EVIDENCE REPORT · DELHI AIR QUALITY

Delhi's Fine-Particle Pollution, in Full
1980 to July 2026

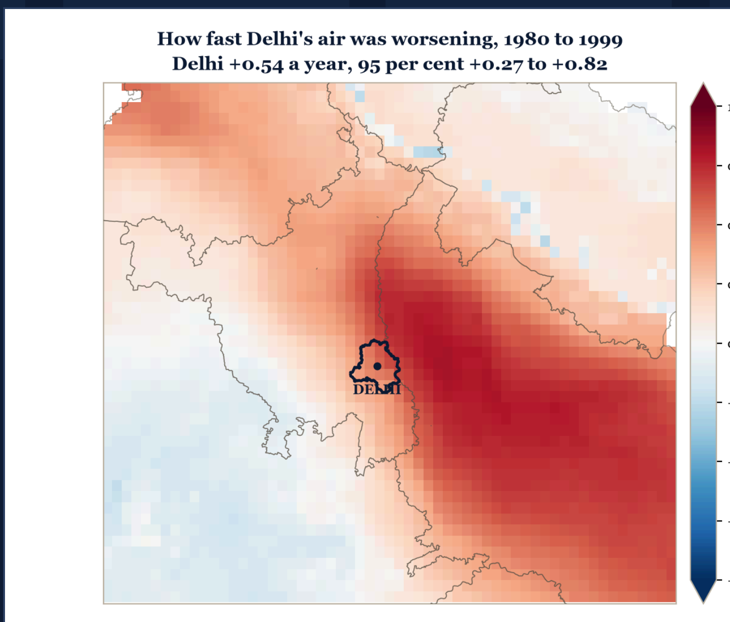
Forty-seven years of the capital's air, reconstructed from satellite and reanalysis to 2022, then carried to the present by Delhi's own monitors, with the weather taken out.

+1.09
micrograms a year added to every winter, weather removed, the same rate on both sides of the 2000 instrument change

0
years in forty-seven at India's own annual standard, and none at the WHO's

17
days out of 3,134 since 2018 that Delhi breathed air the WHO calls safe

3,117
of 3,134 measured days since 2018 breached the WHO twenty-four-hour limit



How fast PM2.5 was rising across North India, 1980 to 1999, with Delhi outlined. A rate within one period, so no instrument change lies inside it. These are the pre-monitor decades, reconstructed rather than measured, as section 11 sets out.

Sources: read 11 August 2026. LongPMInd (Wang et al., ESSD 2024) for 1980 to 2022; CPCB and DPCC monitors via CREA for 2023 to July 2026; ERA5 via the Open-Meteo archive.

Prepared for the **Office of Ajay Maken, MP**

11 August 2026 · analysis in-house

1. What this report does, in plain words

Delhi had no working network of air-quality monitors until well into the 2000s, so a simple question has no direct answer from measurement: how dirty was the city's air before anyone was measuring it, and how does that compare with the air of today? A companion report answered that question to 2022. This one carries the record to July 2026, because the last three and a half years are the ones a reader is actually living in, and because they turn out to settle an argument the earlier record could only open.

Two things had to be built. The reconstruction that reaches back to 1980 stops at December 2022, so the recent years come from Delhi's own monitors, joined to it by a rule chosen on evidence and set out in full in section 9. The weather is then stripped out of the whole forty-seven years at once, since pollution on any day is a contest between what the city emits and what the wind and the mixing depth do with it.

1. **The rise happened once, and nothing since has undone it.** Delhi's fine-particle pollution climbed from about 54 micrograms in every cubic metre of air in the 1980s to about 88 by the 2000s, and it has sat there ever since, at 87 through the 2010s and 88 in this decade.

2. **Eighteen years of that plateau show no direction at all.** The trend since 2010 is +0.18 micrograms a year and cannot be told apart from zero, at $p = 0.66$. Any claim of a recent fall depends entirely on which year the counting starts, which is why section 4 prices every possible start year rather than choosing one.

3. **Delhi has never once met its own law.** In forty-seven years not a single year reached India's annual standard of 40 micrograms, and none came near the World Health Organization's 5. The cleanest year on record, 1983, was still 1.1 times the Indian standard.

4. **Winter is where the record moves, and it moves the wrong way.** While the annual average stood still, the November to January mean has risen +1.62 micrograms a year since 2000, on one observing basis throughout, at $p = 0.028$ and a 95 per cent range of +0.20 to +3.04. That is about 42 micrograms added to every Delhi winter in a quarter of a century, and the winter mean has gone from 148 across the 2000s to 174 across the last ten, a rise of 17 per cent with no seam crossed.

5. **Two faults were found in this office's own working data.** Five months read as exactly zero, four of them in the dirtiest part of 1999, and read naively they would have made 1999 the one year Delhi met its own standard. The published dataset turned out to be sound and two local files damaged, so nothing here is interpolated. Section 10 sets out what was found, what was first concluded and why that was wrong.

Sources: read 11 August 2026. LongPMInd (Wang et al., Earth System Science Data, 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 weather via the Open-Meteo archive. Analysis in-house.

2. The question, and why the answer is not obvious

Fine particles, written PM2.5, are airborne specks smaller than 2.5 micrometres, which makes them about thirty times finer than a human hair. They matter more than any other pollutant for health, because they lodge deep in the lung and then cross into the blood. The limits are strict accordingly. The World Health Organization's 2021 guideline is an annual average no higher than 5 micrograms per cubic metre, with a twenty-four-hour value of 15 that a compliant city may exceed on three or four days a year. Because most of the world is far above that, the same guideline sets four staged interim targets on the way to it, at 35, 25, 15 and 10 as annual means. Delhi is above every one of the four. India's own National Ambient Air Quality Standard is more permissive again, since it allows 40 as an annual average and 60 over twenty-four hours.

Judging whether Delhi is winning or losing needs two things that are easily confused. One is the level, which is what people breathe. The other is the direction, which is what policy can claim credit for. A city can be dangerously polluted while improving, or dangerously polluted while static, and the same annual number is consistent with both. Separating them needs a long record, because a short one measures the weather, and it needs the weather removed, because an unusually still winter can imitate a policy failure exactly.

Sources: read 11 August 2026. World Health Organization Global Air Quality Guidelines, 2021; India's National Ambient Air Quality Standards, notified 2009. LongPMInd (Wang et al., Earth System Science Data, 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 weather via the Open-Meteo archive. Analysis in-house.

3. The long arc, 1980 to July 2026

Across forty-seven years the shape is a staircase with one step in it. Delhi sat near 54 micrograms through the 1980s, which was already many times the WHO guideline, drifted up through the 1990s, then climbed to a plateau near 88 which it has not left since.

Those figures come from the reconstruction, and the reconstruction is not the harshest witness available. On Delhi's own monitors, run by the government, the city breathed 104.3 micrograms in 2024. That is twenty-one times the World Health Organization's guideline and 2.6 times India's own legal limit, with no model anywhere in the chain. The winter that ended in 2026 averaged 200.5, which is 40 times the guideline. The reconstruction used for the long history reads a few micrograms lower, because it averages a ten-kilometre square rather than a set of street addresses, and both figures are printed here so the reader can see the gap rather than be told about it.

Sources: read 11 August 2026. CPCB and DPCC continuous monitors via CREA, fixed panel of continuously reporting CPCB and DPCC stations; the exposure day-counts above use every reporting station; 366 days in 2024. WHO Global Air Quality Guidelines 2021 annual guideline 5; India NAAQS annual 40. Analysis in-house.

That is one city's average, and a reader is entitled to ask whether it is one bad sensor or one bad neighbourhood. It is neither. Across a window reaching from Delhi into Haryana and Uttar Pradesh, 37 government reference monitors held at least three years of daily readings between 2019 and 2025, 10 of them outside Delhi's jurisdiction altogether. **Not one of the 37 met India's own annual standard, and not one met the World Health Organization's.** The cleanest of them all is not in Delhi at all: it is at Hapur, and it reads 64.6 micrograms, which is 1.6 times India's limit and 13 times the WHO's. This is a census of the instruments, not a sample from them, so it needs no model and no significance test: the boundary of Delhi is not what contains this problem.

Sources: read 11 August 2026. CPCB, DPCC, UPPCB, HSPCB and RSPCB continuous monitors via OpenAQ, daily values 2019 to 2025, reference monitors only, each with at least three years of data. Analysis in-house.

How much of that climb is Delhi's air and how much is the instrument that changed in 2000 cannot be separated on this data, so this report will not quote a single figure across the join. The satellite record the reconstruction is built on gained a major instrument in March 2000, and the series steps 15.8 micrograms between 1999 and 2000 alone, which is more than any real year-to-year rise in the record. Fitting one straight line through all forty-six years gives +1.06 micrograms a year; allowing that step its own term drops the same line to +0.23. So the two periods are reported separately, each on one observing basis throughout: **+0.51 a year through 1980 to 1999**, with a 95 per cent range of +0.21 to +0.81, and **+0.11 a year from 2000 to 2025**, range -0.27 to +0.48, which includes no change at all. Delhi rose, and it never came down. What this record cannot support is a precise multiple for how far.

Era	As measured or reconstructed	Weather removed	Times India's 40, as measured	Times the WHO's 5	Trend within the era
The 1980s	51	54	1.3	10.1	+0.78 a year, p = 0.07
The 1990s	56	59	1.4	11.1	+0.26 a year, p = 0.58
The 2000s	88	88	2.2	17.5	+1.45 a year, p = 0.04
The 2010s	90	87	2.3	18.0	+0.73 a year, p = 0.42
The 2020s, to 2025	92	88	2.3	18.4	-1.01 a year, p = 0.60

READING No era since 2000 moves by more than two micrograms from the one before it, which is the plateau stated as a table: whatever happened to Delhi's air happened before this century's second decade began.

CAUTION The step between the 1990s and the 2000s spans the 2000 instrument change described above, so part of it is the satellite record and not Delhi's air. Compare within a basis, never across that boundary.

SOURCE Sources: read 11 August 2026. De-weathered and as-measured annual means over full calendar years. LongPMInd (Wang et al., *Earth System Science Data*, 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 weather via the Open-Meteo archive. Analysis in-house.

THE POINT **Delhi has spent every one of the last two and a half decades between seventeen and nineteen times the level the World Health Organization calls safe.**

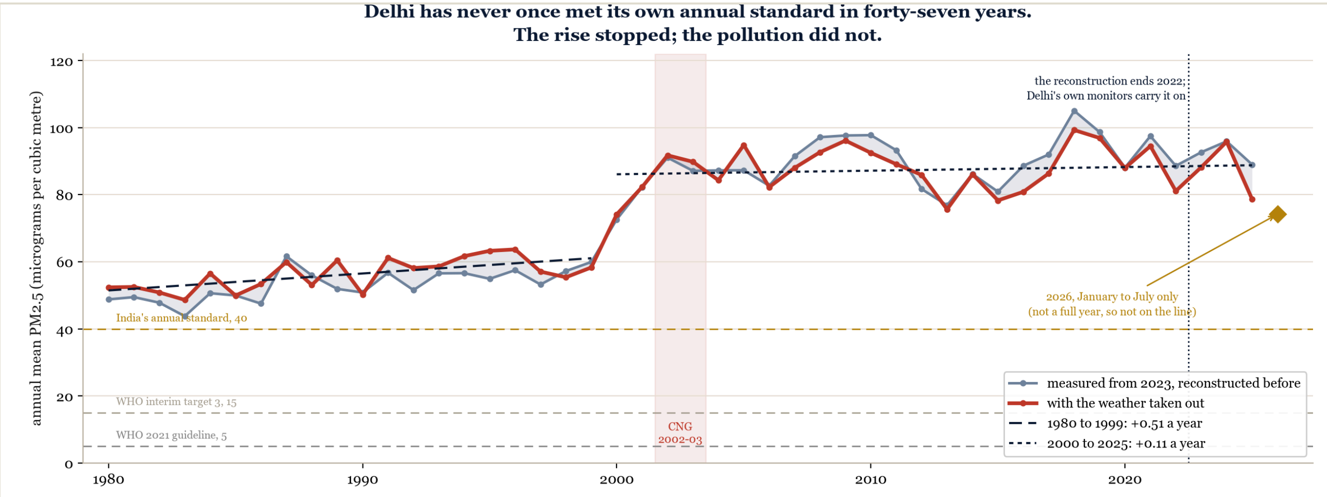


__Figure 1.__ Delhi annual mean PM2.5, 1980 to July 2026. Slate is the record as measured, or as reconstructed for the years before the monitors. Red is the same record with the weather removed. The dotted vertical marks where the reconstruction ends and Delhi's own monitors take over. The two dashed fits are drawn inside one measurement basis each and never across the 2000 instrument change: +0.51 a year to 1999, and +0.11 a year after it, a rate whose range includes no change at all.

READING The two lines run almost together, which is the visual proof that the rise is emissions and not weather, since removing the weather barely moves the path. The second fit is flat, and a flat line here is not a recovery: it is a quarter of a century spent at about eighteen times the WHO guideline, and section 6 shows the winters worsening throughout it.

CAUTION The 2026 diamond is January to July only and is deliberately kept off the line, because a part year that excludes November and December cannot be compared with a full one. Section 7 places it properly.

SOURCE Sources: read 11 August 2026. LongPMInd (Wang et al., *Earth System Science Data*, 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 weather via the Open-Meteo archive. Analysis in-house.

THE POINT **The staircase has one step and then a landing that is now a quarter of a century long, held far above every limit in force.**

Set out by decade the same story is starker, because the averaging removes the year-to-year weather noise that makes a single year arguable. The 1980s and the 1990s belong to one regime and everything from 2000 belongs to another, and the two decades since have added nothing and taken nothing away.

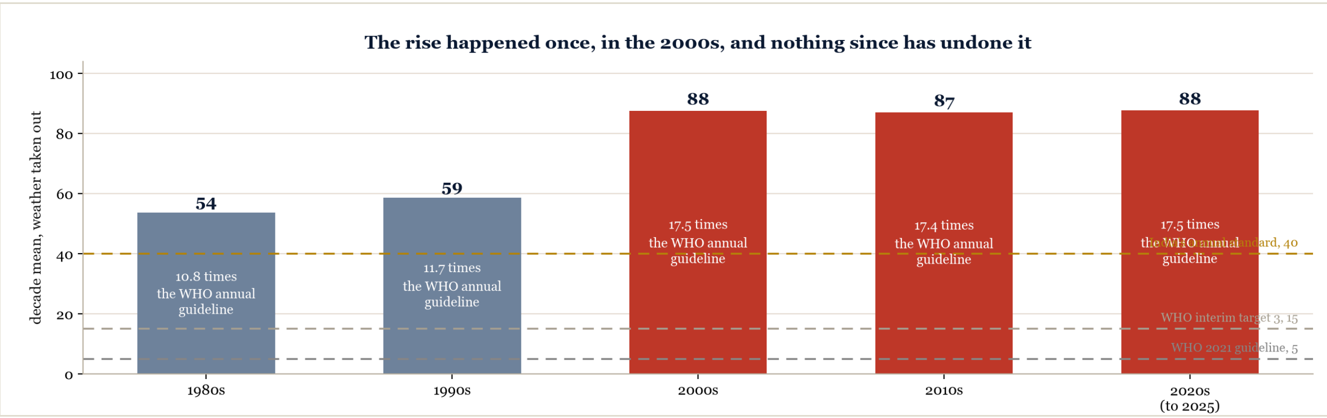


__Figure 2.__ Decade means of de-weathered PM2.5, with each decade expressed as a multiple of the WHO guideline of 5 micrograms. The 2020s bar covers 2020 to 2025, the full years available.

READING Three decades at about the same height is the finding, because the 2000s, the 2010s and the 2020s differ by less than two micrograms, which is well inside the year-to-year noise.

CAUTION The gap between the 1990s bar and the 2000s bar spans the 2000 instrument change, so part of that one step is the satellite record rather than Delhi's air. Every later comparison sits on one basis and is unaffected.

SOURCE Sources: read 11 August 2026. As Figure 1, de-weathered means over full calendar years only. LongPMInd (Wang et al., *Earth System Science Data*, 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 weather via the Open-Meteo archive. Analysis in-house.

THE POINT **Delhi's air got worse once, in a single decade, and it has been held there for a quarter of a century since, although three governments have legislated against it.**

A map of levels is safe only when each panel holds one period on one measurement basis, because then no panel contains a comparison and no arithmetic crosses the instrument change of 2000. Cut that way the record separates into four regimes, each with a different claim on the reader's confidence, and each of them far above every limit in force. The fifth panel restates the same four levels as multiples of the World Health Organization's guideline, which is the unit a reader can act on, and it keeps the two measurement bases apart rather than running them together into a single rising line.

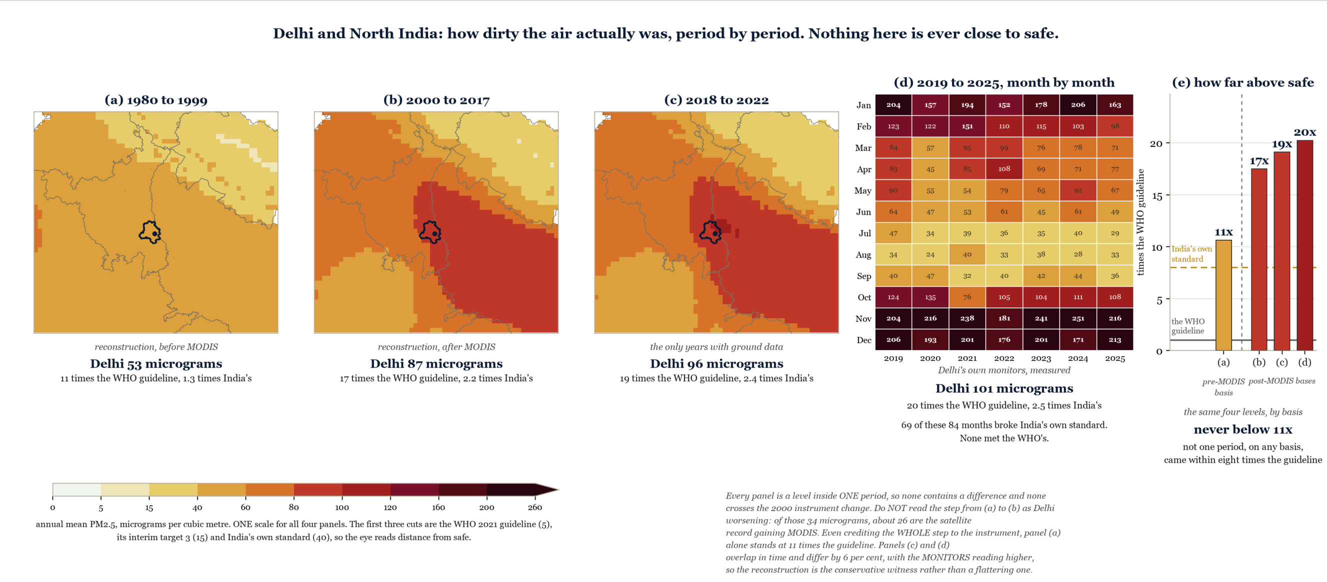


Figure 3. *Annual mean PM2.5 across Delhi and North India, one panel per period, with Delhi outlined. Panels (a) to (c) are the reconstruction, each a level inside a single period on a single basis. Panel (d) is Delhi's own monitors month by month, measured, because no gridded product exists after 2022. Panel (e) restates the four levels as multiples of the WHO guideline, with the two measurement bases kept apart. One colour scale governs every panel, and it is cut at the limits rather than at round numbers.*

READING Nothing here is ever close to safe. The four periods read 11, 17, 19 and 20 times the WHO guideline, so the record never drops below eleven times it on any basis, and the whole plain is dark in every panel rather than Delhi alone.

CAUTION Do not read the step from (a) to (b) as Delhi worsening. It is 34 micrograms, of which about 26 is the satellite record gaining MODIS in 2000. Panels (c) and (d) overlap in time and differ by about six per cent, with the monitors reading HIGHER, so the reconstruction is the conservative witness rather than a flattering one.

SOURCE Sources: read 11 August 2026. Per-cell means within each period on the LongPMInd grid for (a) to (c); CPCB and DPCC monitors via CREA for (d) and (e). LongPMInd (Wang et al., *Earth System Science Data*, 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 weather via the Open-Meteo archive. Analysis in-house.

THE POINT **Even crediting the whole year-2000 step to the instrument, the least-instrumented period in the record still stands at eleven times the guideline.**

4. The plateau has not broken

The extension to July 2026 exists to answer one question, and the answer is no. Fitting a trend to the de-weathered record since 2010 gives +0.18 micrograms a year, with a 95 per cent range of -0.67 to +1.03 that comfortably contains zero. Since 2015 the figure is +0.26, range -1.48 to +2.00. Since 2018 it turns negative at -1.93, range -4.34 to +0.48, which again spans no change, and it is negative for a reason worth naming.

2018 was the dirtiest year in the record. A trend measured from a peak slopes downward whether or not anything real has changed, which is the oldest trap in trend reporting and the easiest for any party in office to fall into honestly. It matters here because the dataset this report is built on says exactly that in its own abstract, reproduced in Annexure A: that 2018 was a turning point, that the National Clean Air Programme was launched that year, and that concentrations fell across most of India between 2018 and 2022. Three things are worth holding together. That decline is not claimed as statistically significant by its authors, and attaches no p-value, where the same sentence marks the pre-2018 rise as significant. For Delhi specifically it rests on five points and does not reach significance. And it ended when the data did: on Delhi's own monitors 2023 and 2024 both read above 2022. The remedy is not to argue about the right start year. It is to price every start year and print the lot, which is the right-hand panel below.

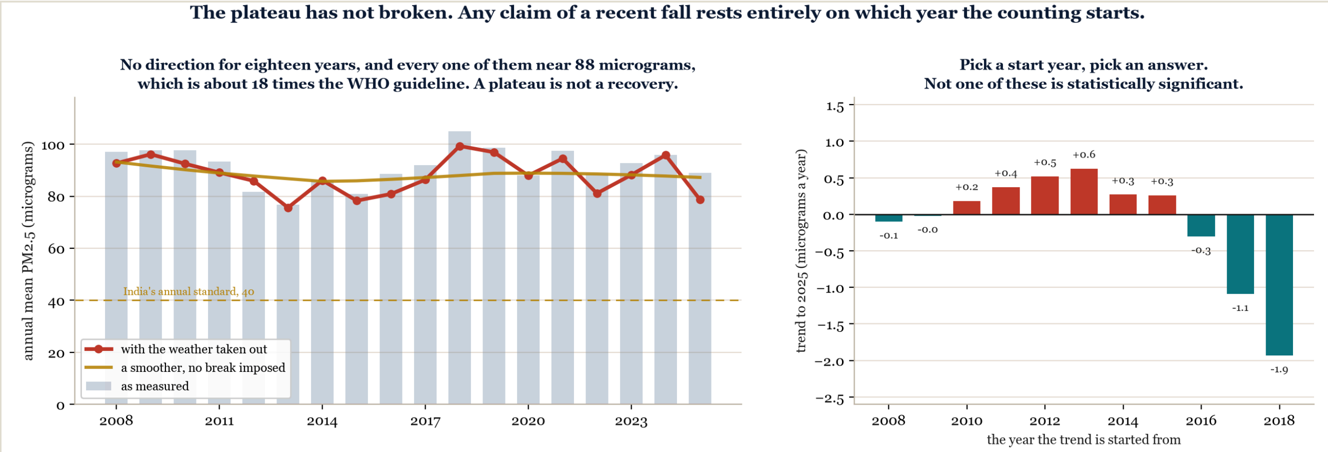


Figure 4. *Left, the last eighteen years as measured and de-weathered, with a smoother that imposes no breakpoint. Right, the trend to 2025 computed from every start year between 2008 and 2018, one bar each.*

READING The sign of the recent trend is a choice, not a finding. Start in 2013 and Delhi is worsening at +0.6 micrograms a year; start in 2018 and it is improving at -1.9; and not one of the eleven bars is distinguishable from zero.

CAUTION Not one of these eleven slopes is distinguishable from zero, so none should be quoted on its own in either direction. They are a sensitivity display rather than a set of claims, which is why they carry p-values and not intervals.

SOURCE Sources: read 11 August 2026. De-weathered annual means, 1980 to 2025, ordinary least squares on the year. LongPMInd (Wang et al., *Earth System Science Data*, 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 weather via the Open-Meteo archive. Analysis in-house.

THE POINT **Anyone claiming Delhi's air is now improving, or now worsening, is reporting their choice of start year rather than the city's air.**

A finding this consequential should not rest on one team's arithmetic, so it was put to a second source. MERRA-2 is NASA's atmospheric reanalysis, produced by a different institution on a different method, and it runs without a break from 1980 to the present. It is not wholly independent of the reconstruction, because the reconstruction takes MERRA-2 composition among its inputs, and that shared parentage is exactly why the two agree on some things and not others. It also reads Delhi far too low in absolute terms, which is a known property of a global model over a dense urban source area, so it is used here for direction alone.

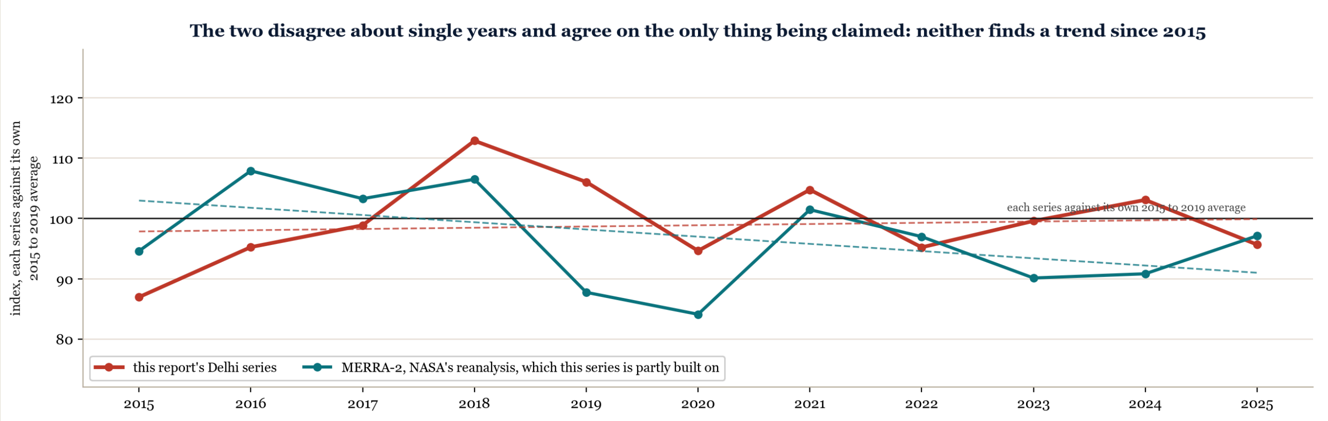


Figure 5. *This report's Delhi series against NASA's MERRA-2 reanalysis, each indexed to its own 2015 to 2019 average so that the two can be compared on direction alone. Dashed lines are the trend through each.*

READING The two lines disagree sharply about individual years, correlating at only 0.20, and they agree on the only claim being made here, which is that neither shows a trend since 2015.

CAUTION A weak year-to-year correlation is a real limitation and is printed rather than hidden. It means the cross-check supports the verdict on the trend and supports nothing at all about any single year.

SOURCE Sources: read 11 August 2026. MERRA-2 M2T1NXAER surface mass, assembled to PM2.5 the standard way, over a 25 km radius of central Delhi, via Google Earth Engine, against this report's own series. Analysis in-house.

THE POINT **Two institutions working on two methods reach one verdict on the question asked, which is that the plateau has not broken.**

5. Winter, where the burden actually falls

A flat annual average is not the same as a still city, and the seasonal record shows why. Delhi's pollution has always been seasonal, because the cold months bring a shallow mixing layer and slack winds that trap emissions near the ground, while the monsoon washes the air and stirs it deep. What has changed is the size of the gap. November now averages 191 micrograms and August about 34, a ratio of 5.6 to one.

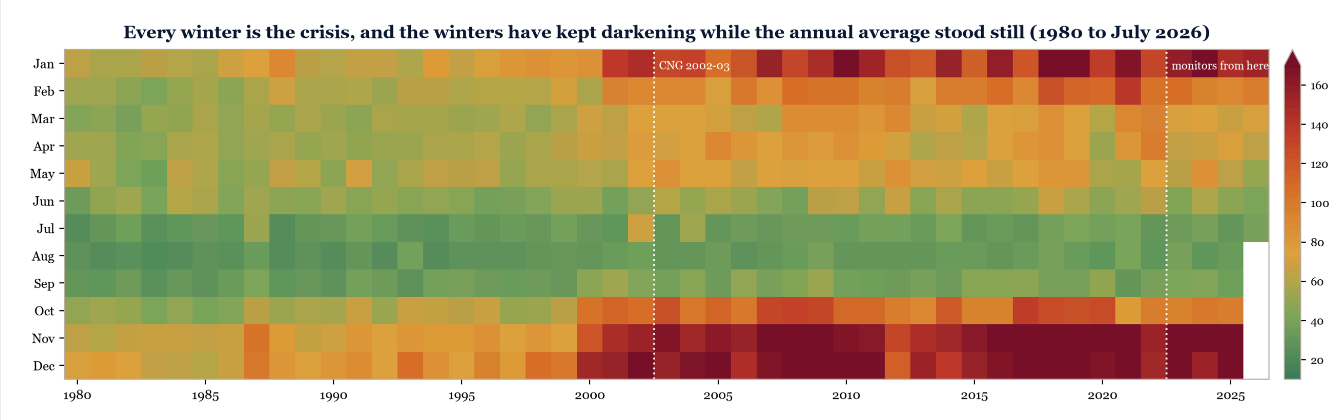


Figure 6. *Every month of every year, 1980 to July 2026. Each column is a year and each row a calendar month, with green cleaner and dark red dirtier. The last four cells of 2026 are blank because the year has not reached them.*

READING The reddening runs along the bottom rows and not the middle ones, which is the whole seasonal argument in one image: the winters darkened while the monsoon months stayed where they were.

CAUTION The colour scale is common to all forty-seven years, so that a cell which looks similar across two decades genuinely is similar. It is not rescaled by decade, which would have flattered the recent years.

SOURCE Sources: read 11 August 2026. As Figure 1, monthly means. LongPMInd (Wang et al., *Earth System Science Data*, 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 weather via the Open-Meteo archive. Analysis in-house.

THE POINT **Delhi does not have an air pollution problem spread evenly across the year, because what it has is a winter emergency sitting on a baseline that is unsafe all year round.**

Separating the two seasons makes the divergence measurable. Across the record the winter mean has risen +1.62 micrograms a year since 2000 as measured, range +0.20 to +3.04, and +1.09 once the weather is removed. Before 2000, on the other observing basis, the de-weathered rate was +1.09, which is the same figure to two decimals, so the winter has risen at a steady pace across the whole record and the pace does not depend on the instrument, while the monsoon months move by only +0.2 and end the record at about 37 micrograms. Since 2010 the winter trend is +0.9 a year, which does not reach significance at p = 0.34 and is worth watching, because it is the only part of the record still pointing anywhere.

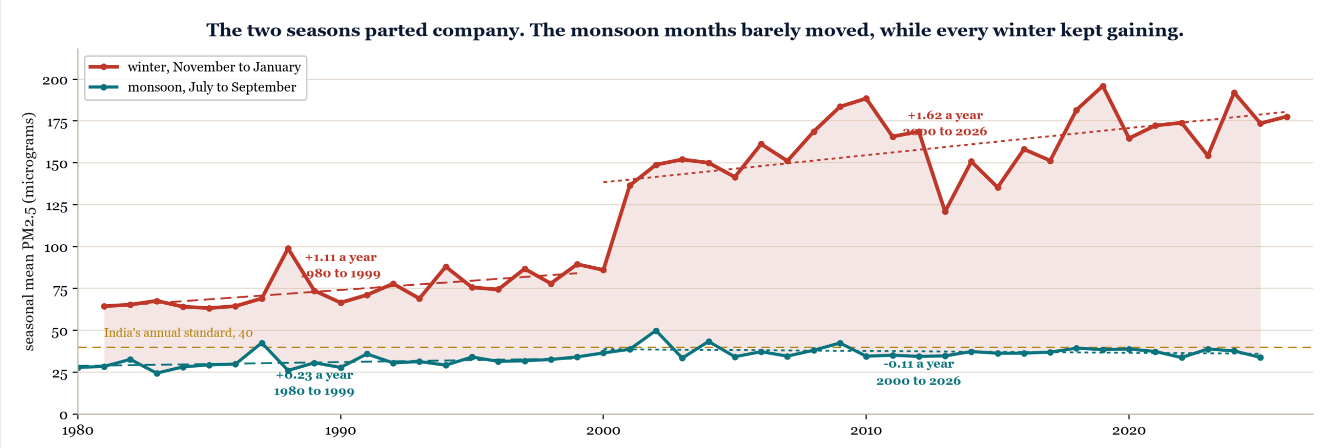


Figure 7. *The November to January mean against the July to September mean, each year from 1980, both as measured. Dashed lines are the least-squares trend through each, so they carry the as-measured slopes rather than the de-weathered ones quoted above. December is counted with the winter it belongs to rather than the calendar year it falls in. Each dashed line is fitted inside one measurement basis and never across the 2000 instrument change, so the rate before that year and the rate after it are quoted separately.*

READING The two seasons parted company around 2000 and have not come back together, since the monsoon months have spent the whole record close to India's annual standard while the winter months are now four times above it.

CAUTION Winter carries the widest uncertainty in any satellite-era reconstruction, because thick haze is exactly the condition a satellite reads worst. The direction is robust; the precise winter level for the pre-monitor years is not.

SOURCE Sources: read 11 August 2026. As Figure 1, seasonal means over complete seasons only. LongPMInd (Wang et al., *Earth System Science Data*, 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 weather via the Open-Meteo archive. Analysis in-house.

THE POINT **Any policy that moves the annual average while leaving November where it is has not touched the part of the year that does the damage.**

6. What Delhi actually breathed

Averages understate a pollution problem, because health damage follows the days rather than the mean. Since 2018, when the monitor network became thick enough to describe the city honestly, Delhi has recorded 3,134 days. Of those, 99.5 per cent breached the WHO twenty-four-hour guideline and 62.1 per cent breached India's own standard of 60. Seventeen days in eight and a half years were clean by the WHO's measure.

Sources: read 11 August 2026. CPCB and DPCC continuous monitors via CREA, 1 January 2018 to 31 July 2026. Analysis in-house.

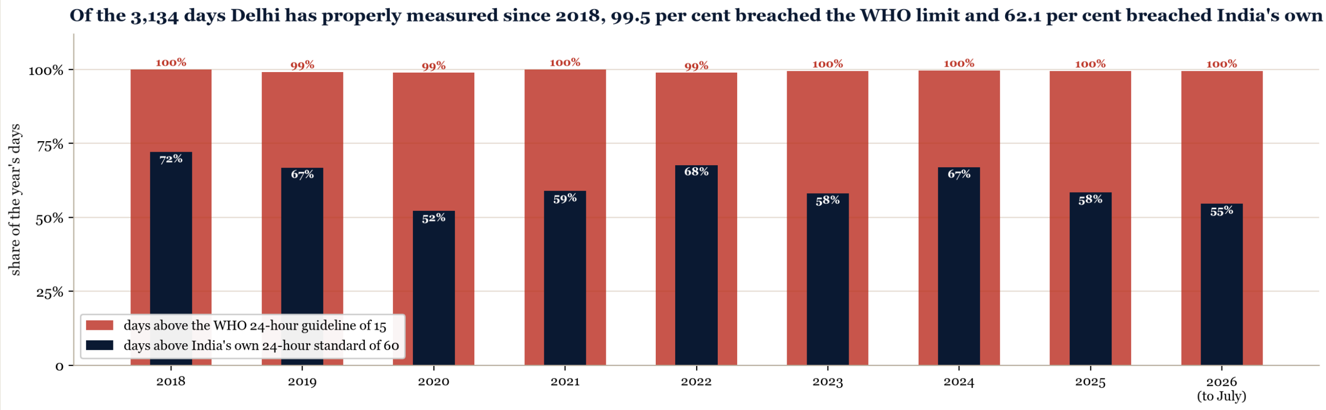


Figure 8. *The share of each year's days above the WHO twenty-four-hour guideline of 15 micrograms and above India's own twenty-four-hour standard of 60. 2026 covers January to July.*

READING The red bars are effectively full every year, so that the WHO measure no longer distinguishes one Delhi year from another. Only the navy bars move, and they move within a band running from 52.2 to 72.1 per cent with no direction to it.

CAUTION Years before 2018 are excluded, not because they were clean but because Delhi was barely measuring them, with a median of four to seven reporting stations a day against about thirty-six now. A breach count over a handful of sites describes the sites and not the city.

SOURCE Sources: read 11 August 2026. CPCB and DPCC continuous monitors via CREA, 2018 to 31 July 2026. A station-day needs at least sixteen valid hours, and the city day is the mean of qualifying stations. Analysis in-house.

THE POINT **The worst day in the window reached 593 micrograms, in 2024, which is nearly forty times the level the WHO calls safe for a single day.**

Year	Days measured	Above the WHO 15	Clean days	Above India's 60	Worst day
2018	365	365	0	263	402
2019	365	362	3	244	553
2020	366	362	4	191	522
2021	365	365	0	215	417
2022	365	361	4	247	400
2023	365	363	2	212	376
2024	366	365	1	245	593
2025	365	363	2	213	384
2026 (to July)	212	211	1	116	337
All years	**3,134**	**3,117**	**17**	**1,946**	**593**

READING Seventeen clean days in eight and a half years is the row that matters. The count of days above India's own standard has moved between 52.2 and 72.1 per cent without settling anywhere.

CAUTION A city daily mean is an average across about thirty-six monitors, so a resident beside a busy road or a construction site breathed worse than every figure printed here, and a resident in a park breathed better.

SOURCE Sources: read 11 August 2026. CPCB and DPCC continuous monitors via CREA, 2018 to 31 July 2026, city daily means under the sixteen-hour completeness rule. Analysis in-house.

THE POINT **Clean air in Delhi is not rare. On the WHO's measure it has effectively stopped occurring.**

7. How to read a year that is only seven months old

2026 is incomplete, and an incomplete Delhi year always looks clean, because the months it is missing are November and December. Comparing seven months against twelve is therefore not a small imprecision but a guaranteed error in one direction. The only lawful comparison is like for like, which means January to July of 2026 against January to July of every year before it.

On that basis 2026 reads 74.2 micrograms, against 81.8 for the same seven months averaged over 2019 to 2025, a difference of -9.3 per cent. It ranks 29th cleanest of 47, so it sits in the middle of the record rather than at either end. July alone reads 39.4 against a recent July average of 36.9, so that the most recent month in hand sits above its own benchmark rather than below it.

Sources: read 11 August 2026. As Figure 1, January to July means only. LongPMInd (Wang et al., Earth System Science Data, 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 weather via the Open-Meteo archive. Analysis in-house.

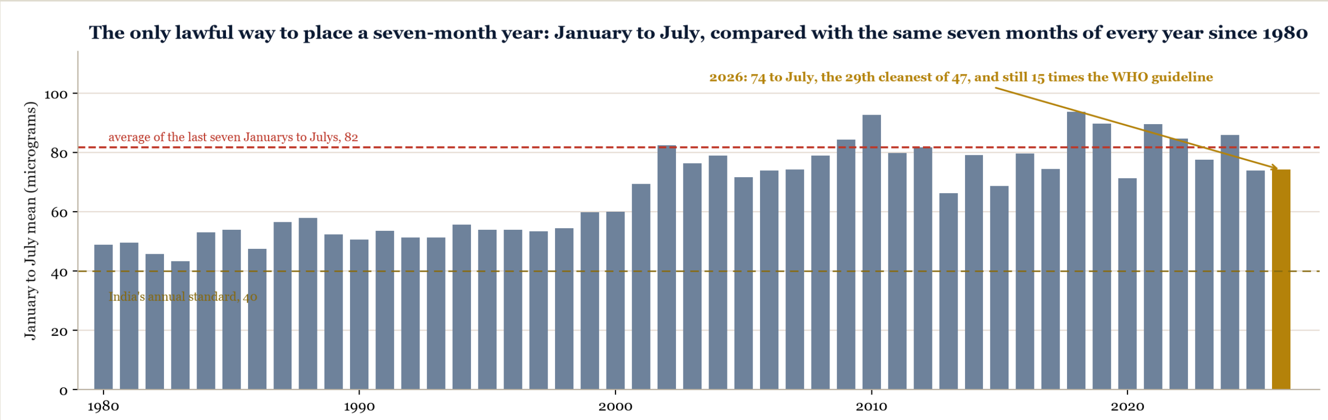


Figure 9. *The January to July mean of every year from 1980, with 2026 in gold. The red line is the average of the same seven months across 2019 to 2025.*

READING 2026 sits below the recent average and near the middle of the record, 29th cleanest of 47, which is what a normal year on a high plateau looks like when it is placed properly, rather than what a turning point looks like.

CAUTION Seven months of one year cannot establish a trend under any circumstances, which is why this chart exists to prevent a false reading of 2026 rather than to supply a confident one.

SOURCE Sources: read 11 August 2026. As Figure 1, January to July means only. LongPMInd (Wang et al., *Earth System Science Data*, 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 weather via the Open-Meteo archive. Analysis in-house.

THE POINT **2026 so far is an ordinary year on the plateau. It is neither the breakthrough nor the collapse that a part-year number can be made to suggest.**

8. The 2002-03 CNG switch, and what this record can and cannot say about it

The most celebrated clean-air measure in Delhi's history belongs to this period, and it is worth setting out what this record can and cannot say about it. Between 1998 and 2003 the Supreme Court, which was hearing the *M. C. Mehta* litigation on the advice of the Bhure Lal Committee, ordered Delhi's entire public transport fleet onto compressed natural gas. A court can order a switch, although it cannot convert a bus. The conversion itself, which meant phasing out some 6,300 diesel buses through operator strikes and street protests, had to be delivered by the elected Delhi government under Chief Minister Sheila Dikshit and her Transport Minister, Ajay Maken.

> The conversion was the contested part, and the contemporaneous record is clear about who carried it. In the Centre for Science and Environment's account, *The Leapfrog Factor* (2003), Delhi's Transport Minister **Ajay Maken**, who "joined the government when the city was burning with riots over the protests against CNG", "helped make the difference by **implementing the court order and becoming a champion of clean air and public transport.**" At the height of the resistance, in April 2002, he held the line: *"We are not blind. We are concerned about the health of citizens, and there is no doubt that CNG is cleaner than diesel. We want to end this controversy once and for all by implementing the order and phasing out diesel buses."*

The air people could feel did improve, and the documented gains are in the gases rather than the particles. Carbon monoxide roughly halved between 1998 and 2003, sulphur dioxide fell about a third, lead fell about three quarters from unleaded petrol, and the black smoke of diesel buses disappeared, because CNG burns without visible soot. Those are measured at ground level, by instruments that were running at the time, and they are not in question.

A CLAIM THIS REPORT HAS WITHDRAWN

An earlier draft said the particles carried a matching fingerprint: that the reconstructed PM2.5 rise slowed sharply at the switch, and that black carbon did the same. A four-triad review on 11 August 2026 established that the bend is an artefact of the satellite record and not of Delhi's air, and it is withdrawn here rather than defended.

MERRA-2, whose aerosol fields feed the reconstruction, began assimilating the MODIS Terra instrument in March 2000, and its fields step and then ramp until Aqua joins in mid-2002. A ramp that ends in 2002 IS a steep rise followed by a flat one, with a bend at 2002-03, whether or not anything happened on the ground. Run the same test at all 27,708 land cells of the dataset and **79 per cent of India shows the same bend**. Lucknow and Kanpur show it harder than Delhi. So does the Thar desert, which has no buses. Measured against Lucknow, Kanpur and Patna, which share Delhi's airshed and had no such conversion, Delhi's own effect has the wrong sign at $p = 0.29$. The bend belongs to the satellite record and to a change the whole subcontinent shows, not to Delhi. Where any part of it is real, it is regional, and it is not a Delhi policy effect.

The dataset's own authors record the same thing in the paper reproduced here as Annexure A: "The PM concentrations jumped in 2000, which can be attributed to the absence of satellite data for MERRA-2 before 2000." This report bound that paper in and did not read its warning closely enough until a review found it.

Sources: read 11 August 2026. Wang et al., Earth System Science Data 16, 3565 to 3577, reproduced as Annexure A, quoted from its section 4.3, Spatial and temporal trends; MERRA-2 via Google Earth Engine; all-cell screen over the LongPMInd grid. Analysis in-house.

One honest consequence deserves stating, because it cuts the other way. With this record's year-to-year scatter, the smallest step it could reliably detect is about 22 micrograms, which is more than a quarter of the period mean. Converting 6,300 buses in a city of this size could never have moved the citywide average that far. **The record was never able to see CNG's true effect, so its silence is not evidence that the reform failed.** What is withdrawn is a corroboration that was never load-bearing, not the achievement it was offered in support of.

Why the air nonetheless stayed dirty is structural rather than a failure of the reform. Vehicle exhaust from the public fleet was never the majority of Delhi's particle mass, since the source studies give roughly co-equal shares to vehicles, to road and construction dust, and to industry, on top of household burning, open waste burning and pollution blown in from the wider region. Converting buses, autos and taxis touched one slice of one of those shares, while Delhi's total vehicle count roughly doubled between 2002 and 2012.

9. How the two records were joined, and why the join holds

The reconstruction and the monitors measure different things, which is why they cannot simply be laid end to end. LongPMInd reports the average over a grid square about ten kilometres across, while the monitors sit at about forty particular addresses, most of which were sited where people and traffic are. The monitors accordingly read higher. Over the 60 months where both records exist and the network is thick, from 2018 to 2022, they correlate at 0.996, so that the relationship is tight and the only open question is what form it takes.

Four candidate rules were fitted on 2018 to 2020 and then judged on 2021 and 2022, which the fitting never saw. The rule with the smallest error on those held-back years was kept, and the losers are printed below so the choice is auditable rather than announced.

Rule tried	Error on the held-back years	Bias	Share of variation	What it assumes
No correction at all	13.42	+8.60	0.931	Reads the monitors as though they were the grid
One ratio	6.25	+0.30	0.985	A single level factor for all months
A ratio for each calendar month	5.42	+0.28	0.989	Twelve factors, one per month
A level and a slope	**4.73**	**+0.50**	**0.991**	**Two numbers, fitted by least squares**

READING Doing nothing is much the worst option, at an error of 13.4 micrograms against 4.7 for the rule chosen, which is why the two records are not simply laid end to end.

CAUTION All four rules leave a positive bias, so all four read the recent years slightly high against what the reconstruction would have said. The chosen rule's bias is +0.50 micrograms, small rather than absent.

SOURCE Sources: read 11 August 2026. Fitted on 2018 to 2020, judged on 2021 to 2022, then re-fitted on 2018 to 2022 before use. LongPMInd (Wang et al., *Earth System Science Data*, 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 weather via the Open-Meteo archive. Analysis in-house.

THE POINT **The join was chosen by a test it could fail, on years withheld from the fitting, rather than by preference.**

The join was then priced in two further ways. Anchoring the rule two years earlier, on 2019 and 2020 alone, and carrying the monitors forward to 2022 lands +2.2 micrograms from what the reconstruction itself says for that year, or +2.5 per cent over two years, which is the honest uncertainty attaching to a carry of this kind. Separately, the whole extension was rebuilt on a fixed panel of 35 stations reporting throughout 2021 to 2026, so that nothing in the recent movement could be an artefact of which monitors happened to be switched on.

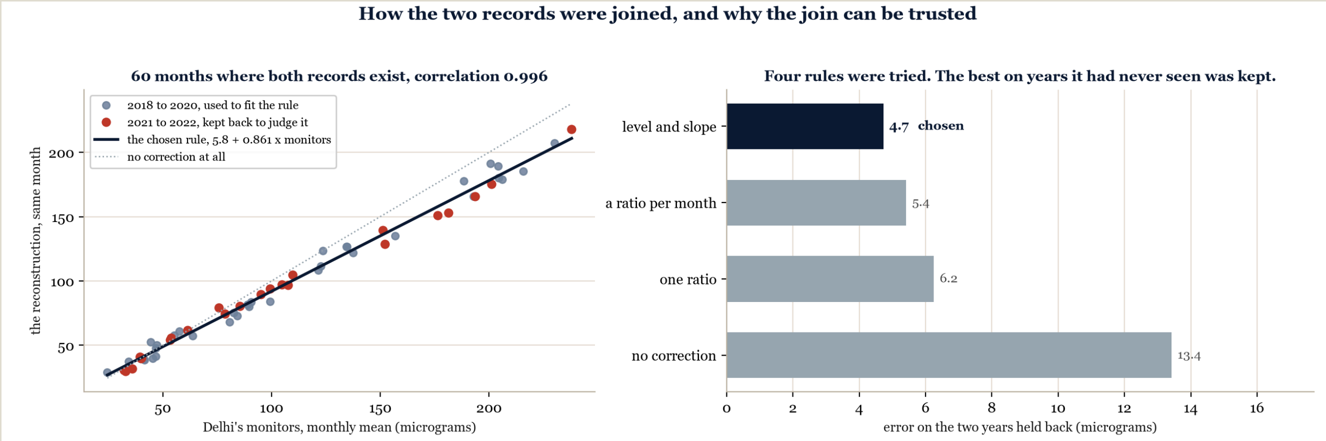


Figure 10. *Left, every month where both records exist, with the months used to fit the rule in slate and the months held back to judge it in red. Right, the error each candidate rule made on those held-back years.*

READING The points sit tightly on a line, and the held-back months fall on it as neatly as the fitted ones, which is what a join looks like when it will survive being extrapolated.

CAUTION A rule validated over two years is being asked to carry three and a half. The drift test above is the reason a 2.5 per cent band belongs on the recent levels, and it is why no claim here rests on a difference smaller than that.

SOURCE Sources: read 11 August 2026. LongPMInd monthly against the CPCB monitor monthly mean for Delhi, 2018 to 2022. Analysis in-house.

THE POINT **The recent years are joined to the history by a rule that was tested before it was trusted, and the size of what could still be wrong is stated rather than assumed, because a join that hides its own error is worth nothing.**

10. Two faults found in the working data, and what was done

Five months that read as zero, and were not missing at all

The Delhi series this office had been working from carried a value of exactly zero for five months: September to December 1999, and December 2006. Zero micrograms per cubic metre does not occur anywhere on earth, still less in Delhi in November. The damage was real, because four of the five are the dirtiest months of their year: the 1999 annual mean came out at 37.6, which sits below India's annual standard of 40, and would have made 1999 the single year in forty-seven that Delhi met its own law. It did not.

The obvious repair was to interpolate the missing months from the years either side, and that repair was made and then withdrawn, because opening the original rasters showed there was nothing missing to repair. **The published dataset is sound.** Two local files were damaged, in two different ways, and both damages announce themselves as zeros rather than as errors. The 1999 raster holds real values for all twelve months, from 32.0 in September to 99.1 in December, and the zeros were manufactured here by an extraction that ran while that file was still downloading. The 2006 raster was itself a truncated download, 6,301,368 bytes against the 6,727,632 the repository publishes, and December 2006 really was blank inside it. Re-fetched and checked against the published checksum, it holds 144.2.

So no value in this report is interpolated, invented or modelled: every month of the reconstruction is the number the published dataset holds. The 1999 annual mean is 59.9 rather than 37.6, and 2006 is 82.7 rather than 70.6. The extraction now refuses to run at all if a raster is the wrong size, and refuses to publish a series containing an impossible value, so a half-written file cannot reach the analysis quietly a second time.

Six months with no mixing height

The ERA5 weather archive carries no boundary-layer height for Delhi from 2 January to 30 June 2024, a gap of 181 days confirmed as missing at source rather than lost in transit. Mixing height is the most important single variable in a Delhi de-weathering, since it sets how deep an air column the day's emissions are stirred into, so dropping it would have cost the analysis five months of 2024 and therefore the whole year. The gap was filled from MERRA-2's independent boundary-layer height, put on the ERA5 scale by regression over the surrounding months, which correlate at 0.73. Dropping those months entirely instead moves the trend since 2010 from +0.18 to +0.01 micrograms a year, so no conclusion here depends on the repair.

Squaring this report with its companion

The companion report to 2022 sits beside this one and was rebuilt from the same repaired series on the same day, so the two now share their history rather than merely resembling it. Across the 43 years both cover, the typical difference between them is 0.6 of a microgram and the largest is 2.4. Nothing in either document turns on any of it.

What little movement remains has one cause, worth naming rather than smoothing over. The de-weathering is fitted across whatever record it is given, so extending this one by three and a half years slightly changes the monthly weather averages every year is measured against, and every de-weathered year shifts a fraction in consequence. The share of variation attributed to weather moves from 11.7 to 12.5 per cent for the same reason. The decade levels barely notice: the 1980s read 53 there against 54 here, and 2013 to 2022 reads 86 against 87.

Sources: read 11 August 2026. LongPMInd (Wang et al., Earth System Science Data, 2024), Delhi grid cell; ERA5 via the Open-Meteo archive; MERRA-2 M2T1NXFLX via Google Earth Engine. Companion report of 25 June 2026, Four Decades of Delhi's Fine-Particle Pollution, 1980 to 2022. Analysis in-house.

11. Limitations, stated openly

- The years before 2000 are an extrapolation, since there were no ground monitors and no MODIS satellite anchor, and the model behind them was trained only on 2018 to 2022. A level step at 2000 of the size measured here is, over forty-six years, itself a trend, so it is the QUANTUM that is uncertain rather than only the early level. The two periods either side of that step are quoted separately throughout, and each is sound on its own basis.
- The reconstruction reports a ten-kilometre area average and the monitors report about forty addresses, which is why the join in section 9 exists and why a residual level difference of a few micrograms is possible in either direction.
- Winter carries the widest error of any season, because thick haze is the condition a satellite reads worst, so the winter finding should be read as a direction rather than as a precise slope.
- Nothing in the forty-seven years is interpolated, modelled or filled by this office: every reconstruction month is the value the published dataset holds, and the one repair made anywhere is the 181 days of mixing height in section 10.

- The de-weathering removes the weather-explained part of each month, which is about 12 per cent of the month-to-month variation. It does not remove long-run changes in Delhi's climate itself, so a slow local trend towards stiller winters would still sit inside the de-weathered line.
- Nothing here apportions blame between sources or between governments, because a concentration record cannot do that. It establishes what the air did and when, which is the prior question.

12. Conclusion

On the government's own monitors, with no model anywhere in the chain, Delhi breathed 104.3 micrograms in 2024 and 97.0 in 2025, the last full year: twenty-one and nineteen times the World Health Organization's guideline, and 2.6 and 2.4 times India's own legal limit. The winter that ended in 2026 averaged 200.5. The long reconstruction says the same thing over a longer span, because Delhi's air has sat at about eighteen times that guideline for a quarter of a century, and it is still there. That is the finding. How it got there matters less than the fact that it has not left.

Delhi's fine-particle pollution rose once and steeply, from the 1990s into the middle of the 2000s, and it has not moved since in either direction. That rise straddles the 2000 instrument change, which is why this report quotes a rate inside each period and no multiple across the join. Forty-seven years of record contain no year at India's own annual standard and no year at the World Health Organization's, and since 2018 the city has recorded seventeen days that the WHO would call clean. No intervention in those forty-seven years has left a mark large enough for this record to detect, and that includes the one most often credited: the CNG conversion of 2002 and 2003, a genuine achievement in what it targeted, which this record was never fine enough to weigh.

The most useful thing the extension adds is a warning about how the recent years will be argued over. The trend since 2010 is indistinguishable from zero, and the sign of any shorter trend can be set at will by choosing a start year, so both a claim of improvement and a claim of collapse can be defended from the same data by anyone willing to pick their own baseline. The defensible statement is narrower and harder. Delhi's air stopped getting worse about eighteen years ago, it has not begun getting better, and its winters have gained about 42 micrograms since 2000 alone.

Sources: read 11 August 2026. LongPMInd (Wang et al., Earth System Science Data, 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 weather via the Open-Meteo archive. Analysis in-house.

13. Sources and references

Primary data. Wang, S., Zhang, M., Zhao, H., Wang, P., Kota, S. H., Fu, Q., Liu, C. and Zhang, H. (2024), "Reconstructing long-term (1980 to 2022) daily ground particulate matter concentrations in India (LongPMInd)", *Earth System Science Data* 16, 3565 to 3577, doi 10.5194/essd-16-3565-2024, reproduced in full as Annexure A. Central Pollution Control Board and Delhi Pollution Control Committee continuous monitors, obtained through the Centre for Research on Energy and Clean Air, 2015 to 11 August 2026. NASA MERRA-2 aerosol and surface-flux reanalyses via Google Earth Engine. ERA5 reanalysis via the Open-Meteo archive.

Standards cited. World Health Organization Global Air Quality Guidelines (2021), annual PM2.5 5 micrograms with an interim target of 15 and a twenty-four-hour guideline of 15 used here as the working limit. India's National Ambient Air Quality Standards (2009), PM2.5 40 micrograms annual and 60 over twenty-four hours.

CNG and Delhi air quality. Narain and Krupnick (2007), *The Impact of Delhi's CNG Program on Air Quality*, Resources for the Future DP 07-06. Guttikunda and colleagues (2023), *What Is Polluting Delhi's Air? 1990 to 2022*,

Sustainability 15(5) 4209. Centre for Science and Environment and EPCA on the loss of the CNG gains. Supreme Court of India, *M. C. Mehta v. Union of India*, 28 July 1998.

The implementation record. Centre for Science and Environment, *The Leapfrog Factor: Clearing the Air in Asian Cities, the Delhi Story* (2003). See especially pages 32 to 33, and the politicians' honour roll at page 48, on the elected Delhi government's delivery of the CNG conversion and on the role of Transport Minister Ajay Maken.

This report's computations are reproducible from the archive under *AQI_Loop/Satellite_AQI*, scripts e0 to e7, together with the series files, the splice statistics and the repair log they write.

ANNEXURE A

Wang, S., Zhang, M., Zhao, H., Wang, P., Kota, S. H., Fu, Q., Liu, C. and Zhang, H. (2024). Reconstructing long-term (1980-2022) daily ground particulate matter concentrations in India (LongPMInd). Earth System Science Data, 16, 3565-3577. Reproduced in full on the following pages as the peer-reviewed source of the primary data used for the years 1980 to 2022 of this report (DOI: 10.5194/essd-16-3565-2024; CC BY 4.0). The years 2023 to July 2026 are not from this dataset: they are Delhi's own CPCB and DPCC monitors, joined to it by the rule set out in section 9.

Earth Syst. Sci. Data, 16, 3565–3577, 2024
https://doi.org/10.5194/essd-16-3565-2024

Open Access Earth System Science Data

Reconstructing long-term (1980–2022) daily ground particulate matter concentrations in India (LongPMInd)

Shuai Wang[1], Mengyuan Zhang[1], Hui Zhao[2], Peng Wang[3,4], Sri Harsha Kota[5], Qingyan Fu[6], Cong Liu[7], and Hongliang Zhang[1,4,8]

[1]Department of Environmental Science and Engineering, Fudan University, Shanghai 200438, China
[2]School of Resources and Environmental Engineering, Jiangsu University of Technology, Changzhou 213001, China
[3]Department of Atmospheric and Oceanic Sciences, Institute of Atmospheric Sciences, Fudan University, Shanghai 200438, China
[4]IRDR ICoE on Risk Interconnectivity and Governance on Weather/Climate Extremes Impact and Public Health, Fudan University, Shanghai 200438, China
[5]Department of Civil Engineering, Indian Institute of Technology, Delhi, 110016, India
[6]Shanghai Academy of Environmental Sciences, Shanghai 200003, China
[7]School of Public Health, Fudan University, Shanghai 200032, China
[8]Institute of Eco-Chongming (IEC), East China Normal University, Shanghai 200062, China

Correspondence: Hongliang Zhang (zhanghl@fudan.edu.cn)

Received: 30 January 2024 – Discussion started: 6 February 2024
Revised: 24 May 2024 – Accepted: 12 June 2024 – Published: 8 August 2024

Abstract. Severe airborne particulate matter (PM, including $PM_{2.5}$ and PM_{10}) pollution in India has caused widespread concern. Accurate PM concentrations are fundamental for scientific policymaking and health impact assessment, while surface observations in India are limited due to scarce sites and uneven distribution. In this work, a simple structured, efficient, and robust model based on the Light Gradient-Boosting Machine (LightGBM) was developed to fuse multisource data and estimate long-term (1980–2022) historical daily ground PM concentrations in India (LongPMInd). The LightGBM model shows good accuracy with out-of-sample, out-of-site, and out-of-year cross-validation (CV) test R^2 values of 0.77, 0.70, and 0.66, respectively. Small performance gaps between $PM_{2.5}$ training and testing (delta RMSE of 1.06, 3.83, and 7.74 $\mu g\,m^{-3}$) indicate low overfitting risks. With great generalization ability, the openly accessible, long-term, and high-quality daily $PM_{2.5}$ and PM_{10} products were then reconstructed (10 km, 1980–2022). This showed that India has experienced severe PM pollution in the Indo-Gangetic Plain (IGP), especially in winter. PM concentrations have significantly increased ($p < 0.05$) in most regions since 2000 (0.34 $\mu g\,m^{-3}\,yr^{-1}$). The turning point occurred in 2018 when the Indian government launched the National Clean Air Programme, and $PM_{2.5}$ concentrations declined in most regions ($-0.78\,\mu g\,m^{-3}\,yr^{-1}$) during 2018–2022. Severe $PM_{2.5}$ pollution caused continuous increased attributable premature mortalities, from 0.73 (95 % confidence interval (CI) [0.65, 0.80]) million in 2000 to 1.22 (95 % CI [1.03, 1.41]) million in 2019, particularly in the IGP, where attributable mortality increased from 0.36 million to 0.60 million. LongPMInd has the potential to support multiple applications of air quality management, public health initiatives, and efforts to address climate change. The daily and monthly $PM_{2.5}$ and PM_{10} concentrations are publicly accessible at https://doi.org/10.5281/zenodo.10073944 (Wang et al., 2023a).

Published by Copernicus Publications.

1 Introduction

Airborne particulate matter (PM, including $PM_{2.5}$ with diameters $< 2.5\,\mu m$ and PM_{10} with diameters $< 10\,\mu m$) not only impacts climate by changing radiation budgets but also has significant adverse effects on human health (Murray et al., 2020; Wang et al., 2012; Yang et al., 2016). India is one of the world's most populous countries, with severe PM pollution resulting from rapid economic development and industrialization over the last few decades. Exposure to $PM_{2.5}$ has become one of the leading causes of health burden in India, contributing to, for example, heart disease, stroke, lung cancer, and premature death (Pandey et al., 2021; Dandona et al., 2017).

Accurate ground PM concentrations are prerequisites for evidence-based policymaking and health impact assessments. The Central Pollution Control Board (CPCB) of India has established and maintained ground-based monitoring networks with ~ 335 continuous ambient air quality monitoring stations (CAAQMSs) currently. However, these monitoring sites are unevenly distributed (mainly located in urban, residential, and industrial areas), with a limited number of sites (monitoring density: ~ 0.6 sites per million population) (Brauer et al., 2019), and many cities even have no monitoring sites (Martin et al., 2019). Therefore, the surface observations alone are not sufficient to support air quality management, especially on a regional scale (Pant et al., 2019; Dey et al., 2020).

Two main approaches have been used for large-scale and long-term $PM_{2.5}$ estimation: scaling methods and statistical methods. Scaling methods use chemical transport models (CTMs) to simulate the association between aerosol optical depth (AOD) and $PM_{2.5}$ and require no ground observations. However, the relationship between $PM_{2.5}$ and AOD is spatially and temporally variable and without the constraints of ground observations, so these methods usually have a large uncertainty (Ma et al., 2022). Compared with scaling methods, statistical methods based on multivariate data fusion have higher prediction accuracy and have been widely used. Statistical models (traditional linear and nonlinear regression and machine learning algorithms) estimate $PM_{2.5}$ concentrations by fitting the relationship between input variables (meteorological factors, emissions, and other relevant variables) and target variables (Wang et al., 2023b; Wei et al., 2021a; X. Ren et al., 2022; Katoch et al., 2023).

Tree-based machine learning (ML) models typically outperform deep learning approaches and traditional machine learning methods in tabular data (e.g., air pollutant observation datasets) and thus have been widely developed and used (Grinsztajn et al., 2022; Sayeed et al., 2022). Wei et al. (2021a) and Li et al. (2021) reconstructed long-term $PM_{2.5}$ data records in China by fusing satellite, meteorological, and emission data using a spatiotemporal tree-based model. Ni et al. (2024) analyzed the contribution of meteorology and emissions to O_3 in China using a chemical transport model (GEOS-Chem) and a tree-based model. Sayeed et al. (2022) improved analysis of the $PM_{2.5}$ concentration in the continental United States using the random forest approach coupled with meteorology and aerosol species of the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2).

Some studies have estimated $PM_{2.5}$ concentrations in India using different methods. Chowdhury et al. (2019) used the $PM_{2.5}$–AOD equation method to estimate $PM_{2.5}$ concentrations in Delhi; however, AOD satellite data suffer from significant non-random misses, especially during cloud cover and on hazy, polluted days, so it is difficult to derive a full-spatiotemporal-coverage PM dataset (Wang et al., 2023b; Bai et al., 2022). Bali et al. (2021) and Dey et al. (2020) estimated total $PM_{2.5}$ in India through empirical coefficients and the MERRA-2 dataset, but these coefficients vary with geographic location and pollution scenarios, which makes the related estimation potentially unreliable. Kumar et al. (2023) analyzed $PM_{2.5}$ for India using a random forest model and showed a gap between train and test scores with a risk of overfitting. In addition, global-scale daily $PM_{2.5}$ concentration datasets (including India) have recently been developed, including GlobalHighAirPollutants (GHAP) (Wei et al., 2023) and the Long-term Gap-free High-resolution Air Pollutants concentration dataset version 2 (LGHAPv2) (Bai et al., 2024). Global monthly $PM_{2.5}$ datasets have also been developed before (Van Donkelaar et al., 2021). These datasets were trained on a global scale and estimated $PM_{2.5}$ concentrations for the India region. The severity of $PM_{2.5}$ pollution in India is much higher than in Europe and the United States (Wei et al., 2023). However, due to the small number of observations recorded, the global model can gain only limited knowledge of $PM_{2.5}$ pollution in India during the training process. Therefore, the reliability and robustness of global model estimates of $PM_{2.5}$ concentrations in India should be systematically assessed. Building a model locally in India can provide a useful comparison, which can also complement the biases in global modeling (e.g., due to focusing more on lightly polluted regions such as Europe and the United States). However, it is challenging to establish long-term, full-coverage, high-accuracy, open-source PM data products locally in India due to insufficient observational data and lack of model robustness because of variations in data distribution across regions and years (Kumar et al., 2023; Dey et al., 2020).

To improve performance, previous models have usually had high complexity, such as numerous trees and leaf nodes (Zhang et al., 2021; Huang et al., 2021). This practice increases the computational resource requirement and is prone to overfitting, leading to a large gap between the performance of the training and testing data (Zhang et al., 2021; Jabbar and Khan, 2015; Ying, 2019). Therefore, it is necessary to minimize model complexity to avoid overfitting. The Light Gradient-Boosting Machine (LightGBM) is an optimized gradient boosting decision tree (GBDT) (Ke et al., 2017), and it has shown superior performance in many fields (Wei et al., 2021b; Yan et al., 2021; Sun et al., 2020; Liang et al., 2020). LightGBM uses a histogram-based decision tree algorithm along with gradient-based one-side sampling (GOSS), which can save memory and computation time (Ke et al., 2017). Our previous study comparing several commonly used machine learning models found that LightGBM has similar performance to the eXtreme Gradient Boosting (XGBoost) with the highest accuracy, but LightGBM was faster and more robust, with the potential to estimate long-term concentrations of PM in India (Wang et al., 2023c).

In this work, a simple structured, efficient, and robust model based on LightGBM was developed to estimate PM concentration. Three cross-validation methods and separate test datasets were designed to evaluate model performance. Long-term (1980–2022) and open-source datasets with a spatial resolution of 10 km of $PM_{2.5}$ and PM_{10} in India were then generated, and the mortalities due to $PM_{2.5}$-induced diseases were also estimated. The concentration datasets could help with pollution formation analysis, assessment of PM health risks, and air quality management in India.

2 Materials and methods

2.1 Data sources

Table 1 shows the multisource datasets used in this study. Ground observations of $PM_{2.5}$ and PM_{10} during 2018–2022 in India were collected from the CPCB air quality monitoring network (https://www.cpcb.nic.in, last access: 29 July 2024). The location of monitoring sites is shown in Fig. S1 in the Supplement. As extreme values affect model robustness, the bottom and top 0.01 % of observation data were excluded. The fifth-generation ECMWF atmospheric reanalysis dataset ERA5-Land covering 1980–2022 was used. Features were selected by their relative importance, which was calculated by their gain, and several meteorological factors with high relative importance were included (Table 1). Data products of the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2), covering 1980–2022 were also collected, with data including aerosol optical depth and aerosol components and precursors (black carbon, organic carbon, sulfate, dust, and SO_2).

2.2 Model building

In this study, LightGBM (Ke et al., 2017), an efficient gradient boosting decision tree (GBDT), was used to estimate $PM_{2.5}$ and PM_{10}, which has been proven to be accurate, fast, and robust in our previous studies (Wang et al., 2023c, d). A grid search cross-validation (CV) method was used to select the optimal hyperparameters. An algorithm for hyperparameter selection (Algorithm S1 in the Supplement) was designed to ensure the model's generalization ability. Looping is performed to increase the model complexity (e.g., number of trees), followed by ending the loop and returning the hyperparameters when the model-predicted RMSE does not decrease significantly (< 0.01) or the difference between the training and predicted RMSE does not increase significantly (< 0.05). Features were selected based on their relative importance. Ten meteorological features, six emission-related features, and total aerosol extinction were used to train LightGBM and estimate PM concentrations (Fig. 1). The meteorological and emission features contributed 64 % and 31 % to the $PM_{2.5}$ prediction.

Meteorology is more important than emissions. Compared to MERRA-2, which has higher uncertainty and lower spatial resolution, ERA5 has higher resolution and accuracy, and its meteorological features can provide richer information and contribute more in model training and thus have higher importance (Muñoz-Sabater et al., 2021; Hersbach et al., 2020). Besides, more meteorological features were used to train the model, thus contributing more to prediction results. The highest importance of surface pressure can be attributed to the important effect on $PM_{2.5}$ concentration and its high data quality (Chen et al., 2020; Bauer et al., 2015).

Three independent CV methods and three metrics (coefficient of determination: R^2; root mean square error: RMSE; mean absolute error: MAE) were designed to evaluate the model's spatiotemporal predictive power. The first is out-of-sample CV, where the dataset is randomly divided into 10 subsets, 1 of which is taken in turn for testing and the remaining 9 of which are used for training; this is repeated 10 times and averaged. The second is out-of-site CV, which is similar to the out-of-sample CV, but the dataset is randomly divided by site. This method can measure the model's spatial predictive power. The third method is interannual out-of-year CV, which sequentially takes 1 year of data for testing and the rest for training. This approach can measure the model's predictive power for the years with no observations. Besides, observations in January–June 2023 were used as a separate test set, and these data were not involved in any of the training and hyperparameter selection processes.

2.3 Mortality estimation

According to the database of the Global Burden of Disease (GBD) study in 2019 (Murray et al., 2020; Vos et al., 2020), annual average concentrations were used to assess long-term exposure to $PM_{2.5}$, and premature deaths were assessed using the following equation:

$$M_{y,i,j} = \frac{\mathrm{RR}_j(C_{y,i}) - 1}{\mathrm{RR}_j(C_{y,i})} \times P_{y,i} \times I_{y,j}, \quad (1)$$

where $M_{y,i}$ represents the mortality attributable to cause j due to long-term $PM_{2.5}$ exposure in year y in region i, $\mathrm{RR}(C_{y,i})$ represents the relative risk of cause j for year y in region i, $P_{y,i}$ represents the population j in year y in region i, and I_y represents the baseline mortality in year y.

Table 1. Summary of the ERA5, MERRA-2, and ground observation data used in this study.

Type	Variable	Description	Spatial resolution	Temporal resolution
ERA5	SSRD	Surface solar radiation	$0.1° \times 0.1°$	Hourly
	BLH	Boundary layer height	$0.25° \times 0.25°$	Hourly
	EVAP	Evaporation	$0.1° \times 0.1°$	Hourly
	TEMP2	2 m air temperature	$0.1° \times 0.1°$	Hourly
	DEWP2	2 m dew point temperature	$0.1° \times 0.1°$	Hourly
	SP	Surface pressure	$0.1° \times 0.1°$	Hourly
	TPREC	Total precipitation	$0.1° \times 0.1°$	Hourly
	TCLOUD	Total cloud cover	$0.25° \times 0.25°$	Hourly
	UWIND10	10 m u component of wind	$0.1° \times 0.1°$	Hourly
	VWIND10	10 m v component of wind	$0.1° \times 0.1°$	Hourly
MERRA-2	BCSMASS	Black carbon surface mass concentration	$0.5° \times 0.625°$	Hourly
	OCSMASS	Organic carbon surface mass concentration	$0.5° \times 0.625°$	Hourly
	DUSMASS25	Dust $PM_{2.5}$ surface mass concentration	$0.5° \times 0.625°$	Hourly
	DUSMASS	Dust surface mass concentration	$0.5° \times 0.625°$	Hourly
	SO2SMASS	Sulfur dioxide surface mass concentration	$0.5° \times 0.625°$	Hourly
	SO4SMASS	Sulfate surface mass concentration	$0.5° \times 0.625°$	Hourly
	TOTEXTTAU	Total aerosol extinction [550 nm]	$0.5° \times 0.625°$	Hourly
Observation	$PM_{2.5}$, PM_{10}	Particulate matter	Point	Hourly

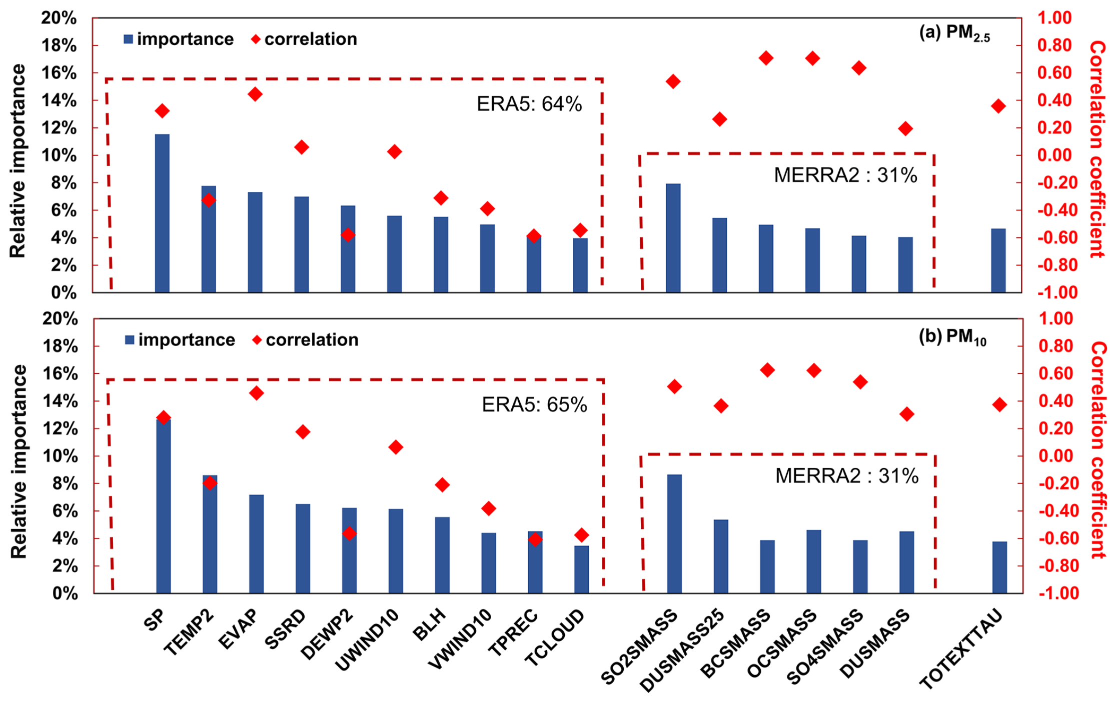


Figure 1. Relative importance and correlation coefficient for the $PM_{2.5}$ and PM_{10} estimate models. A description of the features is given in Table 1.

$PM_{2.5}$ exposure-related deaths due to ischemic heart disease (CVD_IHD), chronic stroke (CVD_stroke), obstructive pulmonary disease (RESP_COPD), lung cancer (NEO_LUNG), lower respiratory infections (LRI), and type 2 diabetes mellitus (T2_DM) were estimated. The gridded population data were obtained from WorldPop datasets (https://www.worldpop.org, last access: 29 July 2024). The relative risk is a discrete value obtained from GBD 2019 and is the relative risk corresponding to each $PM_{2.5}$ concentration level. Details about the calculation of relative risk can be found in GBD 2019 (Murray et al., 2020). Annual baseline mortality (2000–2019) and the risk of cause-specific deaths

Table 2. Summary of the LongPMInd dataset.

Data description	LongPMInd dataset
Data type	Gridded
File format	NetCDF
Species	$PM_{2.5}$, PM_{10}
Spatial reference	WGS 84
Horizontal resolution	$0.1° \times 0.1°$ ($\approx$ 10 km $\times$ 10 km)
Horizontal coverage	India ([5.0° N, 40.0° N], [60° E, 100° E])
Temporal resolution	Daily, monthly, and yearly
Temporal coverage	1980–2022

at different $PM_{2.5}$ levels were obtained from GBD 2019. The minimum-risk exposure level for the health effects of $PM_{2.5}$ is in the range of 2.4 to 5.9 µg m^{-3}.

3 Data availability

The LongPMInd dataset, including daily $PM_{2.5}$ and PM_{10} concentration (10 km) for India during 1980–2022, is publicly accessible. All data are provided in NetCDF format and can be downloaded at https://doi.org/10.5281/zenodo.10073944 (Wang et al., 2023a).

4 Results

4.1 Long-term India $PM_{2.5}$ dataset

Applying the trained LightGBM model to the large input dataset constructed for the years 1980–2022, the long-term high-quality daily $PM_{2.5}$ and PM_{10} products of India (LongPMInd) are reconstructed. Table 2 summarizes basic information about the dataset; the data are provided in NetCDF format with a spatial resolution of 10 km. The LongPMInd dataset, to the best of our knowledge, is the first open-source, long-term (i.e., 1980–2022), and relatively high accuracy dataset covering the whole of India. The daily, monthly, and yearly $PM_{2.5}$ and PM_{10} datasets are publicly available at https://doi.org/10.5281/zenodo.10073944 (Wang et al., 2023a).

4.2 Model performance

Table 3 shows the training and testing results of out-of-sample CV, out-of-site CV, and out-of-year CV for daily $PM_{2.5}$ and PM_{10}. Overall, the model shows good accuracy, with out-of-sample CV R^2 values of 0.77 and 0.76 and RMSE of 29.57 and 51.63 µg m^{-3} for daily $PM_{2.5}$ and PM_{10}. Monthly predictions show better performance, with out-of-sample CV R^2 values of 0.87 and 0.86 and RMSE of 17.65 and 31.26 µg m^{-3} for monthly $PM_{2.5}$ and PM_{10}. More importantly, out-of-sample CV results of training and testing showed small accuracy gaps with RMSE and MAE of 1.06 (4 %) and 0.51 (3 %) µg m^{-3} for $PM_{2.5}$ and 1.52 (3 %) and 0.9 (3 %) µg m^{-3} for PM_{10}, reflecting good generalization ability. Out-of-site CV measures the model's predictive ability for unobserved areas. The spatially validated R^2 and RMSE for $PM_{2.5}$ and PM_{10} were 0.70 and 0.65 and 31.73 and 51.37 µg m^{-3}, respectively, indicating the model's ability to fill the unobserved areas accurately. The small performance gap between out-of-site CV training and testing also reflects good spatial generalization ability. Observations before 2018 are limited due to the number and quality of sites. Out-of-year CV was used to evaluate LightGBM prediction performance, which was conducted by sequentially taking 1 year of data for testing and the rest for training. The model's prediction accuracy for unobserved years decreases slightly compared to out-of-sample CV (R^2 decreases by 14 %, and RMSE increases by 20 %) due to differences in data distributions among years (Fig. S2). Notably, most predictions are consistent with observations, with most data samples evenly distributed around the 1 : 1 line (Fig. S3) but with underestimation for high PM levels and overestimation for low PM levels (slopes of 0.75 and 0.74 and intercepts of 16.45 and 35.79 µg m^{-3} for daily $PM_{2.5}$ and PM_{10} predictions). Monthly predictions show better agreement with observations, with slopes of 0.84 and 0.82 and intercepts of 10.26 and 23.53 µg m^{-3} for monthly $PM_{2.5}$ and PM_{10}. The underestimation and overestimation indicate the potential unreliability of model predictions for extreme pollution and extreme clean days. This can be attributed to the small proportion of data records for extreme pollution and clean days.

Observations from January to June 2023 were used for testing and were not involved in any training or hyperparameter selection processes (Fig. S4 and Table S1 in the Supplement). Six representative regions were selected for the analysis including Delhi and Uttar Pradesh (the Indo-Gangetic Plain – IGP – region), Gujarat (western India region), Madhya Pradesh (central India region), West Bengal (eastern India region), and Andhra Pradesh (southern India region). The model shows accurate prediction ability with RMSE of 33.58 and 64.25 µg m^{-3} for $PM_{2.5}$ and PM_{10}, respectively, in India. The model can capture the decreasing trend in PM concentration from January to June in different regions of India but with some biases, e.g., overestimation of $PM_{2.5}$ in Uttar Pradesh on 8 January and underestimation of haze pollution in Gujarat on 19 February. The large RMSE of $PM_{2.5}$ prediction in Uttar Pradesh (32.72 µg m^{-3}) could be attributed to the complexity of pollution causes in the region as well as insufficient observation data. The small RMSE (8.34 µg m^{-3}) of $PM_{2.5}$ prediction in Andhra Pradesh can be related to the light haze pollution and small fluctuation in $PM_{2.5}$ concentration.

4.3 Spatial and temporal trends

First, spatial patterns of $PM_{2.5}$ and PM_{10} are analyzed (Figs. S5 and S6). The Indo-Gangetic Plain (IGP) and western arid regions show high levels of $PM_{2.5}$ and PM_{10}, es-

Table 3. Training and testing results of out-of-sample CV, out-of-site CV, and out-of-year CV for daily $PM_{2.5}$ and PM_{10} (2018–2022). RMSE and MAE unit: $\mu g\,m^{-3}$.

Species	Type	R^2		RMSE ($\mu g\,m^{-3}$)		MAE ($\mu g\,m^{-3}$)	
		Test	Train	Test	Train	Test	Train
$PM_{2.5}$	out-of-sample	0.77	0.79	29.57	28.51	18.76	18.25
	out-of-site	0.70	0.79	31.73	27.90	20.32	17.78
	out-of-year	0.66	0.79	35.35	27.61	21.54	17.61
PM_{10}	out-of-sample	0.76	0.77	51.63	50.11	35.42	34.52
	out-of-site	0.65	0.77	57.37	49.42	39.92	33.94
	out-of-year	0.66	0.78	60.65	49.06	40.74	33.72

pecially for years after 2000. Low PM concentrations were observed in southern India. The high terrain in the northern and southern IGP is unfavorable for pollutant dispersion. Intense human activities in the IGP (population > 700 million) emit large quantities of primary PM and gas pollutants (SO_2 and nitrogen oxide), and this, coupled with unfavorable dispersion conditions, leads to severe PM pollution (Dey et al., 2020; Maheshwarkar et al., 2022). Both $PM_{2.5}$ and PM_{10} concentrations show a north-to-south (high-to-low) distribution, consistent with the population distribution and corresponding anthropogenic emissions (Upadhyay et al., 2020; Dey et al., 2020).

Figure 2 shows the spatial patterns of seasonal $PM_{2.5}$ and PM_{10} anomalies. The highest PM levels occurred in winter, especially in the IGP (positive anomaly > $20\,\mu g\,m^{-3}$ relative to the annual mean during 1980–2022). This enhancement is related to additional anthropogenic emissions (from space and water heating of households especially in cold places like the IGP) and stable meteorological conditions (low boundary layer height and low wind speed) (Pandey et al., 2014; Tiwari et al., 2013). During the pre-monsoon season (March–April–May), favorable meteorological conditions (increased boundary layer height due to increased temperature and wind speeds) reduce $PM_{2.5}$ concentrations in the IGP area (Dey et al., 2020). During the monsoon season (June to September), rainfall enhances PM deposition, resulting in a substantial reduction in PM concentrations. With the end of the monsoon (post-monsoon season, October and November), less rainfall, lower temperatures, extensive open biomass burning (for heating), and reduced boundary layer heights exacerbate PM pollution (Nagpure et al., 2015; Kumari et al., 2021).

The long-term trends of aerosols in India can be better examined given the advantage of long temporal coverage of the LongPMInd dataset. The monthly $PM_{2.5}$ and PM_{10} anomalies from 1980 to 2022 in India and typical regions were first calculated (Figs. 3 and S7). PM concentrations slowly increased in India ($0.19\,\mu g\,m^{-3}\,yr^{-1}$) before 2000; in the IGP and eastern India, they increased by 0.43 and $0.26\,\mu g\,m^{-3}\,yr^{-1}$, respectively. The PM concentrations jumped in 2000, which can be attributed to the absence of satellite data for MERRA-2 before 2000 (Buchard et al., 2017). The MERRA-2 dataset before 2000 could not provide the same level of data quality as in the later period, further leading to a systematic bias in the model estimates. With accelerated industrialization, anthropogenic emissions of primary particulate matter (PPM) and precursors of secondary aerosols (e.g., SO_2, NO, and NH_3) have increased since 2000 (Pandey et al., 2014; Nagpure et al., 2015), leading to significant increases in PM concentrations in most regions ($p <$ 0.05), except for western India (Figs. 4 and S8). $PM_{2.5}$ increased by 0.50 and $0.46\,\mu g\,m^{-3}\,yr^{-1}$ in the IGP and eastern India during 2000–2017. In early 2018, the Indian government launched the National Clean Air Programme (NCAP). The interventions were grouped into transport, industry, waste management, domestic, and construction activities as well as road dust and others (Ganguly et al., 2020). Emissions declined rapidly, and $PM_{2.5}$ concentrations declined significantly in the IGP ($1.63\,\mu g\,m^{-3}\,yr^{-1}$), western India ($1.22\,\mu g\,m^{-3}\,yr^{-1}$), and southern India ($0.52\,\mu g\,m^{-3}\,yr^{-1}$). However, PM concentrations in east-central India showed an increasing trend (Fig. 4), which may be related to emissions from mining activities and related industries and thermal power plants (Upadhyay et al., 2020).

4.4 Health burden analysis

The health burden of $PM_{2.5}$ was estimated from 2000–2019 following the rapid increase in $PM_{2.5}$ concentrations after 2000. Using GBD 2019, premature deaths attributed to $PM_{2.5}$ exposure were calculated for six diseases, comprising ischemic heart disease (CVD_IHD), chronic stroke (CVD_stroke), obstructive pulmonary disease (RESP_COPD), lung cancer (NEO_LUNG), lower respiratory infections (LRI), and diabetes mellitus type 2 (T2_DM) (Murray et al., 2020; Vos et al., 2020).

Figure 5 shows the changes in annual average $PM_{2.5}$ concentrations and corresponding attributed deaths, and Table S2 shows the related uncertainties. $PM_{2.5}$ concentrations showed a fluctuating upward trend with a continuous increase in attributable premature mortality, from 0.73 (95 % confidence interval (CI) [0.65, 0.80]) million prema-

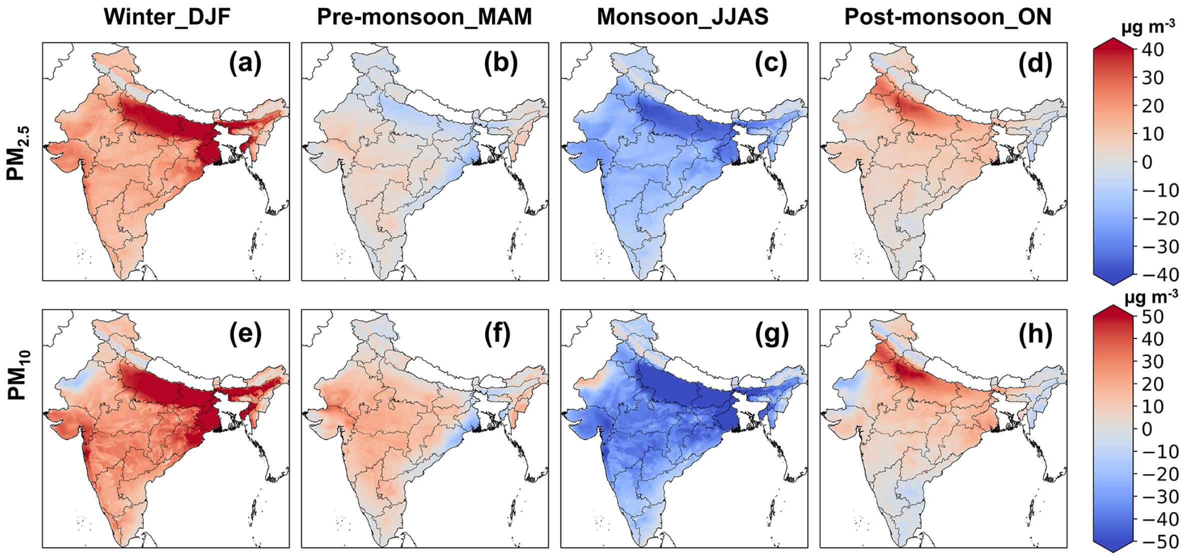


Figure 2. Spatial patterns of seasonal $PM_{2.5}$ and PM_{10} anomalies (the difference between the seasonal mean and annual mean) in India during 1980–2022.

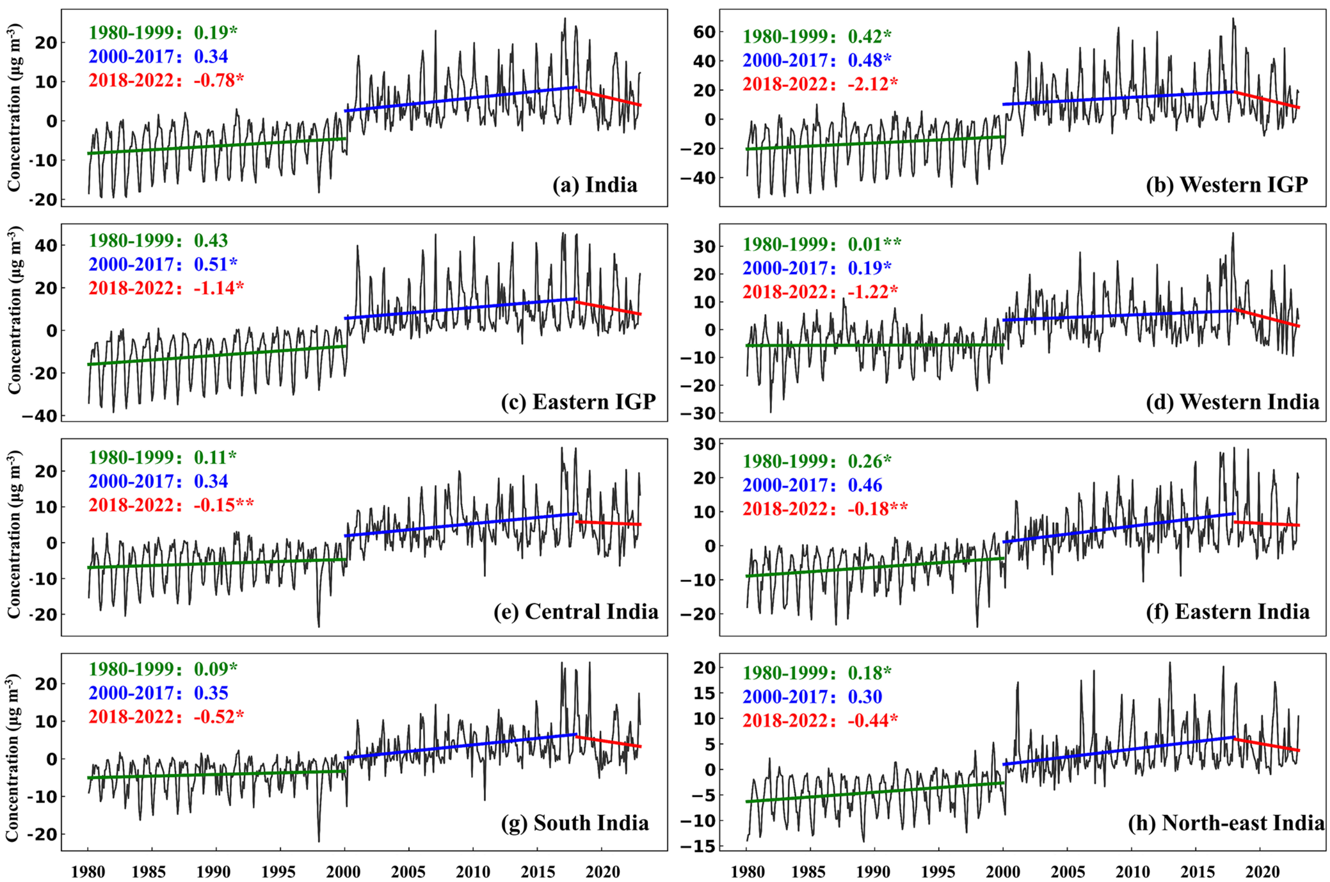


Figure 3. Time series of the monthly $PM_{2.5}$ anomaly from 1980 to 2022 in India (overall and regionally). The colored straight lines are the linear regression trend ($\mu g\, m^{-3}\, yr^{-1}$) for different periods in China, and * represent the significance of the trends (* mean $p < 0.05$ and ** mean $p < 0.01$).

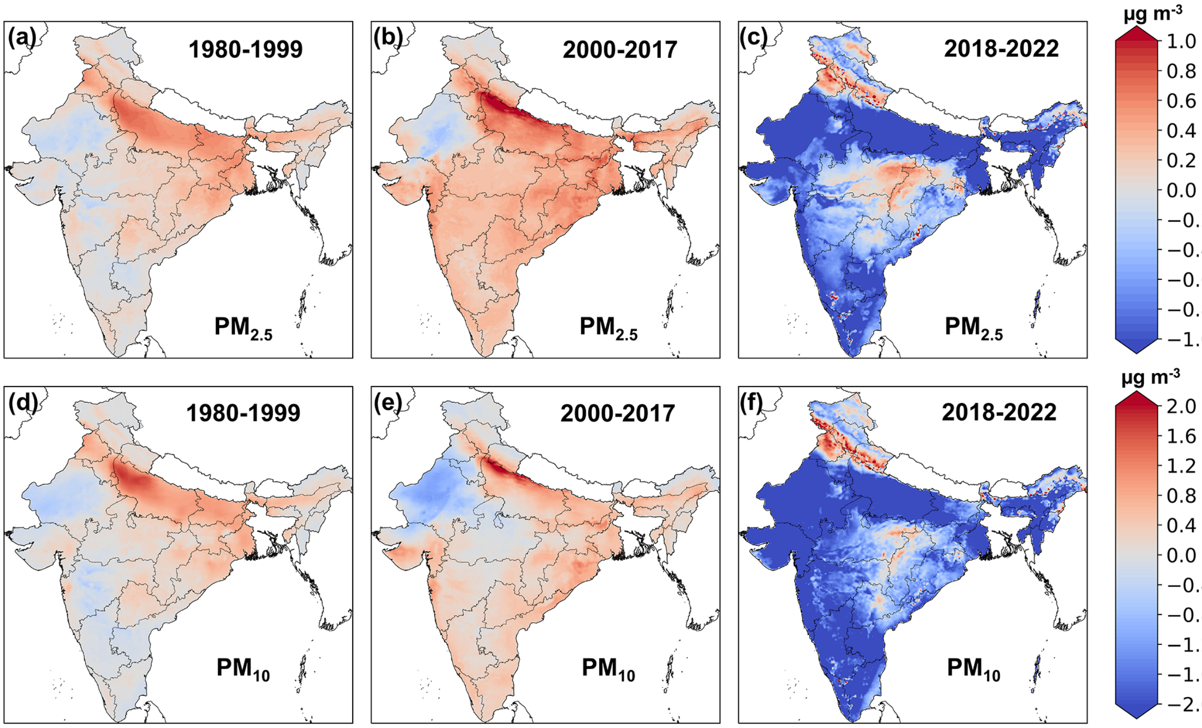


Figure 4. Spatial patterns of annual changes for $PM_{2.5}$ and PM_{10} ($\mu g\, m^{-3}\, yr^{-1}$) during different periods (1980–1999, 2000–2017, and 2018–2022). Publisher's remark: please note that the above figure contains disputed territories.

ture deaths in 2000 to 1.22 (95 % CI [1.03, 1.41]) million in 2019, with CVD_IHD, CVD_stroke, RESP_COPD, NEO_LUNG, LRI, and T2_DM causing an annual average of 0.35 million, 0.21 million, 0.21 million, 0.02 million, 0.12 million, and 0.04 million premature deaths, respectively. $PM_{2.5}$-attributable deaths were counted by region (Fig. 5). The IGP had the highest attributable premature deaths, increasing from 0.36 million in 2000 to 0.60 million in 2019, due to high population density coupled with severe haze pollution (Dey et al., 2020; Pandey et al., 2021).

To reduce premature deaths from $PM_{2.5}$ exposure, policies to mitigate $PM_{2.5}$ pollution should be implemented. In addition, appropriate health advice and enhanced medical facilities to reduce baseline mortality are also important to reduce the health burden (Maji et al., 2023). India has experienced rapid urbanization and large-scale population migration, which introduces uncertainty into health risk estimates for $PM_{2.5}$ (Shi et al., 2020). Country-level baseline disease rates were used, so regional differences were not accounted for due to lack of data, which could introduce some error. In addition, uncertainties in relative risk, population, and $PM_{2.5}$ concentrations may also introduce errors into health risk estimates.

4.5 Model complexity

Model complexity can be measured by the number of parameters the model has. As model complexity increases, the

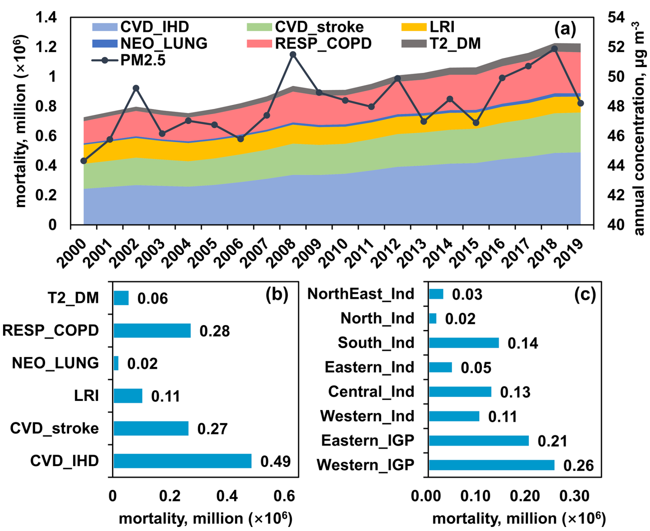


Figure 5. Annual mortalities due to $PM_{2.5}$-induced diseases in India during 2000–2019, including ischemic heart disease (CVD_IHD), chronic stroke (CVD_stroke), obstructive pulmonary disease (RESP_COPD), lung cancer (NEO_LUNG), lower respiratory infections (LRI), and diabetes mellitus type 2 (T2_DM). Panels **(b)** and **(c)** show statistical results for causes and regions.

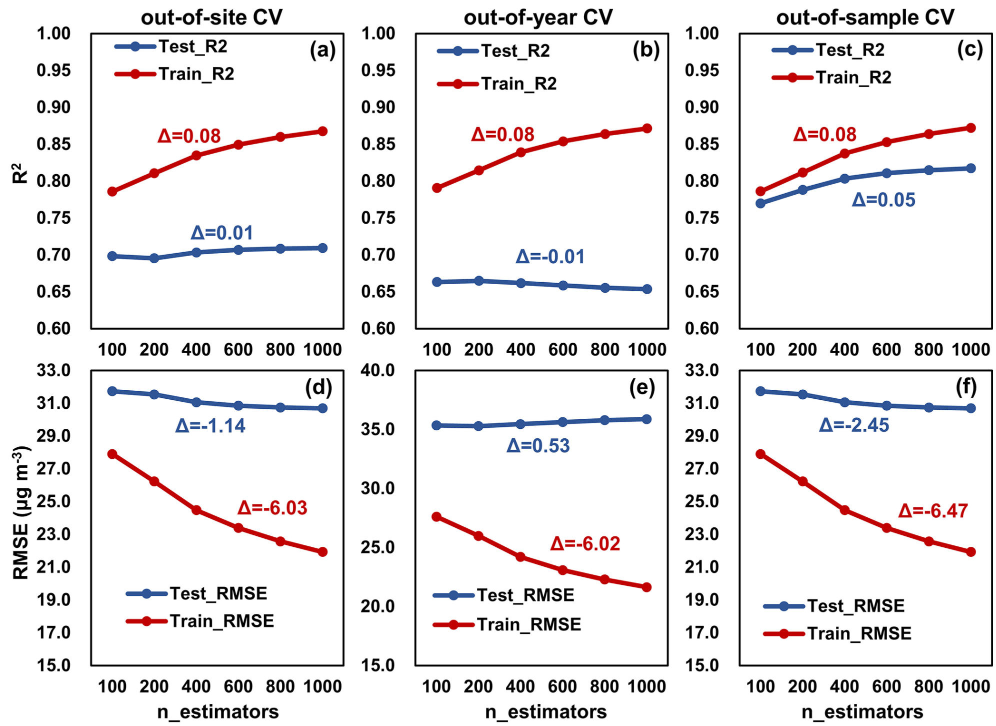


Figure 6. Three CV results of the model complexity test for $PM_{2.5}$ estimation. The n_estimators is the number of trees, representing the complexity of LightGBM. Δ is the difference between the metrics with n_estimators = 1000 and n_estimators = 100. Units of RMSE and MAE are $\mu g\, m^{-3}$.

model is more capable of learning complex patterns in the data, but at the same time, it may lead to overfitting and inaccurate predictions of new and unseen data (Hu et al., 2021). The impact of the complexity of the tree-based LightGBM model on the performance of training and testing is analyzed (Fig. 6). The number of trees (n_estimators) was used as a complexity proxy, and the other hyperparameters were kept constant. All three cross-validation results show that the increase in model complexity improves the model's fitting ability, increasing R^2 and decreasing RMSE. However, the increase in complexity did not improve the model's predictive performance. With n_estimators increasing from 100 to 1000, there was no significant change in R^2 for the out-of-site and out-of-year CV (−0.01 to 0.01), and the RMSE for the out-of-year CV on the contrary increased by 0.53. Out-of-sample CV showed an improvement in R^2 but with a limited reduction in RMSE (−2.45). Therefore, using only out-of-sample CV to select hyperparameters and evaluate the model is limiting, and out-of-site and out-of-year CV allows a more objective evaluation of the model's generalization ability.

4.6 Uncertainties

Uncertainty in this study comes from two main sources: the machine learning model and the dataset used. Firstly, machine learning is essentially based on probability theory and is influenced by the distribution pattern of the target variable ($PM_{2.5}$ and PM_{10}) (Yang et al., 2021; Breiman, 2001). Due to the low frequency of extreme pollution scenarios, the model suffers from the problem of smoothing predictions, i.e., underestimating high-pollution scenarios (Wei et al., 2021a; Yu et al., 2023; Geng et al., 2021). In addition, machine learning has limitations in describing atmospheric physical and chemical processes, and it is difficult to fit complex, logistically long processes, such as secondary aerosol generation (Stirnberg et al., 2021; Li et al., 2023). Attempts have been made to incorporate physical constraints into neural networks to improve interpretability, but this approach is limited to spatially continuous two-dimensional data (Geiss et al., 2022). Other studies have shown that chemical reaction processes can be described by neural networks, but it is still a challenge to efficiently couple them with CTMs (Huang et al., 2022; Huang and Seinfeld, 2022).

The second aspect is the uncertainties caused by the datasets. First, the label (observations) and corresponding features (MERRA-2 and ERA5) have a long-tailed distribution with few high-pollution records, so there is an issue of imbalance regression (Yang et al., 2021). The model was trained with a bias towards denser observations, leading to the underestimation of high-pollution scenarios. For the problem of imbalanced regression, there are currently mainly data-based solutions and model-based solutions (J. Ren et al., 2022). Data-based solutions require acquiring more data or changing the data distribution by resampling. Model-based solutions increase the weighting of fewer samples (high-pollution scenarios) by modifying the loss function. Both methods can improve the accuracy of fewer samples, but they are not suitable for the task of this study because the distribution of the data was altered. Therefore, more observations should be collected in the future to increase observations recorded for high-pollution scenarios and mitigate the problem of imbalanced regression. In addition, observational data can only be collected for recent years (2018–2022), which may lead to uncertainties when inferring PM concentrations for historical years. In out-of-year validation, the gap between training and testing is mainly attributed to the difference in the data distribution among years (data drift, Fig. S2). Besides, changes in climate and human activities over the decades may affect the relationship among emissions, meteorology, and PM concentrations, resulting in extra uncertainty (concept drift).

Secondly, the uncertainty in the input feature sets (ERA5 and MERRA-2) also affects the estimation results. The uncertainty in ERA5, a widely used meteorological reanalysis dataset, has been systematically analyzed. ERA5 has good accuracy for most meteorological factors, exceeding other reanalysis data (Muñoz-Sabater et al., 2021; Hersbach et al., 2020). With MODIS data as a reference, the global mean surface temperature of ERA5-Land shows lower uncertainty (Muñoz-Sabater et al., 2021). For precipitation, ERA5 shows 77 % correlation with monthly mean Global Precipitation Climatology Project (GPCP) data (Hersbach et al., 2020). Compared to the pre-assimilation data, ERA5-Land provides an improved fit to tropospheric winds and humidity (Hersbach et al., 2020).

MERRA-2 is a global air pollution reanalysis dataset, published and maintained by NASA; it has been widely used for PM pollution studies in the Indian region, and its reliability has been extensively analyzed (Gueymard and Yang, 2020; Navinya et al., 2020; Buchard et al., 2017). For MERRA-2 AOD, evaluation using AERONET observations showed that MERRA-2 outperformed the Copernicus Atmosphere Monitoring Service (CAMS) in most regions (Gueymard and Yang, 2020). Kumar et al. (2023) predicted ground-level $PM_{2.5}$ concentrations in India using only MERRA-2 and machine learning methods, proving the reliability of MERRA-2 data.

In addition, before 2000, there were no assimilated satellite data for MERRA-2, and nitrate was not provided in the MERRA-2 aerosol reanalysis datasets (Buchard et al., 2017). These issues may be detrimental to the accuracy of the LongPMInd dataset. Previous studies in India have shown that $PM_{2.5}$ estimates based on MERRA-2 and empirical formulas suffer from inaccuracies due to issues such as the absence of nitrate, which can be improved by tree-based modeling (Sayeed et al., 2022). The model trained in this study relies heavily on ERA5 (64 % relative contribution) with a minor contribution from MERRA-2 (36 %). Although tree-based models can improve $PM_{2.5}$ estimation accuracy and the inclusion of ERA5 meteorological features reduces the model's dependence on MERRA-2, model accuracy may decrease for areas dominated by nitrate emissions and for years before 2000.

Supplement. This file contains the research domain, feature importance, spatial and temporal patterns of $PM_{2.5}$ and PM_{10}, and uncertainty in estimated annual mortalities. The supplement related to this article is available online at: https://doi.org/10.5194/essd-16-3565-2024-supplement.

Author contributions. SW: methodology, software, writing (original draft). MZ: visualization, validation. HuZ: data curation, methodology. PW: methodology, writing (reviewing and editing). SHK: data curation. QF: writing (reviewing and editing). CL: writing (reviewing and editing). HoZ: conceptualization, funding acquisition, supervision, writing (reviewing and editing).

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. We thank the editor Jing Wei, the reviewer Yang Yang, and the anonymous reviewer for their helpful comments and suggestions for improving this paper. In addition, we would like to thank all the organizations and groups that provided important datasets for this study.

Financial support. This work was supported by the National Key R&D Program of China (grant no. 2022YFC3701105), the co-fund DFG–NSFC Sino-German AirChanges project (grant no. 448720203), the National Natural Science Foundation of China (grant nos. 42077194 and 42061134008), and the Shanghai International Science and Technology Partnership Project (grant no. 21230780200).

Review statement. This paper was edited by Jing Wei and reviewed by Yang Yang and one anonymous referee.

References

Bai, K., Li, K., Ma, M., Li, K., Li, Z., Guo, J., Chang, N.-B., Tan, Z., and Han, D.: LGHAP: the Long-term Gap-free High-resolution Air Pollutant concentration dataset, derived via tensor-flow-based multimodal data fusion, Earth Syst. Sci. Data, 14, 907–927, https://doi.org/10.5194/essd-14-907-2022, 2022.

Bai, K., Li, K., Shao, L., Li, X., Liu, C., Li, Z., Ma, M., Han, D., Sun, Y., Zheng, Z., Li, R., Chang, N.-B., and Guo, J.: LGHAP v2: a global gap-free aerosol optical depth and $PM_{2.5}$ concentration dataset since 2000 derived via big Earth data analytics, Earth Syst. Sci. Data, 16, 2425–2448, https://doi.org/10.5194/essd-16-2425-2024, 2024.

Bali, K., Dey, S., and Ganguly, D.: Diurnal patterns in ambient $PM_{2.5}$ exposure over India using MERRA-2 reanalysis data, Atmos. Environ., 248, 118180, https://doi.org/10.1016/j.atmosenv.2020.118180, 2021.

Bauer, P., Thorpe, A., and Brunet, G.: The quiet revolution of numerical weather prediction, Nature, 525, 47–55, https://doi.org/10.1038/nature14956, 2015.

Brauer, M., Guttikunda, S. K., Nishadh, K. A., Dey, S., Tripathi, S. N., Weagle, C., and Martin, R. V.: Examination of monitoring approaches for ambient air pollution: A case study for India, Atmos. Environ., 216, 116940, https://doi.org/10.1016/j.atmosenv.2019.116940, 2019.

Breiman, L.: Random Forests, Mach. Learn., 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Buchard, V., Randles, C. A., da Silva, A. M., Darmenov, A., Colarco, P. R., Govindaraju, R., Ferrare, R., Hair, J., Beyersdorf, A. J., Ziemba, L. D., and Yu, H.: The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part II: Evaluation and Case Studies, J. Climate, 30, 6851–6872, https://doi.org/10.1175/JCLI-D-16-0613.1, 2017.

Chen, Z., Chen, D., Zhao, C., Kwan, M.-P., Cai, J., Zhuang, Y., Zhao, B., Wang, X., Chen, B., Yang, J., Li, R., He, B., Gao, B., Wang, K., and Xu, B.: Influence of meteorological conditions on $PM_{2.5}$ concentrations across China: A review of methodology and mechanism, Environ. Int., 139, 105558, https://doi.org/10.1016/j.envint.2020.105558, 2020.

Chowdhury, S., Dey, S., Di Girolamo, L., Smith, K. R., Pillarisetti, A., and Lyapustin, A.: Tracking ambient $PM_{2.5}$ build-up in Delhi national capital region during the dry season over 15 years using a high-resolution (1 km) satellite aerosol dataset, Atmos. Environ., 204, 142–150, https://doi.org/10.1016/j.atmosenv.2019.02.029, 2019.

Dandona, L., Dandona, R., Kumar, G. A., Shukla, D., Paul, V. K., Balakrishnan, K., Prabhakaran, D., Tandon, N., Salvi, S., and Dash, A.: Nations within a nation: variations in epidemiological transition across the states of India, 1990–2016 in the Global Burden of Disease Study, Lancet, 390, 2437–2460, 2017.

Dey, S., Purohit, B., Balyan, P., Dixit, K., Bali, K., Kumar, A., Imam, F., Chowdhury, S., Ganguly, D., Gargava, P., and Shukla, V. K.: A Satellite-Based High-Resolution (1 km) Ambient $PM_{2.5}$ Database for India over Two Decades (2000–2019): Applications for Air Quality Management, Remote Sens.-Basel, 12, 3872, https://doi.org/10.3390/rs12233872, 2020.

Ganguly, T., Selvaraj, K. L., and Guttikunda, S. K.: National Clean Air Programme (NCAP) for Indian cities: Review and outlook of clean air action plans, Atmospheric Environment: X, 8, 100096, https://doi.org/10.1016/j.aeaoa.2020.100096, 2020.

Geiss, A., Silva, S. J., and Hardin, J. C.: Downscaling atmospheric chemistry simulations with physically consistent deep learning, Geosci. Model Dev., 15, 6677–6694, https://doi.org/10.5194/gmd-15-6677-2022, 2022.

Geng, G., Xiao, Q., Liu, S., Liu, X., Cheng, J., Zheng, Y., Xue, T., Tong, D., Zheng, B., and Peng, Y.: Tracking Air Pollution in China: Near Real-Time $PM_{2.5}$ Retrievals from Multisource Data Fusion, Environ. Sci. Technol., 55, 12106–12115, 2021.

Grinsztajn, L., Oyallon, E., and Varoquaux, G.: Why do tree-based models still outperform deep learning on typical tabular data?, Adv. Neur. In., 35, 507–520, 2022.

Gueymard, C. A. and Yang, D.: Worldwide validation of CAMS and MERRA-2 reanalysis aerosol optical depth products using 15 years of AERONET observations, Atmos. Environ., 225, 117216, https://doi.org/10.1016/j.atmosenv.2019.117216, 2020.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hu, X., Chu, L., Pei, J., Liu, W., and Bian, J.: Model complexity of deep learning: a survey, Knowl. Inf. Syst., 63, 2585–2619, https://doi.org/10.1007/s10115-021-01605-0, 2021.

Huang, C., Hu, J., Xue, T., Xu, H., and Wang, M.: High-Resolution Spatiotemporal Modeling for Ambient $PM_{2.5}$ Exposure Assessment in China from 2013 to 2019, Environ. Sci. Technol., 55, 2152–2162, 2021.

Huang, J., Zhou, Y., and Yong, W.-A.: Data-driven discovery of multiscale chemical reactions governed by the law of mass action, J. Comput. Phys., 448, 110743, https://doi.org/10.1016/j.jcp.2021.110743, 2022.

Huang, Y. and Seinfeld, J. H.: A neural network-assisted Euler integrator for stiff kinetics in atmospheric chemistry, Environ. Sci. Technol., 56, 4676–4685, 2022.

Jabbar, H. and Khan, R. Z.: Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study), Computer Science, Communication and Instrumentation Devices, 70, 978–981, 2015.

Katoch, V., Kumar, A., Imam, F., Sarkar, D., Knibbs, L. D., Liu, Y., Ganguly, D., and Dey, S.: Addressing Biases in Ambient $PM_{2.5}$ Exposure and Associated Health Burden Estimates by Filling Satellite AOD Retrieval Gaps over India, Environ. Sci. Technol., 57, 19190–19201, https://doi.org/10.1021/acs.est.3c03355, 2023.

Ke, G. L., Meng, Q., Finley, T., Wang, T. F., Chen, W., Ma, W. D., Ye, Q. W., and Liu, T. Y.: LightGBM: A Highly Efficient Gradient Boosting Decision Tree, 31st Annual Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, Dec 04-09, WOS:000452649403021, 2017.

Kumar, V., Malyan, V., Sahu, M., Biswal, B., Pawar, M., and Dev, I.: Spatiotemporal analysis of fine particulate matter for India (1980–2021) from MERRA-2 using ensemble machine learning, Atmos. Pollut. Res., 14, 101834, https://doi.org/10.1016/j.apr.2023.101834, 2023.

Kumari, S., Verma, N., Lakhani, A., and Kumari, K. M.: Severe haze events in the Indo-Gangetic Plain during post-monsoon: Synergetic effect of synoptic meteorology and crop residue burning emission, Sci. Total Environ., 768, 145479, https://doi.org/10.1016/j.scitotenv.2021.145479, 2021.

Li, H., Yang, Y., Wang, H., Li, B., Wang, P., Li, J., and Liao, H.: Constructing a spatiotemporally coherent long-term $PM_{2.5}$ concentration dataset over China during 1980–2019 using a machine learning approach, Sci. Total Environ., 765, 144263, https://doi.org/10.1016/j.scitotenv.2020.144263, 2021.

Li, T., Zhang, Q., Peng, Y., Guan, X., Li, L., Mu, J., Wang, X., Yin, X., and Wang, Q.: Contributions of Various Driving Factors to Air Pollution Events: Interpretability Analysis from Machine Learning Perspective, Environ. Int., 173, 107861, https://doi.org/10.1016/j.envint.2023.107861, 2023.

Liang, W., Luo, S., Zhao, G., and Wu, H.: Predicting hard rock pillar stability using GBDT, XGBoost, and LightGBM algorithms, Mathematics, 8, 765, https://doi.org/10.3390/math8050765, 2020.

Ma, Z., Dey, S., Christopher, S., Liu, R., Bi, J., Balyan, P., and Liu, Y.: A review of statistical methods used for developing large-scale and long-term $PM_{2.5}$ models from satellite data, Remote Sens. Environ., 269, 112827, https://doi.org/10.1016/j.rse.2021.112827, 2022.

Maheshwarkar, P., Ralhan, A., Sunder Raman, R., Tibrewal, K., Venkataraman, C., Dhandapani, A., Kumar, R. N., Mukherjee, S., Chatterje, A., Rabha, S., Saikia, B. K., Bhardwaj, A., Chaudhary, P., Sinha, B., Lokhande, P., Phuleria, H. C., Roy, S., Imran, M., Habib, G., Azharuddin Hashmi, M., Qureshi, A., Qadri, A. M., Gupta, T., Lian, Y., Pandithurai, G., Prasad, L., Murthy, S., Deswal, M., Laura, J. S., Chhangani, A. K., Najar, T. A., and Jehangir, A.: Understanding the Influence of Meteorology and Emission Sources on $PM_{2.5}$ Mass Concentrations Across India: First Results From the COALESCE Network, J. Geophys. Res.-Atmos., 127, e2021JD035663, https://doi.org/10.1029/2021JD035663, 2022.

Maji, K. J., Namdeo, A., and Bramwell, L.: Driving factors behind the continuous increase of long-term $PM_{2.5}$-attributable health burden in India using the high-resolution global datasets from 2001 to 2020, Sci. Total Environ., 866, 161435, https://doi.org/10.1016/j.scitotenv.2023.161435, 2023.

Martin, R. V., Brauer, M., van Donkelaar, A., Shaddick, G., Narain, U., and Dey, S.: No one knows which city has the highest concentration of fine particulate matter, Atmospheric Environment: X, 3, 100040, https://doi.org/10.1016/j.aeaoa.2019.100040, 2019.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Murray, C. J., Aravkin, A. Y., Zheng, P., Abbafati, C., Abbas, K. M., Abbasi-Kangevari, M., Abd-Allah, F., Abdelalim, A., Abdollahi, M., and Abdollahpour, I. J. T. L.: Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019, Lancet, 396, 1223–1249, 2020.

Nagpure, A. S., Ramaswami, A., and Russell, A.: Characterizing the spatial and temporal patterns of open burning of municipal solid waste (MSW) in Indian cities, Environ. Sci. Technol., 49, 12904–12912, 2015.

Navinya, C. D., Vinoj, V., and Pandey, S. K.: Evaluation of $PM_{2.5}$ Surface Concentrations Simulated by NASA's MERRA Version 2 Aerosol Reanalysis over India and its Relation to the Air Quality Index, Aerosol Air Qual. Res., 20, 1329–1339, https://doi.org/10.4209/aaqr.2019.12.0615, 2020.

Ni, Y., Yang, Y., Wang, H., Li, H., Li, M., Wang, P., Li, K., and Liao, H.: Contrasting changes in ozone during 2019–2021 between eastern and the other regions of China attributed to anthropogenic emissions and meteorological conditions, Sci. Total Environ., 908, 168272, https://doi.org/10.1016/j.scitotenv.2023.168272, 2024.

Pandey, A., Sadavarte, P., Rao, A. B., and Venkataraman, C.: Trends in multi-pollutant emissions from a technology-linked inventory for India: II. Residential, agricultural and informal industry sectors, Atmos. Environ., 99, 341–352, https://doi.org/10.1016/j.atmosenv.2014.09.080, 2014.

Pandey, A., Brauer, M., Cropper, M. L., Balakrishnan, K., Mathur, P., Dey, S., Turkgulu, B., Kumar, G. A., Khare, M., and Beig, G.: Health and economic impact of air pollution in the states of India: the Global Burden of Disease Study 2019, The Lancet Planetary Health, 5, e25–e38, 2021.

Pant, P., Lal, R. M., Guttikunda, S. K., Russell, A. G., Nagpure, A. S., Ramaswami, A., and Peltier, R. E.: Monitoring particulate matter in India: recent trends and future outlook, Air Qual. Atmos. Hlth., 12, 45–58, 2019.

Ren, J., Zhang, M., Yu, C., and Liu, Z.: Balanced mse for imbalanced visual regression, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 21–24 June 2022, New Orleans Ernest N. Morial Convention Center, New Orleans, Louisiana, 7926–7935, https://arxiv.org/abs/2203.16427 (last access: 30 July 2024), 2022.

Ren, X., Mi, Z., Cai, T., Nolte, C. G., and Georgopoulos, P. G.: Flexible Bayesian Ensemble Machine Learning Framework for Predicting Local Ozone Concentrations, Environ. Sci. Technol., 56, 3871–3883, https://doi.org/10.1021/acs.est.1c04076, 2022.

Sayeed, A., Lin, P., Gupta, P., Tran, N. N. M., Buchard, V., and Christopher, S.: Hourly and Daily $PM_{2.5}$ Estimations Using MERRA-2: A Machine Learning Approach, Earth Space Sci., 9, e2022EA002375, https://doi.org/10.1029/2022EA002375, 2022.

Shi, G., Lu, X., Deng, Y., Urpelainen, J., Liu, L.-C., Zhang, Z., Wei, W., and Wang, H.: Air pollutant emissions induced by population migration in China, Environ. Sci. Technol., 54, 6308–6318, 2020.

Stirnberg, R., Cermak, J., Kotthaus, S., Haeffelin, M., Andersen, H., Fuchs, J., Kim, M., Petit, J.-E., and Favez, O.: Meteorology-driven variability of air pollution (PM_1) revealed with explainable machine learning, Atmos. Chem. Phys., 21, 3919–3948, https://doi.org/10.5194/acp-21-3919-2021, 2021.

Sun, X., Liu, M., and Sima, Z.: A novel cryptocurrency price trend forecasting model based on LightGBM, Financ. Res. Lett., 32, 101084, https://doi.org/10.1016/j.frl.2018.12.032, 2020.

Tiwari, S., Srivastava, A. K., Bisht, D. S., Parmita, P., Srivastava, M. K., and Attri, S.: Diurnal and seasonal variations of black carbon and $PM_{2.5}$ over New Delhi, India: Influence of meteorology, Atmos. Res., 125, 50–62, 2013.

Upadhyay, A., Dey, S., and Goyal, P.: A comparative assessment of regional representativeness of EDGAR and ECLIPSE emission inventories for air quality studies in India, Atmos. Environ., 223, 117182, https://doi.org/10.1016/j.atmosenv.2019.117182, 2020.

van Donkelaar, A., Hammer, M. S., Bindle, L., Brauer, M., Brook, J. R., Garay, M. J., Hsu, N. C., Kalashnikova, O. V., Kahn, R. A., Lee, C., Levy, R. C., Lyapustin, A., Sayer, A. M., and Martin, R. V.: Monthly Global Estimates of Fine Particulate Matter and Their Uncertainty, Environ. Sci. Technol., 55, 15287–15300, https://doi.org/10.1021/acs.est.1c05309, 2021.

Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Abbasi, M., Abbasifard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., and Abdelalim, A.: Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019, Lancet, 396, 1204–1222, 2020.

Wang, K. C., Dickinson, R. E., Wild, M., and Liang, S.: Atmospheric impacts on climatic variability of surface incident solar radiation, Atmos. Chem. Phys., 12, 9581–9592, https://doi.org/10.5194/acp-12-9581-2012, 2012.

Wang, S., Kota, S. H., and Zhang, H.: LongPMInd: long-term (1980–2022) daily ground particulate matter datasets in India, Zenodo [data set], https://doi.org/10.5281/zenodo.10073944, 2023a.

Wang, S., Wang, P., Qi, Q., Wang, S., Meng, X., Kan, H., Zhu, S., and Zhang, H.: Improved estimation of particulate matter in China based on multisource data fusion, Sci. Total Environ., 869, 161552, https://doi.org/10.1016/j.scitotenv.2023.161552, 2023b.

Wang, S., Wang, P., Zhang, R., Meng, X., Kan, H., and Zhang, H.: Estimating particulate matter concentrations and meteorological contributions in China during 2000–2020, Chemosphere, 330, 138742, https://doi.org/10.1016/j.chemosphere.2023.138742, 2023c.

Wang, S., Zhang, M., Gao, Y., Wang, P., Fu, Q., and Zhang, H.: Diagnosing drivers of $PM_{2.5}$ simulation biases from meteorology, chemical composition, and emission sources using an efficient machine learning method, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-1531, 2023d.

Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., and Cribb, M.: Reconstructing 1-km-resolution high-quality $PM_{2.5}$ data records from 2000 to 2018 in China: spatiotemporal variations and policy implications, Remote Sens. Environ., 252, 112136, https://doi.org/10.1016/j.rse.2020.112136, 2021a.

Wei, J., Li, Z., Pinker, R. T., Wang, J., Sun, L., Xue, W., Li, R., and Cribb, M.: Himawari-8-derived diurnal variations in ground-level $PM_{2.5}$ pollution across China using the fast space-time Light Gradient Boosting Machine (LightGBM), Atmos. Chem. Phys., 21, 7863–7880, https://doi.org/10.5194/acp-21-7863-2021, 2021b.

Wei, J., Li, Z., Lyapustin, A., Wang, J., Dubovik, O., Schwartz, J., Sun, L., Li, C., Liu, S., and Zhu, T.: First close insight into global daily gapless 1 km $PM_{2.5}$ pollution, variability, and health impact, Nat. Commun., 14, 8349, https://doi.org/10.1038/s41467-023-43862-3, 2023.

Yan, J., Xu, Y., Cheng, Q., Jiang, S., Wang, Q., Xiao, Y., Ma, C., Yan, J., and Wang, X.: LightGBM: accelerated genomically designed crop breeding through ensemble learning, Genome Biol., 22, 1–24, 2021.

Yang, X., Zhao, C. F., Zhou, L. J., Wang, Y., and Liu, X. H.: Distinct impact of different types of aerosols on surface solar radiation in China, J. Geophys. Res.-Atmos., 121, 6459–6471, https://doi.org/10.1002/2016jd024938, 2016.

Yang, Y., Zha, K., Chen, Y., Wang, H., and Katabi, D.: Delving into deep imbalanced regression, International Conference on Machine Learning, 18–24 July 2021, Honolulu, Hawaii, 65 USA, 11842–11851, https://arxiv.org/abs/2102.09554 (last access: 30 July 2024), 2021.

Ying, X.: An overview of overfitting and its solutions, J. Phys. Conf. Ser., 1168, 022022, https://doi.org/10.1088/1742-6596/1168/2/022022, 2019.

Yu, W., Ye, T., Zhang, Y., Xu, R., Lei, Y., Chen, Z., Yang, Z., Zhang, Y., Song, J., and Yue, X.: Global estimates of daily ambient fine particulate matter concentrations and unequal spatiotemporal distribution of population exposure: a machine learning modelling study, The Lancet Planetary Health, 7, e209–e218, 2023.

Zhang, T. N., He, W. H., Zheng, H., Cui, Y. P., Song, H. Q., and Fu, S. L.: Satellite-based ground $PM_{2.5}$ estimation using a gradient boosting decision tree, Chemosphere, 268, 128801, https://doi.org/10.1016/j.chemosphere.2020.128801, 2021.

FOR THE PEER REVIEWER

How this record was made

Everything a statistician or an atmospheric scientist needs to judge this report, on one page: the two sources, the join, the de-weathering, the tests and what they cannot cover.

The two sources

Years 1980 to 2022 come from LongPMInd, a published daily PM2.5 reconstruction for India on a ten-kilometre grid, built with gradient-boosted trees on CPCB ground measurements, satellite aerosol optical depth, MERRA-2 composition and ERA5 weather. Its reported skill on whole years withheld from training is about R^2 0.66, BUT that test runs only inside 2018 to 2022, which are the only years it was trained on, so it measures interpolating one year from its four neighbours and says nothing about 1985. Years 2023 to July 2026 come from Delhi's own CPCB and DPCC continuous monitors.

The city day

A station-day is retained with at least sixteen valid hours and a value between 0 and 1000 micrograms. The city day is the unweighted mean of qualifying station-days, and a month is reported only with at least twenty valid city-days.

The join

Four mappings were fitted on 2018 to 2020 and scored on 2021 to 2022, out of time. The winner was re-fitted on the full 2018 to 2022 overlap.

$$\hat{L}_m = 5.805 + 0.8615 \cdot O_m$$

O is the monitor city monthly mean, L̂ the same month on the reconstruction's scale; n = 60 months, r = 0.996

Held-out error 4.73 micrograms against 13.42 for the uncorrected series. Anchoring on 2019 to 2020 and carrying to 2022 drifts +2.5 per cent, which is the band that belongs on the extension.

The de-weathering

Monthly meteorological normalisation. Each variable is expressed as an anomaly from its own calendar-month mean over the whole record, the PM anomaly is regressed on the weather anomalies, and the weather-explained part is subtracted, so the trend, the level and the seasonal cycle survive.

$$P_t^{dw} = P_t - \tilde{X}_t\hat{\beta}$$

$\tilde{X}$ are anomalies of wind speed, the two wind components, mixing height, temperature, humidity, rainfall and surface pressure; R^2 = 0.125

The tests

Trends are ordinary least squares on annual means. Every trend this report rests an argument on is printed with a 95 per cent interval as well as a p-value, because a p-value says only whether zero can be excluded while an interval says how large the change might be. The start-year bars in Figure 4 and the within-era column of the table on page 4 are sensitivity displays rather than claims, and carry p-values only. Where a series runs fifteen years or more and its residuals carry lag-one correlation above 0.2, the interval is widened for that autocorrelation on the Weatherhead adjustment.

The standard error is multiplied by the square root of (1 plus φ) over (1 less φ), and the degrees of freedom become n(1 less φ)/(1 plus φ), less two.

φ is the lag-one residual correlation. Both conditions must hold, because on fewer than fifteen points φ is itself noise.

Newey-West standard errors are NOT used on these annual series. On five and eight point windows they do not widen the interval, they narrow it until zero falls outside, which would manufacture a significant improvement out of noise. The recent trend is additionally reported from every start year between 2008 and 2018, because a single start year is a choice and this record's peak falls at 2018.

What these intervals do not cover

They price sampling and fitting error alone. They do not price the reconstruction's own model error in the years before the monitors, the change of measurement basis at the join, monitor siting, or the difference between an area average and a network of addresses. The record's direction is robust to all four; a level quoted to better than a few micrograms is not.

Sources: read 11 August 2026. LongPMInd (Wang et al., ESSD 2024) for 1980 to 2022; CPCB and DPCC continuous monitors via CREA for 2023 to July 2026; ERA5 and MERRA-2 reanalyses. Series files, the splice statistics, the repair log and every script are open to a reviewer on request. Analysis in-house for the Office of Ajay Maken, MP.