Data Methodology

Indonesia air quality tracker

10 September 2026

How the air quality readings shown on this tracker are collected, cleaned, and published, and how the population-exposure and thermal anomalies layers are produced.

Air quality recordings

Data sources

All readings originate from government-operated monitoring stations across Indonesia. The tracker draws from three networks:

Operator	Coverage	Reporting
Ministry of Environment and Forestry (KLHK)	~117 stations nationwide	Hourly readings and 24-hour rolling averages
Meteorological Agency (BMKG)	~26 stations, major cities	Hourly readings
Provincial environmental agencies (DLH)	12 reference-grade sites in Jakarta	Half-hourly readings

Counts are of sites that currently pass every check below. DKI Jakarta's portal publishes around 125 sites in total; the great majority are low-cost or unclassified PM2.5-only sensors, which are excluded from the tracker.

Processing pipeline

Station readings pass through a fixed sequence of steps before reaching the tracker. Measurements that are flagged as suspect are ignored.

1- Collection

Raw readings are fetched from each network's reporting system at station level, with no upstream pre-processing.

2- Quality control

Each reading is checked against a series of rules. Readings that fail any check are flagged as invalid and excluded from published averages. See the full list of checks below.

3- Unit conversion

All published values are in micrograms per cubic metre (µg/m^3). Stations that report only the national air quality index (ISPU) are converted to µg/m^3 using the official breakpoint table (Permen LHK P.14/2020). A direct measurement always takes priority over a converted index value.

4- Daily averaging

A station's daily mean is calculated from its valid readings for that day. A minimum of 12 distinct hours must be covered for a daily value to be published; days with too few readings are excluded rather than given a misleading average. When using KLHK 24-hour rolling data, we take the last value of the day.

5- Instrument merging

Several identifiers can describe one physical device: the same analyser published on an hourly and a 24-hour channel, or a station republished by a second network. These are collapsed into a single instrument before anything is averaged, so one device counts once. See below.

6- City aggregation

The city-level value is the unweighted mean across every reference-grade instrument in that city that passes the checks above, after merging. Different networks contribute to the same series.

To stop a city's line from stepping when its composition changes rather than its air, an instrument must cover at least 20% of the displayed window to join the basket. Where no instrument clears that bar, the best-covered one is kept, so a city is never emptied. Jakarta is published as its five districts plus an average across them, which requires at least three districts reporting.

Quality checks

Every reading passes through these checks, in order. A reading that fails any check is flagged with the reason and excluded from all averages. The original value is retained in the database.

Check	What it catches
Missing values	Null or absent readings
Negative values	Readings below zero, including the −9999 error sentinel some stations emit
Zero values	Exact zeros, which for air pollutant concentrations indicate a reporting fault rather than truly clean air
Out-of-range index	ISPU index values outside the 0–500 scale defined by regulation
Physical limits	Concentrations above what is physically plausible (e.g. PM2.5 above 2,000 $\mu g/m^3$)

Consecutive repeats	An identical value reported for 24+ consecutive hours, typically a stuck instrument
Flatline	Zero variation over 48+ hours
Low variance	Near-constant readings over a full week (stuck sensor reporting with slight noise)
Statistical outlier	A reading that deviates sharply from the station's own 30-day behaviour, measured using the median absolute deviation (MAD).

A second set judges the instrument, over its whole recent record, and withdraws a device from publication for as long as the fault lasts:

Check	**Period judged**	**What it catches**
Dropout	A day	A daily mean at or below 1 µg/m^3, which no operating monitor produces
Sentinel	A day	The value 1000.0 exactly, an error code some stations emit
Frozen	A month	Almost no variation across the month, or across any run of 20 reporting days within it
Mostly dropout	A month	Half or more of the month's days sitting at the floor
Near-zero flatline	A month	A median below 3 µg/m^3 sustained over 20 days or more, typically a sensor reading clean air while its neighbours record haze

Population exposure

Population exposure is mapped from TAP (Tracking Air Pollution), a daily global PM2.5 product at about 11 km resolution that fuses ground measurements, satellite aerosol readings and chemical transport modelling (Geng et al., 2021).

Checked against Indonesian monitors, TAP reads systematically low, by an amount that varies from place to place and is larger for the share of PM2.5 it attributes to fire smoke: in September 2026 the monitors read about 1.2 times TAP's non-fire PM2.5 and about 2.3 times its fire PM2.5, nationally. We correct it with two scaling factors, one for the smoke share and one for the rest, fitted by geographically weighted regression on the last 30 days of paired monitor and TAP values.

TAP arrives about three days late, so the newest days start from its latest fields, move with the monitors' readings since, and are marked provisional until it catches up.

The corrected surface is weighted by the Gridded Population of the World v4.11 grid (CIESIN, 2018), with each province scaled so that it sums to Statistics Indonesia's (BPS) mid-2025 estimate, 284 million nationally (Badan Pusat Statistik, 2023)

The result is a distribution: how many people are exposed to concentrations in each air quality band, plus a single population-weighted mean concentration. Bands follow the Indonesian ISPU standard (Good, Moderate, Unhealthy, Very unhealthy, Hazardous).

Thermal anomalies

We use NASA's VNP14A1 product: the daily thermal anomaly composite from the VIIRS instrument aboard the Suomi-NPP satellite, gridded at 1 km from the instrument's 750 m infrared bands (NASA LP DAAC, 2023). It identifies pixels radiating substantially more heat than their surroundings, which is an actively burning fire, and records each detection's fire radiative power in megawatts. Fire radiative power is the rate at which a fire releases energy, and it scales with how much biomass is being consumed. Summing it across a region measures how intensely the landscape is burning rather than how many fires happened to be spotted.

References

Badan Meteorologi, Klimatologi, dan Geofisika (BMKG). Kualitas udara: konsentrasi partikulat (PM2.5). https://www.bmkg.go.id/kualitas-udara/pm25

Badan Pusat Statistik (2023). Proyeksi Penduduk Indonesia 2020–2050, Hasil Sensus Penduduk 2020. Jakarta: BPS-Statistics Indonesia.

Center for International Earth Science Information Network (CIESIN), Columbia University (2018). Gridded Population of the World, Version 4 (GPWv4.11): Population Count. NASA SEDAC.

Dinas Lingkungan Hidup Provinsi DKI Jakarta. Udara Jakarta. https://udara.jakarta.go.id/. Historical ISPU datasets published through Satu Data Jakarta, https://satudata.jakarta.go.id/

Geng G, Xiao Q, Liu S, et al. (2021). Tracking air pollution in China: near real-time PM2.5 retrievals from multisource data fusion. Environmental Science & Technology 55:12106–12115.

GADM Global Administrative Areas (2022). GADM database of Global Administrative Areas, version 4.1.

Kementerian Lingkungan Hidup dan Kehutanan (2020). Peraturan Menteri Lingkungan Hidup dan Kehutanan No. P.14/MENLHK/SETJEN/KUM.1/7/2020 tentang Indeks Standar Pencemar Udara. Jakarta. Real-time index and station data at https://ispu.menlhk.go.id/

NASA Land Processes Distributed Active Archive Center (2023). VIIRS/NPP Thermal Anomalies and Fire Daily L3 Global 1 km SIN Grid, Version 2 (VNP14A1). Sioux Falls: LP DAAC. https://doi.org/10.5067/VIIRS/VNP14A1.002

Data acknowledgements

We are grateful to the Tracking Air Pollution (TAP) team at Tsinghua University for the PM2.5 product on which the exposure estimates are built. Ground measurements come from the monitoring networks of the Indonesian Ministry of Environment and Forestry, the Meteorological Agency (BMKG) and the DKI Jakarta environment agency. Fire detections come from NASA's VIIRS instrument, distributed by the Land Processes Distributed Active Archive Center.