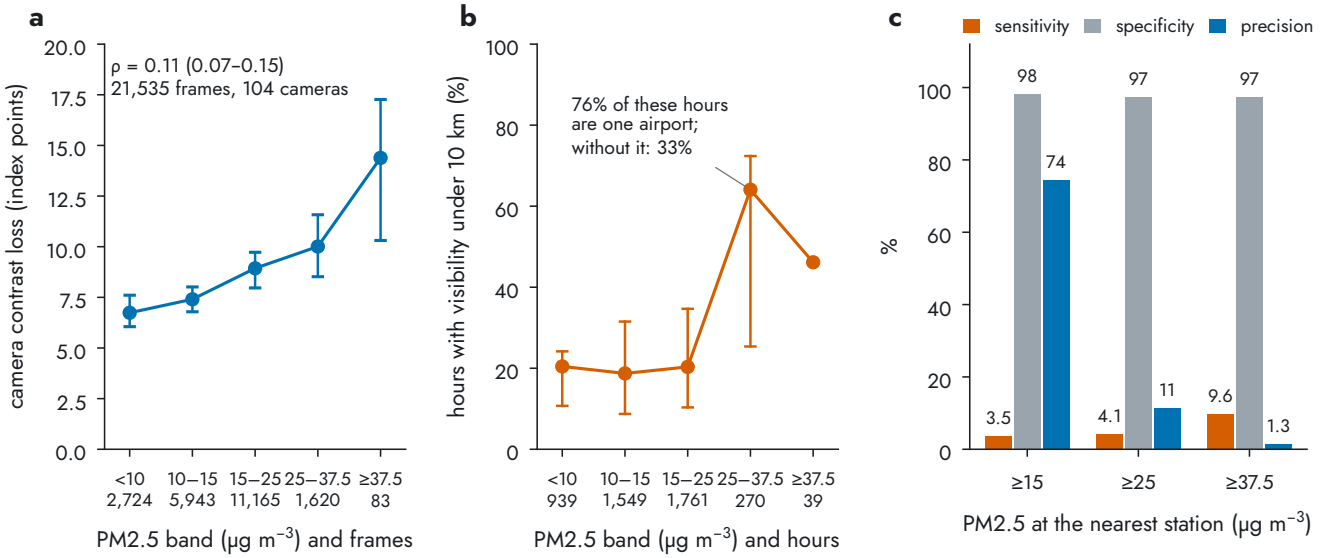


Fig. 7. What cameras and airport visibility can and cannot say. a, Camera contrast loss (a relative index, not a concentration) by PM2.5 band at the nearest station; bars are 95% cluster-bootstrap intervals over cameras. **b**, Share of airport-hours with visibility below 10 km, by band of PM2.5 at the nearest station. **c**, Sensitivity, specificity and precision of the "possible haze" camera flag at three thresholds.

ation towards Transboundary Haze Pollution Control with Means of Implementation (2016-2020) (Haze-free Roadmap). hazeportal.asean.org (Internet Archive snapshot 2022-12-01). https://web.archive.org/web/20221201101343/https://hazeportal.asean.org/action/roadmap-on-asean-cooperation-towards-transboundary-haze-pollution-control-with-means-of-implementation/ (accessed 2026-11-04)

[14] ASEAN (2023) Media Release: 17th ASEAN Ministerial Meeting on the Environment and 18th Meeting of the COP to the ASEAN Agreement on Transboundary Haze Pollution (Vientiane, 23 August 2023). ASEAN Secretariat. https://asean.org/wp-content/uploads/2023/08/Media-Release-AMME17_COP18-Adopted_23Aug2023_final.pdf (accessed 2026-11-04)

[15] Nikam J, Archer D, Nopsert C (2021) Air quality in Thailand: Understanding the regulatory context (SEI working paper). Stockholm Environment Institute. https://www.sei.org/wp-content/uploads/2021/02/210212c-killeen-archer-air-quality-in-thailand-wp-2101e-final.pdf (accessed 2026-11-04)

[16] Poapongsakorn N, Pantakua K, Ratchakom S (2023) Challenges in tackling PM2.5 crisis (TDRI Insight; first published in the Bangkok Post, 12 April 2023). Thailand Development Research Institute (TDRI). https://tdri.or.th/en/2023/04/challenges-in-tackling-pm2-5-crisis/ (accessed 2026-11-04)

[17] Marks D, Miller MA (2022) A transboundary political ecology of air pollution: Slow violence on Thailand's margins. *Environmental Policy and Governance* **32**(4), 305-319. doi:10.1002/eet.1976

[18] Saberian S, Heyes A, Rivers N (2017) Alerts work! Air quality warnings and cycling. *Resource and Energy Economics* **49**, 165-185. doi:10.1016/j.reseneeco.2017.05.004

[19] Janjua S, Powell P, Atkinson R, Stovold E, Fortescue R (2021) Individual-level interventions to reduce personal exposure to outdoor air pollution and their effects on people with long-term respiratory conditions. *Cochrane Database of Systematic Reviews* **2021**(8). doi:10.1002/14651858.cd013441.pub2

[20] Riley R, de Preux L, Capella P, Mejia C, Kajikawa Y, de Nazelle A (2021) How do we effectively communicate air pollution to change public attitudes and behaviours? A review. *Sustainability Science* **16**(6), 2027-2047. doi:10.1007/s11625-021-01038-2

[21] Lee HH, Bar-Or RZ, Wang C (2017) Biomass burning aerosols and the low-visibility events in Southeast Asia. *Atmospheric Chemistry and Physics* **17**(2), 965-980. doi:10.5194/acp-17-965-2017

[22] Reid JS, Hyer EJ, Johnson RS, Holben BN, Yokelson RJ, Zhang J, et al. (2013) Observing and understanding the Southeast Asian aerosol system by remote sensing: An initial review and analysis for the Seven Southeast Asian Studies (7SEAS) program. *Atmospheric Research* **122**, 403-468. doi:10.1016/j.atmosres.2012.06.005

[23] Lin NH, Tsay SC, Maring HB, Yen MC, Sheu GR, Wang SH, et al. (2013) An overview of regional experiments on biomass burning aerosols and related pollutants in Southeast Asia: From BASE-ASIA and the Dongsha Experiment to 7-SEAS. *Atmospheric Environment* **78**, 1-19. doi:10.1016/j.atmosenv.2013.04.066

[24] Reid JS, Xian P, Hyer EJ, Flatau MK, Ramirez EM, Turk FJ, et al. (2012) Multi-scale meteorological conceptual analysis of observed active fire hotspot activity

Table 3. Pre-specified tests for the 2026–27 season. Baselines are this paper's results. "Pass" is stated before the season; "if not" is what the system does about a failure. CI is a 95% interval; CAMS is the bias-corrected forecast.

	Question	Baseline (this paper)	Pass	If not
T1	Is a severe alert handed to Telegram within 35 minutes of being raised? (A guard on the batching, not a measure of delivery.)	100% of 485 alerts; median 7 s	At least 95% in the season	Shorten the batch gap, or move the busiest hours to a priority queue
T2	Are PCD alerts followed by a bad day?	27% of alerts followed by a station-day mean above 37.5 (CI 19%–37%)	For alerts issued only when at least three of the last six hourly values were above the line, lower CI limit at least 50%; replayed on the season's data if the rule is not live	Keep hourly alerts to a "watch" tier and push only the daily mean
T3	Does the corrected forecast beat carrying the last observation forward two days ahead?	MAE difference +0.50 (CI +0.19–+0.81), worse than carrying forward	Upper CI limit of the difference below zero	Stop issuing the second-day warning; show it as a range, not a value
T4	Does upwind fire predict next-day $PM_{2.5}$ in the upper north, February–April?	$\rho = 0.53$ (CI over days 0.10–0.73; 9 days), 0.07 without day-centring	$\rho \geq 0.25$ with the lower CI limit above zero, over at least 60 days	Describe the indicator as a map of sources only, and drop it from the verdict
T5	Do cameras track $PM_{2.5}$ within a camera over time?	Median within-camera $\rho = 0.04$	Median $\rho \geq 0.30$ among cameras with at least 20 judged frames	Remove the rule that lets cameras raise a province to WATCH
T6	Does low visibility at airports separate smoky from clean hours?	64% of hours under 10 km at 25–37.5 against 20% below 10; one airport supplies 76%	Gap of at least 20 points, cluster CI above zero, no airport over 30% of hours	Keep visibility as a reason on the card and never in the score
T7	Does success mean data?	Two token-gated sources logged 100% success and wrote no rows	Zero token-gated sources that skip silently	A skipped source is reported as "not configured" on the status page

and smoke optical depth in the Maritime Continent. *Atmospheric Chemistry and Physics* **12**(4), 2117-2147. doi:10.5194/acp-12-2117-2012

[25] Marlier ME, DeFries RS, Voulgarakis A, Kinney PL, Randerson JT, Shindell DT, et al. (2012) El Niño and health risks from landscape fire emissions in southeast Asia. *Nature Climate Change* **3**(2), 131-136. doi:10.1038/nclimate1658

[26] Nurdiati S, Sopaheluwakan A, Septiawan P (2022) Joint Pattern Analysis of Forest Fire and Drought Indicators in Southeast Asia Associated with ENSO and IOD. *Atmosphere* **13**(8), 1198. doi:10.3390/atmos13081198

[27] Marvin MR, Palmer PI, Yao F, Latif MT, Khan MF (2024) Uncertainties from biomass burning aerosols in air quality models obscure public health impacts in Southeast Asia. *Atmospheric Chemistry and Physics* **24**(6), 3699-3715. doi:10.5194/acp-24-3699-2024

[28] Barkjohn KK, Gantt B, Clements AL (2021) Development and application of a United States-wide correction for PM2.5 data collected with the PurpleAir sensor. *Atmospheric Measurement Techniques* **14**(6), 4617-4637. doi:10.5194/amt-14-4617-2021

[29] Giordano MR, Malings C, Pandis SN, Presto AA, McNeill VF, Westervelt DM, et al. (2021) From low-cost sensors to high-quality data: A summary of challenges and best practices for effectively calibrating low-cost particulate matter mass sensors. *Journal of Aerosol Science* **158**, 105833. doi:10.1016/j.jaerosci.2021.105833

[30] Morawska L, Thai PK, Liu X, Asumadu-Sakyi A, Ayoko G, Bartonova A, et al. (2018) Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone?. *Environment International* **116**, 286-299. doi:10.1016/j.envint.2018.04.018

[31] Castell N, Dauge FR, Schneider P, Vogt M, Lerner U, Fishbain B, et al. (2017) Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?. *Environment International* **99**, 293-302. doi:10.1016/j.envint.2016.12.007

[32] Dejchanchaiwong R, Tekasakul P, Saejio A, Limna T, Le TC, Tsai CJ, et al. (2023) Seasonal Field Calibration of Low-Cost PM2.5 Sensors in Different Locations with Different Sources in Thailand. *Atmosphere* **14**(3), 496. doi:10.3390/atmos14030496

[33] Chiang Mai University Climate Change Data Center (CCDC) (n.d.) CMU CCDC website - DustBoy PM2.5 network pages (status, API, installation guide). cmuccdc.org. https://www.cmuccdc.org/ (accessed 2026-11-04)

[34] Schroeder W, Oliva P, Giglio L, Csiszar IA (2014) The New VIIRS 375 m active fire detection data product: Algorithm description and initial assessment. *Remote Sensing of Environment* **143**, 85-96. doi:10.1016/j.rse.2013.12.008

[35] Davies DK, Ilavajhala S, Wong MM, Justice CO (2009) Fire Information for Resource Management System: Archiving and Distributing MODIS Active Fire Data. *IEEE Transactions on Geoscience and Remote Sensing* **47**(1), 72-79. doi:10.1109/tgrs.2008.2002076

[36] NASA LANCE (n.d.) Fire Information for Resource Management System (FIRMS). firms.modaps.eosdis.nasa.gov. https://firms.modaps.eosdis.nasa.gov/ (accessed 2026-11-04)

[37] Holben BN, Eck TF, Slutsker I, Tanré D, Buis JP, Setzer A, et al. (1998) AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization. *Remote Sensing of Environment* **66**(1), 1-16. doi:10.1016/s0034-4257(98)00031-5

[38] Giles DM, Sinyuk A, Sorokin MG, Schafer JS, Smirnov A, Slutsker I, et al. (2019) 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 (AOD) measurements. *Atmospheric Measurement Techniques* **12**(1), 169-209. doi:10.5194/amt-12-169-2019

[39] Schuster GL, Dubovik O, Holben BN (2006) Angstrom exponent and bimodal aerosol size distributions. *Journal of Geophysical Research: Atmospheres* **111**(D7), 2005JD006328. doi:10.1029/2005jd006328

[40] Ångström A (1929) On the Atmospheric Transmission of Sun Radiation and on Dust in the Air. *Geografiska Annaler* **11**(2), 156-166. doi:10.1080/20014422.1929.11880498

[41] Geo-Informatics and Space Technology Development Agency (GISTDA) (n.d.) Disaster Platform - Fire (hotspot service). disaster.gistda.or.th. https://disaster.gistda.or.th/fire (accessed 2026-11-04)

[42] Tam Roy Pao project (site hosted by the Hydro-Informatics Institute) (2024) Tam Roy Pao - open burning / burn-scar monitoring system. tamroypao.hii.or.th. https://tamroypao.hii.or.th/ (accessed 2026-11-04)

[43] Suwanprasit C, Shahnawaz (2024) Mapping burned areas in Thailand using Sentinel-2 imagery and OBIA techniques. *Scientific Reports* **14**(1), 9609. doi:10.1038/s41598-024-60512-w

[44] Inness A, Ades M, Agustí-Panareda A, Barré J, Benedictow A, Blechschmidt AM, et al. (2019) The CAMS reanalysis of atmospheric composition. *Atmospheric Chemistry and Physics* **19**(6), 3515-3556. doi:10.5194/acp-19-3515-2019

[45] Tan Z, Latif MT, Ashfold MJ (2023) Assessment of Malaysia-wide PM2.5 Forecasts from a Global Model. *Aerosol and Air Quality Research* **23**(9), 220444. doi:10.4209/aaqr.220444

[46] Wu C, Li K, Bai K (2020) Validation and Calibration of CAMS PM2.5 Forecasts Using In Situ PM2.5 Measurements in China and United States. *Remote Sensing* **12**(22), 3813. doi:10.3390/rs12223813

[47] Ali MA, Bilal M, Wang Y, Nichol JE, Mhawish A, Qiu Z, et al. (2022) Accuracy assessment of CAMS and MERRA-2 reanalysis PM2.5 and PM10 concentrations over China. *Atmospheric Environment* **288**, 119297. doi:10.1016/j.atmosenv.2022.119297

[48] Glahn HR, Lowry DA (1972) The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. *Journal of Applied Meteorology* **11**(8), 1203-1211. doi:10.1175/1520-0450(1972)011<1203:tuomos>2.0.co;2

[49] Stein AF, Draxler RR, Rolph GD, Stunder BJB, Cohen MD, Ngan F (2015) NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System. *Bulletin of the American Meteorological Society* **96**(12), 2059-2077. doi:10.1175/bams-d-14-00110.1

[50] Koschmieder H (1924) Theorie der horizontalen Sichtweite. *Beitraege zur Physik der freien Atmosphaere* **12**, 33-53. https://library.metoffice.gov.uk/portal/Default/en-GB/RecordView/Index/457195 (accessed 2026-11-04)

[51] Fei Y, Liao J, Zhang Z (2023) Consistency and Discrepancy between Visibility and PM2.5 Measurements: Potential Application of Visibility Observation to Air Quality Study. *Sensors* **23**(2), 898. doi:10.3390/s23020898

[52] He K, Sun J, Tang X (2011) Single Image Haze Removal Using Dark Channel Prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**(12), 2341-2353. doi:10.1109/tpami.2010.168

[53] Shin J, Kim D, Noh Y (2022) Estimation of Aerosol Extinction Coefficient Using Camera Images and Application in Mass Extinction Efficiency Retrieval. *Remote Sensing* **14**(5), 1224. doi:10.3390/rs14051224

[54] Laohakiat S, Klerkkidakan S, Wiwatwattana N (2023) Visually Estimating and Forecasting PM2.5 Levels Using Hybrid Architecture Deep Neural Network. *CURRENT APPLIED SCIENCE AND TECHNOLOGY*, e0258074. doi:10.55003/cast.2023.258074

[55] Wu Y, Wang X, Wang M, Liu X, Zhu S (2024) Time-Series Forecasting of PM2.5 and PM10 Concentrations Based on the Integration of Surveillance

Images. *Sensors* **25**(1), 95. doi:10.3390/s25010095
[56] Theerawasttanasiri N, Pattara-anantanop N, Chamnankong K, Samdangsuk I, Winun A (2024) Accuracy of estimating a human-sighted visual range for short-term particulate matter concentration in Chiang Mai, Thailand. *International Journal of Public Health Asia Pacific*, 46-58. doi:10.62992/vfg2zf31
[57] Janke K (2014) Air pollution, avoidance behaviour and children's respiratory health: Evidence from England. *Journal of Health Economics* **38**, 23-42. doi:10.1016/j.jhealeco.2014.07.002
[58] Mahajan S, Chung MK, Martinez J, Olaya Y, Helbing D, Chen LJ (2022) Translating citizen-generated air quality data into evidence for shaping policy. *Humanities and Social Sciences Communications* **9**(1), 122. doi:10.1057/s41599-022-01135-2
[59] Pollution Control Department (Thailand) (2023) Notification of the Pollution Control Department: Air Quality Index of Thailand B.E. 2566 (2023) (dated 16 May 2023; Royal Gazette Vol. 140, Special Part 157 Ng, 3 July 2023). Royal Gazette. https://ratchakitcha.soc.go.th/documents/140D157S0000000000300.pdf (accessed 2026-11-04)
[60] National Environment Board (Thailand) (2022) Notification of the National Environment Board: Setting the standard of particulate matter not exceeding 2.5 microns in ambient air (signed 23 June 2022; Royal Gazette Vol. 139, Special Part 163 Ng, 8 July 2022, p. 21-22). Royal Gazette (copy hosted by Pollution Control Department, pcd.go.th). https://www.pcd.go.th/wp-content/uploads/2022/07/pcdnew-2022-07-12_02-24-56_120798.pdf (accessed 2026-11-04)
[61] World Health Organization (2021) WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization, Geneva (ISBN 9789240034228 electronic; 9789240034211 print). https://www.who.int/publications/i/item/9789240034228 (accessed 2026-11-04)
[62] Efron B (1979) Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics* **7**(1), 1-26. doi:10.1214/aos/1176344552
[63] Wilson EB (1927) Probable Inference, the Law of Succession, and Statistical Inference. *Journal of the American Statistical Association* **22**(158), 209-212. doi:10.1080/01621459.1927.10502953
[64] Brown LD, Cai TT, DasGupta A (2001) Interval Estimation for a Binomial Proportion. *Statistical Science* **16**(2), 101-133. doi:10.1214/ss/1009213286
[65] Spearman C (1904) The Proof and Measurement of Association between Two Things. *The American Journal of Psychology* **15**(1), 72. doi:10.2307/1412159
[66] Jolliffe IT, Stephenson DB (2012) Forecast Verification: A Practitioner's Guide in Atmospheric Science. Wiley, 2nd edition. doi:10.1002/9781119960003
[67] Wilks DS (2019) Statistical Methods in the Atmospheric Sciences. Elsevier (Academic Press), 4th edition; ISBN 9780128158234. doi:10.1016/c2017-0-03921-6