

Coal plants persist as a large barrier to the global solar energy transition

Received: 9 July 2025

Accepted: 14 April 2026

Published online: 15 May 2026

 Check for updates

Rui Song^{1,2}✉, Feng Yin³, Jan-Peter Muller²✉, Adam C. Povey⁴,
Basudev Swain¹, Chenchen Huang⁵ & Roy G. Grainger¹✉

The global energy transition depends on solar photovoltaic (PV) power displacing fossil fuels to deliver projected climate and air quality benefits. However, aerosol pollution from co-located coal plants actively suppresses PV energy production. Here a global, facility-level dataset shows that aerosols reduced global PV generation by 5.8% in 2023 (111 TWh). From 2017 to 2023, annual aerosol-induced PV energy losses from existing systems were, on average, equivalent to one-third of the energy added by new PV installations. In China, aerosols caused the largest PV energy losses worldwide, reducing national PV generation by 7.7% in 2023. The corresponding annual loss-to-growth ratio averaged 38% and frequently exceeded 50%. Despite continued coal expansion, PV energy losses have declined by 1.4% yr⁻¹ since 2017 owing to stricter emission controls. By contrast, the USA, where co-location of solar and coal plants is limited, experienced only 3.1% aerosol-induced PV loss. Given the slow pace of global coal phase-out, these results reveal a constraint on solar performance that, if unaccounted for, could lead to a systematic overestimation of the transition's contribution to climate and air quality goals.

Limiting global warming to 1.5 °C requires a 45% reduction in greenhouse gas emissions from 2010 levels by 2030 and net-zero emissions by 2050^{1,2}. This demands a rapid transition from fossil fuels to renewable energy. Since 2011, renewables have expanded at an average annual rate of 6.1%, supplying 29.1% (8,440 TWh) of global electricity in 2023³. Expansion potential varies across technologies: hydropower⁴, geothermal⁵ and tidal⁶ energy face environmental and geographic constraints, whereas solar and wind offer greater scalability with fewer such obstacles⁷. Solar photovoltaic (PV) technology has been leading this transition, driven by advances in cell materials⁸, automated manufacturing⁹, large-scale deployment¹⁰ and sharply declining costs. Over the past decade, PV module prices have declined sharply, making it the most cost-effective electricity source¹¹, supported by subsidies, tax incentives and favourable trade policies¹². In 2023, PV accounted for 75% of the 510-GW increase in global renewable capacity¹³. This

rapid growth has prompted projections of substantial climate and air quality benefits, assuming that new PV capacity displaces coal-fired generation^{14,15}. Yet, the extent to which renewables displace fossil fuels in practice remains unclear^{16,17}. Clarifying whether PV expansion translates into fossil fuel displacement and delivers the anticipated benefits is essential for tracking progress towards net-zero targets.

Solar PV offers a low-carbon energy pathway, with life-cycle greenhouse gas emissions that are an order of magnitude lower than those of coal-fired generation equipped with carbon capture and storage¹⁸. From 2009 to 2019, global PV deployment avoided an estimated 1.3 Gt of carbon dioxide (CO₂) emissions. Meeting 40% of electricity demand with PV from 2020 to 2060 could mitigate up to 205 Gt (ref. 19). Rooftop PV alone could reduce global temperatures by 0.05–0.13 °C by 2050 through avoided emissions¹⁴. Substituting fossil fuels with PV reduces air pollutants such as sulfur dioxide (SO₂), nitrogen oxides (NO_x) and

¹National Centre for Earth Observation, Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, UK. ²Mullard Space Science Laboratory, Department of Space and Climate Physics, University College London, Holmbury St Mary, Surrey, UK. ³National Centre for Earth Observation, Department of Geography, University College London, London, UK. ⁴National Centre for Earth Observation, School of Physics and Astronomy, University of Leicester, Leicester, UK. ⁵School of Management, University of Bath, Bath, UK. ✉e-mail: rui.song@physics.ox.ac.uk; j.muller@ucl.ac.uk; don.grainger@physics.ox.ac.uk

fine particulate matter (PM_{2.5})^{20,21}, bringing immediate health benefits²², particularly in regions with high baseline pollution¹⁵. These climate and air quality gains depend on the extent to which solar electricity substitutes for rather than supplements fossil-fuel-based generation. While full substitution maximizes mitigation, even apparent supplementation can confer benefits by offsetting the expansion in fossil fuel use that would otherwise accompany rising energy demand. Conversely, if fossil fuel use remains unchanged or expands alongside renewables, the transition's effective mitigation potential is reduced²³. In practice, this displacement process has proven highly inefficient, with evidence suggesting that globally, over six units of renewable energy may be needed to displace one unit of fossil energy¹⁷. This inefficiency largely reflects rising overall energy demand and systemic dependence on fossil-fuel infrastructure, which reduces the substitution between renewable and fossil energy. This dependence is further reinforced by persistent policy incentives for fossil fuels: despite frequent pledges, fossil fuel subsidy reforms since 2016 have often proved short lived and, in most major subsidising countries, support for coal has either remained unchanged or increased²⁴.

This coal resurgence not only sustains emissions but also impairs solar performance by degrading air quality as the atmospheric emissions from coal-fired power generation directly reduce surface irradiance. When these coal plants are brought online as backup during periods of low solar output, their emissions can further intensify pollution and prolong dimming episodes, delaying the rebound of surface irradiance even when meteorological conditions improve²⁵. This reduction occurs as particles scatter and absorb incoming radiation^{26–28}, a direct effect that has been found to lower annual PV energy yields by over 20% in heavily polluted regions such as eastern China and northern India^{26,29}. Aerosols also modify cloud microphysics, altering cloud reflectivity and coverage, and further suppress surface irradiance through indirect effects²⁹. These losses risk disrupting a positive feedback loop, in which solar deployment should reduce fossil fuel reliance and improve air quality, which in turn enhances PV performance to reinforce the energy transition^{15,26}. This mechanism presents a diagnostic opportunity: trends in aerosol-induced PV losses may indicate whether solar is displacing coal or merely expanding alongside it. Detecting this signal remains challenging. Most existing studies^{26–29} analyse broad irradiance patterns without accounting for spatial and temporal variation in PV deployment. Such approaches are insufficient because PV siting reflects not only solar resource availability but also policy incentives, infrastructure and land-use constraints^{30,31}. Capturing the signal of fossil displacement therefore requires a spatially explicit, facility-level analysis that links atmospheric conditions to actual trends in PV energy generation and losses.

The application of machine learning to high-resolution Earth observation data, such as from Sentinel-2, has produced global¹⁰ and national inventories of PV systems^{32–34}. While these datasets successfully map the spatial distribution of facilities, their utility for tracking energy trends is fundamentally limited by ambiguity in the detected panel footprints. This ambiguity can result in estimation errors of up to a factor of two and varies substantially with surface type (Supplementary Figs. 2 and 3). Detection is most accurate in deserts, where dark panels contrast sharply with bright terrain. Accuracy decreases in mountainous regions, where PV arrays follow complex topography, and in heterogeneous land-use areas such as croplands, fishponds and greenhouses, where spectral mixing is more common. When aggregated over large areas, footprint and classification errors introduce regional biases to estimates of PV generation. These mapping inaccuracies are a primary obstacle to quantifying PV energy generation and losses from aerosols, particularly those associated with fossil fuel combustion, thus obscuring the signal needed to detect fossil fuel displacement.

To overcome this limitation, a three-step framework to generate a global dataset tracking facility-level PV energy generation and calculate

losses driven by clouds and aerosols (Fig. 1) has been developed. The framework achieves comprehensive coverage of PV installations that were operational as of early 2024 by identifying candidate sites from a combination of existing inventories, crowd-sourced data and a custom machine learning model applied to global satellite imagery. It produces unprecedented spatial accuracy through a dedicated segmentation model, the Segment Anything Model (SAM)³⁵, to precisely extract PV facility footprints. This facilitates the reliable estimation of facility-level energy generation and the losses from clouds and aerosols by integrating the footprints with atmospheric reanalysis data and a validated PV model. This high-fidelity dataset allows the quantification of aerosol-induced energy losses as a diagnostic signal of interaction between solar deployment and fossil fuel displacement. The signal is used to evaluate whether, where and to what extent solar expansion is offsetting coal-fired generation, and to assess whether the projected climate and air quality benefits of the solar transition are being realized in practice.

Results

Global database of facility-level PV generation and losses

The three-step approach (Methods) produced a global dataset of 140,945 PV facilities, each with facility-level estimates of energy generation and losses from clouds and aerosols, with aerosol-induced losses representing the direct radiative effects of aerosols suspended in the atmosphere. In 2023, these data show that average PV energy losses reached 26.9% of potential output under optimal conditions, comprising 21.1% from clouds and 5.8% from aerosols. For global generation of 1,911 TWh, this amounts to a global energy loss of 515 TWh, equivalent to the annual output of 84 medium-sized (1 GW) coal plants operating at a typical capacity of 70%. Loss rates varied regionally (Supplementary Fig. 6 and Supplementary Tables 3–5) due to differences in climate and pollution; China, the world's largest PV generator, experienced a 28.0% total loss rate, of which 7.7% was from aerosols, compared with a 22.7% total loss with a 3.1% aerosol loss in the USA. At the facility level, a single site in Qinghai Province, China (36.09° N, 100.49° E), lost 3.6 TWh in 2023 alone, equivalent to 0.7% of global PV output losses.

This facility-level dataset overcomes a key limitation of earlier assessments that relied on spatially uniform assumptions of PV distribution. The impact of this methodological improvement is substantial. For instance, previous studies estimated that aerosols reduce PV output in China by as much as 20–25% (refs. 26,29), as those analyses assessed potential solar irradiance without considering the actual distribution of PV facilities. By contrast, constrained by the precise locations of operational systems, our analysis finds a national average loss of 7.7%. By accurately linking PV footprints to electricity generation, our facility-level data resolves this overestimate and provides a robust foundation for assessing energy trends and detecting the signal of fossil fuel displacement.

High loss-to-growth ratios constrain net solar energy gains

Aerosol-induced losses from the existing PV fleet are equivalent to nearly one-third of the annual energy generated by new PV installations globally, representing a magnitude not previously quantified and unexpectedly high. While clouds are the dominant source of atmospheric reduction in insolation, aerosols have a disproportionate impact in densely populated, industrialized regions where PV deployment is concentrated. Between 2017 and 2023, new PV installations added 246.6 TWh of generation per year on average, while aerosol-related losses from existing systems reached 74.0 TWh annually. To compare these quantities across regions and years, we use the loss-to-growth ratio (R_{LG}), defined as the annual aerosol-induced energy loss from existing systems divided by the annual energy generated from new PV capacity (Fig. 2). Globally, R_{LG} averaged 30.0% during the study period (Fig. 2b), varying substantially across regions (Fig. 2c–e and

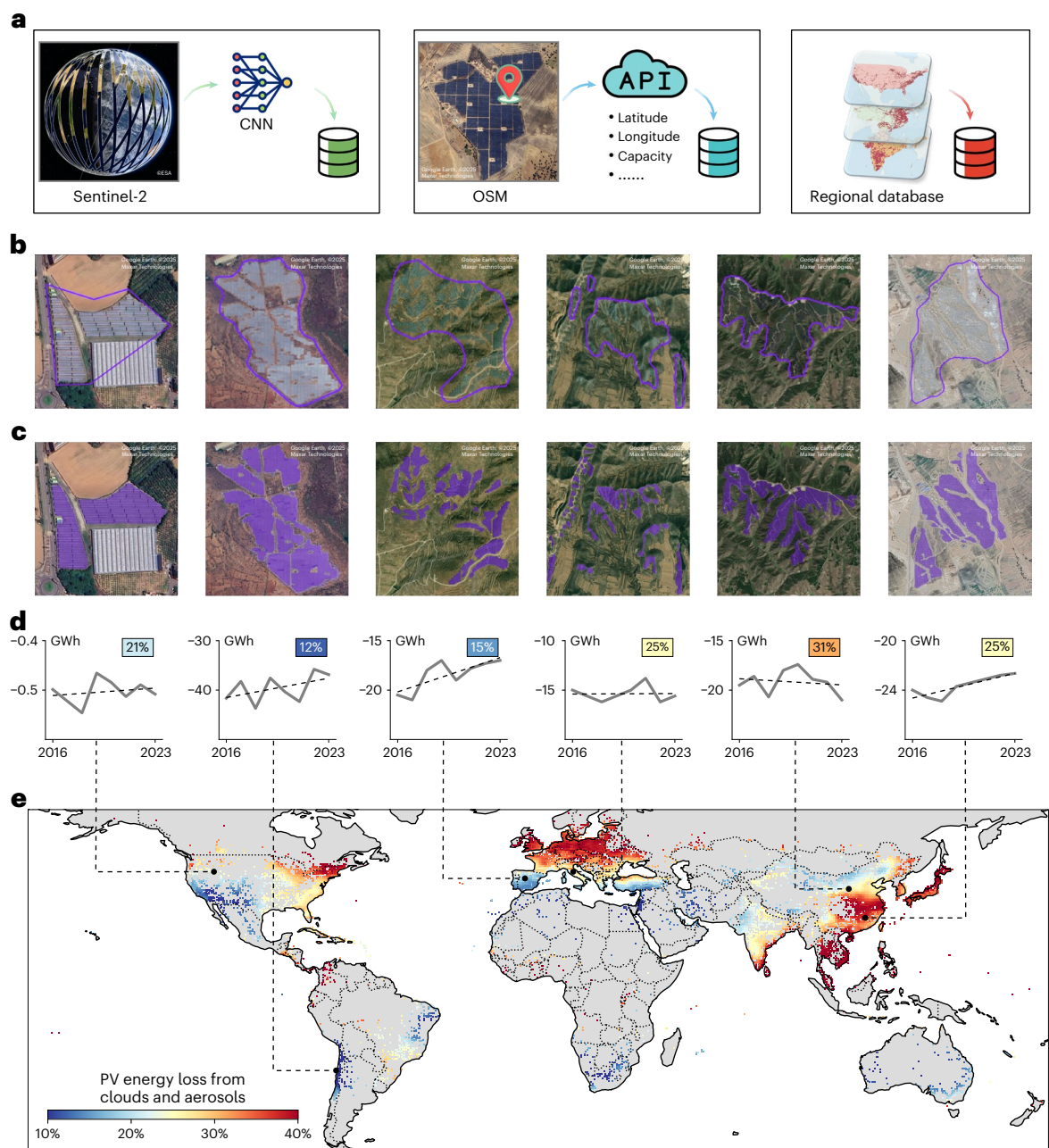


Fig. 1 | A three-step framework for tracking global facility-level PV energy generation and climate-related losses from clouds and aerosols. a, Step 1: global identification of PV facilities. An initial inventory is created by integrating OSM data, regional databases^{10,32,33} and a custom-trained CNN applied to global-scale Sentinel-2 imagery. API, application programming interface. **b**, Examples of initial PV polygons (inventory updated to early 2024) overlaid on Google Earth satellite imagery; polygon boundaries are shown as purple outlines, illustrating substantial false positives and boundary errors. **c**, Step 2: precise extraction of PV footprints.

For each facility, Meta's SAM is applied to high-resolution imagery to isolate the active panel arrays, shown as purple shading. **d**, Step 3: facility-level energy estimation. The improved footprints are integrated with reanalysis climate data in a validated energy model to estimate time series of electricity generation and the distinct losses from clouds and aerosols. Example plots show total percentage losses from 2016 onwards. **e**, A global map of PV energy losses from clouds and aerosols (%). This map visualizes the aggregated losses from clouds and aerosols across all facilities in 2023, illustrating one key output of the final database.

Supplementary Fig. 7); lower values indicate the delivery of more net energy to the grid and thus a greater contribution to decarbonization.

In China, aerosol-induced losses remain substantial, reducing national PV generation by 7.7% in 2023 due to persistent pollution and dense PV deployment (Fig. 2f). In 2023, the nation produced 793.5 TWh of PV electricity, representing 41.5% of the global total, while its 61.3 TWh of aerosol-induced losses accounted for a disproportionate 54.9% of the global total, exceeding that of all other countries combined. The North China Plain continues to be the primary hotspot³⁶,

but increasing losses are also observed in western and southern regions as PV installations expand geographically. Over the full study period, China's annual R_{LC} averaged 38.4%, exceeding 50% in three consecutive years and peaking at 62.1% in 2021 before falling to 26.0% by 2023 (Fig. 2c). The years of high R_{LC} (2018–2021) coincided with slower PV deployment, initially triggered by subsidy reductions and regulatory uncertainty under the mid-2018 '531' policy³⁷, and later compounded by coronavirus disease 2019-related supply-chain disruptions in 2020³⁸. A subsequent policy-driven rebound in PV growth

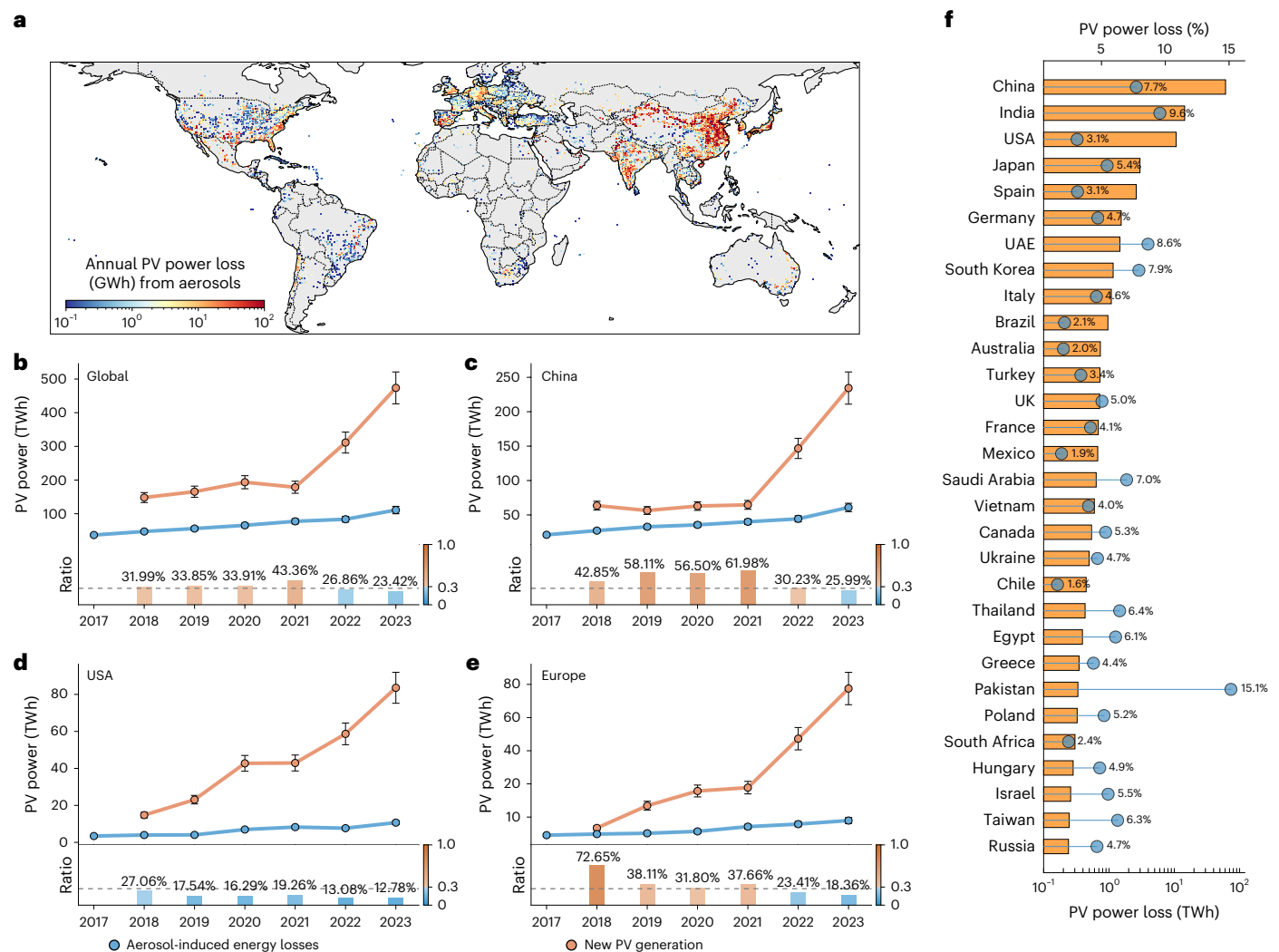


Fig. 2 | Aerosol-induced energy losses offset a substantial fraction of new PV generation globally. **a**, The global distribution of PV energy losses due to atmospheric aerosols in 2023, aggregated from facility-level simulations and mapped at $0.5^\circ \times 0.5^\circ$ resolution. **b–e**, The annual PV loss-to-growth ratios (R_{LG}) from 2017 to 2023 for the globe (**b**), China (**c**), the USA (**d**) and Europe, including the UK (**e**). In each panel, the lines represent annual energy from new PV additions (orange) versus aerosol-induced losses from existing systems (blue), with the

corresponding R_{LG} value for each year as a bar. In **b–e**, the annual totals are derived from facility-level simulations across independent PV facilities (global dataset $n = 140,945$ facilities). The error bars indicate $\pm 10\%$ relative model uncertainty associated with climate reanalysis and energy modelling. **f**, A country-level ranking of total aerosol-induced PV energy losses in 2023, showing absolute losses (TWh, bottom axis) and losses as a percentage of total PV generation (top axis).

from late 2021 improved national performance and reduced the R_{LG} accordingly.

In the USA, aerosol-induced losses are considerably lower than in other major PV regions, reducing national PV generation by 3.1% in 2023 (Fig. 2f). That year, the USA accounted for 18.1% of global PV generation but only 9.6% (10.7 TWh) of global total aerosol-induced losses (Fig. 2d). This smaller burden reflects both natural and anthropogenic factors, with overall aerosol loading lower than in other major PV regions. In the western USA, dust storms and more frequent wildfires reduce surface irradiance, though their impacts are episodic and spatially limited³⁹. By contrast, the eastern regions experience more persistent aerosol pollution from population centres and industrial activities. Although the West benefits from higher solar irradiance ($5\text{--}7\text{ kWh m}^{-2}\text{ day}^{-1}$), the East, with lower irradiance ($4\text{--}5\text{ kWh m}^{-2}\text{ day}^{-1}$), hosts a larger share of PV installations due to greater electricity demand and previous stronger policy incentives. Continued PV expansion in these relatively low-aerosol areas has kept the national R_{LG} consistently below the global average throughout the study period, reaching a minimum of 12.8% in 2023 (Fig. 2d). Europe (including the UK) shows

moderate losses, with an average annual R_{LG} of 28.4% (Fig. 2e), close to the global mean.

Limited evidence for PV air quality feedback beyond China

Paradoxically, while suffering the most from total aerosol-induced losses, China is the only major PV-producing region showing evidence of a sustained decline in these losses. This decline may indicate the early stages of a positive feedback between solar deployment and air quality improvement, whereby cleaner air conditions enhance solar performance and further support renewable expansion. The specific drivers of this trend are analysed in the next section. The annual aerosol-induced energy losses from 2013 to 2023 across a fixed network comprising all facilities operational by year-end 2023 (Fig. 3) is simulated to isolate the direct impact of aerosols from changes in PV network expansion. This approach enables an assessment of aerosol effects that is independent of PV network growth. The results confirm that, despite having the highest absolute losses, China is the only region exhibiting a consistent decline, with its losses falling by 0.96 TWh yr^{-1} ($\sim 1.4\%$ annually). By stark contrast, losses in the USA and Europe trended upwards by 0.15 TWh

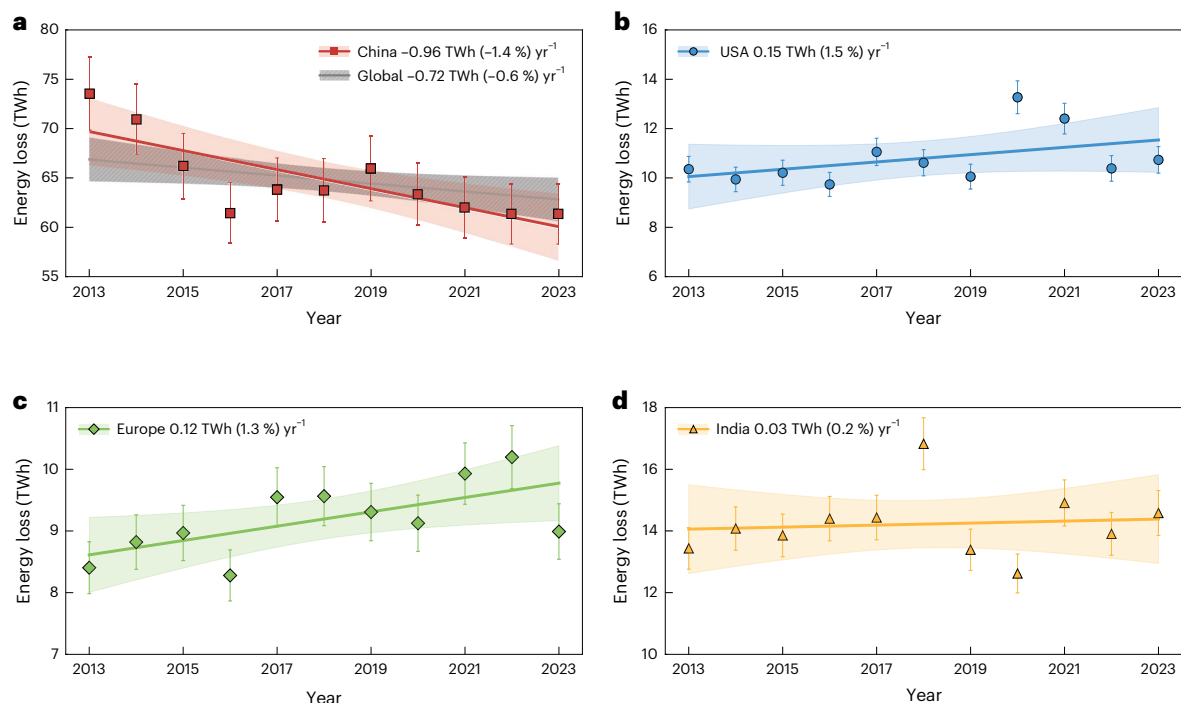


Fig. 3 | Only China shows a sustained decline in aerosol-induced PV energy loss. a–d, Standardized annual energy losses (in TWh) from aerosols between 2013 and 2023, estimated under fixed PV network conditions, for China (a), the USA (b), Europe, including the UK (c) and India (d), aggregated from facility-level datasets. The error bars indicate $\pm 5\%$ relative uncertainty associated with energy

modelling. The solid lines show linear regression fits; the shaded bands indicate the 95% confidence interval of the fitted trend. In a, the global trend, normalized to China's mean value, is shown as a grey hatched band. The use of a fixed network (based on all facilities operational by year-end 2023) isolates the impact of changing atmospheric conditions from the effects of network expansion.

(1.5%) and 0.12 TWh yr^{-1} (1.3%), respectively, despite their early adoption of renewable energy. India continues to experience persistently high losses due to severe air pollution, with no clear trend. The modest global decline of $-0.72 \text{ TWh yr}^{-1}$ (-0.6%) was therefore driven almost entirely by the improvements in China. These findings suggest that while this positive feedback is anticipated to accelerate the energy transition^{15,26}, this virtuous cycle has yet to materialize at a global scale, even during the recent period of rapid PV expansion.

Influence of coal emissions on solar PV generation

China's unique status as both the region most affected by aerosol-induced PV losses and the only one showing a sustained improvement warrants an investigation into the underlying drivers. Our analysis of aerosol composition, sector-specific attribution, temporal pollutant trends and spatial correlations identifies coal-fired power generation as the dominant factor. First, the composition of performance-reducing aerosols points to coal. Sulfate aerosols, formed through the oxidation of SO_2 (a primary coal emission), account for 46.2% (28.3 TWh annually) of total aerosol-related PV losses (Fig. 4a). Carbonaceous aerosols, also partly from coal combustion, contribute an additional 11.3 TWh (18.4%). While dust contributes about one-third of total losses, its impact is highly localized to desert regions with a smaller share of PV capacity. A regional breakdown confirms the pervasive influence of coal, showing that sulfate aerosols dominate losses across most of China: from 8.9 TWh yr^{-1} in the North China Plain to 9.1 TWh yr^{-1} in arid, dust-dominated regions (for example, Northwest and Inner Mongolia Plateau) (Fig. 4b). To further isolate the role of coal emissions, a sector-specific attribution experiment was conducted using the GEOS-Chem chemical transport model (Methods). This analysis shows that 29.0% of aerosol-induced PV energy losses in China can be attributed to coal-fired power plants, reinforcing the dominant influence of coal-related aerosols identified in Fig. 4.

Second, the observed decline in pollutants over China's PV sites does not reflect a coal phase-out, but rather the effectiveness of aggressive emission controls on a growing coal fleet⁴⁰. Coal-fired power plants are major sources of $\text{PM}_{2.5}$, SO_2 and NO_2 pollution in China^{41,42}, and our analysis of effective pollutant values (Methods) confirms that these stricter standards have had a measurable impact on the atmospheric pollutants that reduce solar generation. From 2013 to 2023, effective aerosol optical depth (AOD) declined by 0.0052 yr^{-1} (1.7% annually), tropospheric NO_2 by 9.1×10^{13} molecules per square centimetre per year (2.4%) and planetary boundary layer (PBL) SO_2 by 1.4×10^{14} molecules per square centimetre per year (1.4%) (Fig. 4c–e). This trend is paradoxical as it occurred while China's coal-fired electricity output increased from 4,093 TWh to 5,857 TWh⁴³. This was achieved through fleet-wide modernization involving the phasing out of older, inefficient plants^{44,45} while simultaneously accelerating the construction of new capacity, which reached its highest level in nearly a decade in 2024⁴⁶. To further quantify this dynamic, we analysed the relative contributions of ultralow-emission (ULE) retrofits and coal-plant retirements to the decline in pollutant emissions over China's PV regions. Using unit-level data on coal-plant capacities, retrofit progress and retirements, we reconstructed annual SO_2 emissions from 2014 to 2023 (Supplementary Note 3). The results show that ULE retrofits account for 91% of the total SO_2 reduction from the coal-power sector during this period, while coal retirements contribute 9%. These findings confirm that the observed improvements in atmospheric conditions above China's PV sites were driven primarily by emission-intensity reductions from widespread ULE adoption rather than by large-scale coal phase-out.

Third, the spatial distribution of PV losses in China mirrors that of its coal-fired power capacity⁴⁷. Facility-level data reveal widespread co-location of PV installations and coal plants, extending into western desert regions often perceived as dedicated hubs for renewable expansion (Fig. 5). The influence of coal in these areas is further confirmed by the composition of the aerosol-induced losses; even where dust is

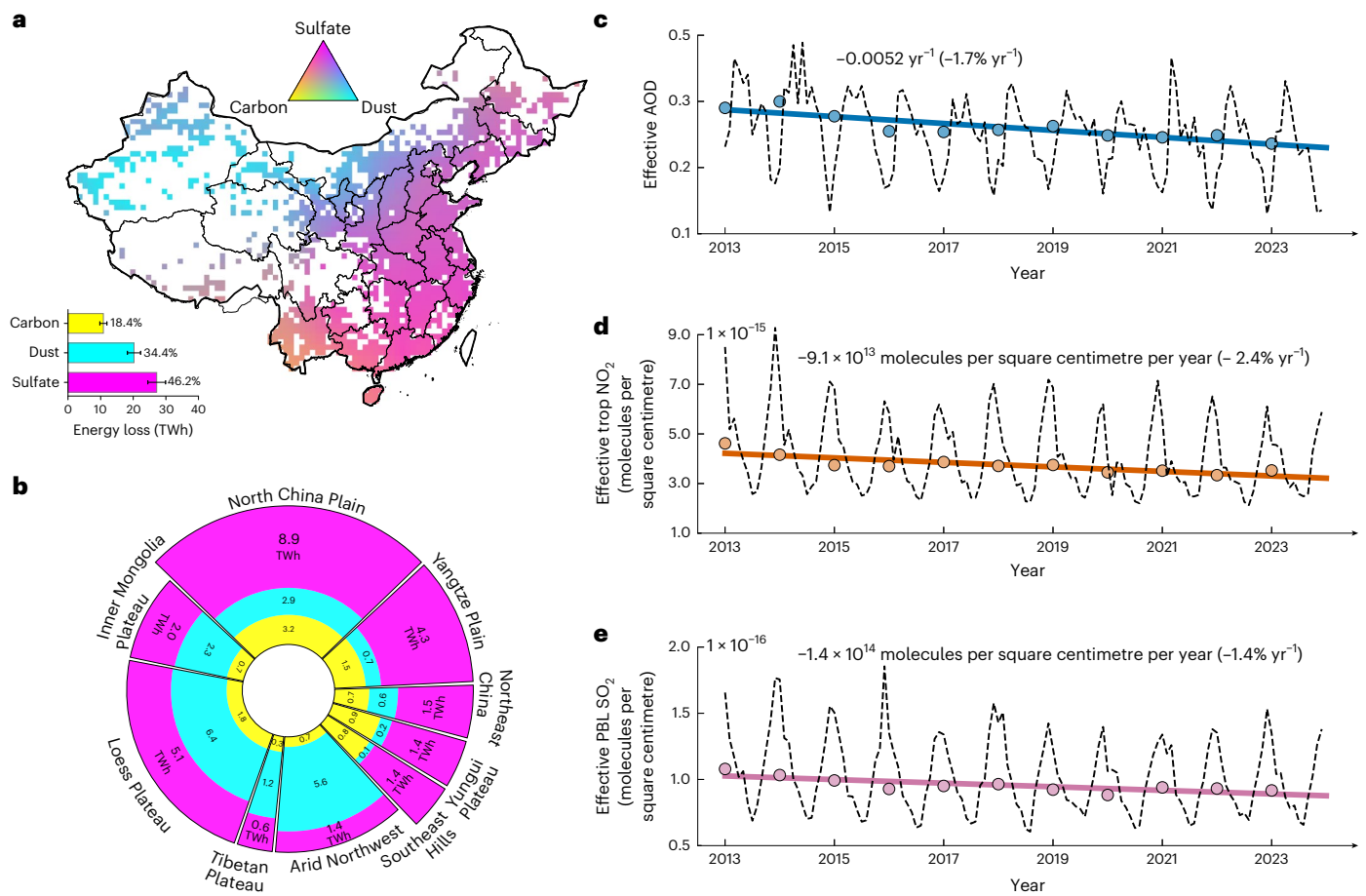


Fig. 4 | Coal-related pollutants dominate aerosol-induced PV losses in China but show declining trends. a, The contribution of major aerosol types to total PV energy losses. The losses are classified into sulfate, carbonaceous and dust aerosols, which together account for 99% of total aerosol-related losses, excluding only sea salt. The error bars indicate $\pm 10\%$ relative model uncertainty associated with climate reanalysis and energy modelling. **b**, The regional breakdown of PV output losses by aerosol type. The surface area of each slice of

the pie chart is proportional to its total aerosol-induced PV losses in that region, while its elongation indicates the dominance of sulfate aerosols, of which coal combustion is a main source. **c–e**, Trends in effective values of atmospheric pollutants at PV sites from 2013 to 2023, based on a fixed 2023 PV network. Consistent declines are shown in AOD (c), tropospheric NO_2 column amounts (d) and PBL SO_2 column amounts (e).

the dominant aerosol type, sulfate-driven losses remain high (Fig. 4b). To quantify this spatial relationship, we aggregated coal capacity and aerosol-induced PV losses to a $1.0^\circ \times 1.0^\circ$ grid and calculated the bivariate Moran's I statistic⁴⁸ on the top 30% of grid cells for each variable. This analysis yielded a significant positive correlation ($I = 0.5654$, $P = 0.007$), confirming that the highest-capacity coal regions tend to coincide with the areas of greatest PV loss. By contrast, the USA exhibits limited co-location of its solar and coal capacity (Extended Data Fig. 1). This geographical separation is also a key factor behind the nation's low R_{LC} , a finding supported by the corresponding Moran's I statistic, which shows no spatial correlation ($I \approx 0$). Supplementary Fig. 8 further quantifies this spatial relationship by showing that, across different ranges of annual generation ($E_{\text{year}} < 10$ GWh, 10–100 GWh and > 100 GWh), PV facilities in China are consistently located much closer to coal-fired power plants, with distance distributions peaking at 20–30 km, than those in the USA, which peak beyond 100 km.

Discussion

Our findings reveal a previously unquantified physical interaction that constrains the global solar energy transition: aerosol emissions from coal-fired power plants directly reduce the energy output of PV installations to a degree that measurably alters regional energy yields. While many recent studies document the accelerating expansion of PV

capacity as a primary metric of progress^{10,49}, our work shows that this physical dynamic represents a fundamental challenge to realizing the associated climate and air quality benefits that are widely projected under idealized scenarios^{14,15}. Using a global, facility-level dataset, it becomes clear that this effect is real: every year, aerosol-related losses are equivalent to almost one-third of the energy generated by newly installed PV capacity worldwide. In China, where coal and solar have expanded most rapidly in parallel, this ratio can exceed 50%. These estimates capture only the direct radiative effects of aerosols, while additional indirect impacts through aerosol–cloud interactions, which generally lead to negative radiative forcing by increasing cloud reflectivity and persistence⁵⁰, would probably further increase solar energy losses, meaning that estimates presented in this study represent a conservative lower bound. This demonstrates that the effectiveness of the energy transition cannot be evaluated by installed capacity alone, but must also account for the spatially embedded, performance-degrading interactions with the remaining fossil fuel fleet.

The persistence of coal power thus constitutes more than just a market-based challenge for renewables⁵¹; it creates a direct physical barrier that degrades solar asset performance. The global 5.8% aerosol-induced reduction in PV generation demonstrates that air pollution is actively eroding realized climate benefits and reducing the value of new solar investments. To further quantify how this barrier

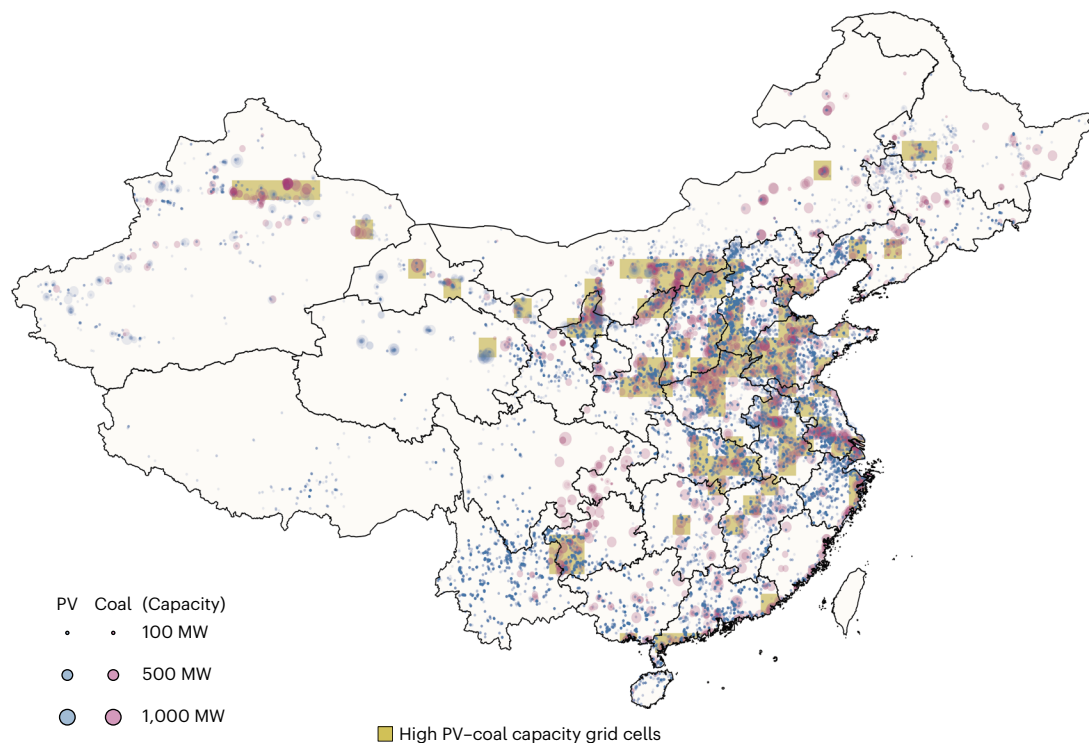


Fig. 5 | Widespread co-location of coal power and solar PV capacity in China. The map shows the spatial distribution of coal-fired power plants (purple markers) and PV facilities (blue markers) across China. PV capacity is estimated using a constant

solar capacity factor of 18%. The yellow grid cells highlight the top 30% of locations where high-capacity coal and PV installations are co-located, indicating regions with high potential for interaction between the two energy sources.

influences the efficiency of the solar transition, the R_{LG} is introduced as a comparative metric that evaluates the resilience of regional solar expansion. R_{LG} highlights a considerable, previously hidden cost of the co-existence of coal and solar infrastructure, a cost most evident in China, which suffers from the highest ratio among major economies (average of 38.4% and peaks exceeding 60% in recent years). This high loss is primarily driven by coal emissions, with sulfate aerosols from coal-derived SO_2 accounting for 46.2% of all aerosol-induced PV losses in the country.

China's situation, however, is particularly instructive because it is also the only major region where these performance losses are in sustained decline. Our fixed-network analysis shows that from 2013 to 2023, aerosol-induced losses in China fell by 1.4% annually, a trend that contrasts sharply with the rising losses in the USA (1.5% yr^{-1}) and Europe (1.3% yr^{-1}). This improvement occurred not because of a coal phase-out, as coal capacity continued to expand in China⁴⁶, but is instead a direct cobenefit of stricter pollutant emission standards^{44,45}. The evidence for this is multifaceted: the decline in losses corresponds directly with falling pollutant levels over PV sites, including effective AOD (−1.7% yr^{-1}), tropospheric NO_2 (−2.4% yr^{-1}) and PBL SO_2 (−1.4% yr^{-1}); and a significant positive spatial correlation ($I = 0.5654$) connects coal capacity to these losses. This provides direct evidence that while cleaning up existing coal plants yields partial benefits, a full phase-out would more decisively remove this performance barrier. The importance of this spatial coal–solar interaction is further illustrated by the contrasting case of the USA. The USA exhibits the lowest R_{LG} because there is limited co-location of its solar and coal fleets. This geographical separation results in negligible spatial correlation between coal capacity and aerosol-induced PV losses. The US case demonstrates that the aerosol-induced performance barrier is not an inherent property of solar power, but is a direct consequence of the proximate operation of coal-fired power plants, a barrier that would be removed by their phase-out.

Looking forward, the physical interaction between coal-based aerosols and solar PV performance is likely to become an increasingly critical constraint on the global energy transition. A fundamental challenge for grids with high renewable penetration is the need for dispatchable power to maintain stability during periods of low solar output⁵². This stability challenge is being amplified by accelerating electricity demand, which the International Energy Agency projects will grow by nearly 4% annually through 2027, driven by industrial expansion, electrification and the rapid growth of artificial intelligence applications⁵³. Together, these pressures for both grid reliability and rising demand are promoting many economies to extend their reliance on coal. In China, this dynamic is evident in the continued approval of new coal plants as backup capacity for renewables⁵⁴, a trend also exported abroad through Chinese and Indian investment in coal projects in Africa⁵⁵. In the near term, the persistence of large coal capacity remains a challenge for the solar transition: while existing coal plants continue to provide backup power, their operation sustains fossil dependence and constrains the benefits of solar deployment. Retrofitting coal plants to ULE standards can reduce SO_2 , NO_x and $PM_{2.5}$ emissions per facility; however, given China's vast coal capacity, total emissions remain among the highest globally, and these retrofits do little to mitigate CO_2 emissions⁵⁶. Ultimately, faster coal retirement is essential for achieving the objectives of the energy transition, while coordinated policy measures should address short-term system reliability without reinforcing long-term fossil dependence. Our results highlight the direct physical cost of this dependence: the same coal plant retained or newly built will continue to degrade the performance of the solar assets intended to replace them. A complete assessment of future energy pathways must therefore account for these cross-sectoral dynamics, integrating climate, air quality and energy system models to ensure that the expected benefits of renewable investments are fully realized.

Methods

Three-step framework for PV generation and losses

A three-step framework was developed to estimate facility-level PV energy generation and its associated major losses from clouds and aerosols globally. This framework introduces two main advances over previous methods: it combines multiple sources to assemble a more complete global PV dataset and it improves accuracy by decoupling the initial identification of PV sites from the precise extraction of panel footprints.

A key challenge is that global-scale detection is inherently difficult because PV installations must be distinguished from diverse non-PV surfaces, often leading to boundary errors. For this reason, separating the initial detection from the final segmentation improves accuracy over single-stage models that perform both tasks simultaneously. Our process therefore begins with Step 1: identifying candidate PV locations using a combination of existing inventories^{10,32,33}, crowd-sourced records and a convolutional neural network (CNN) classifier trained on global-scale Sentinel-2 imagery.

This is followed by Step 2, where, in contrast to broad detection, segmentation applied to satellite imagery with identified PV sites is more tractable, as panels show distinct visual and spectral contrast with their surroundings. Therefore, high-resolution imagery from confirmed sites was processed using the SAM³⁵, developed by Meta, to extract precise panel boundaries using a ‘few-shot’ learning approach. To evaluate the spatial accuracy of the resulting dataset, comparisons were then made with two existing global PV inventories, as described in Supplementary Table 1.

In Step 3, these high-accuracy panel footprints were integrated with atmospheric data from the Modern-Era Retrospective Analysis for Research and Applications (MERRA-2) re-analysis in a validated PV model. This allowed the estimation of a time series of facility-level electricity generation, and the distinct losses caused by clouds and aerosols for each of the 140,945 facilities in our database.

Identification of PV facilities from multiple data sources

We first identified potential PV facilities using OpenStreetMap (OSM), a global crowd-sourced geospatial digital database. While OSM data are extensively used for mapping and spatial analysis, its data quality and completeness vary by region. In particular, PV facilities may be mapped at the construction area level, often overestimating panel coverage and including planned rather than operational sites. Candidate PV facilities were extracted by querying OSM with the attributes: power=plant, plant:source=solar and plant:method=photovoltaic. Global queries were regionally partitioned and run in parallel on UK's JASMIN supercomputer, which provided the necessary computational power for large-scale geospatial processing.

To supplement OSM, we integrated publicly available PV datasets generated by machine learning-based analyses of satellite imagery. One such resource is the global inventory by Kruitwagen et al.¹⁰, which used a U-Net model trained on Sentinel-2 imagery to detect PV installations worldwide as of 2019. Their dataset consists of polygons outlining each identified PV site. From these polygons, we extracted the central coordinates and bounding box of each potential site to guide the next stage of our analysis.

For China, Feng et al.³³ employed a random forest classifier on Google Earth Engine using 2020 Sentinel-2 imagery, training province-specific models to improve detection accuracy. This dataset captured older installations missing from global inventories but introduced more false positives. After clustering the 10-m resolution GeoTIFF PV masks into individual facilities, we extracted each PV facility's central coordinates and bounding information. For India, we incorporated the PV dataset from Ortiz et al.³², which combined a U-Net detection model with hard negative mining to improve detection accuracy. This database includes 1,363 validated and grouped PV installations updated to 2020.

To identify new or previously unrecorded PV installations, we trained a U-Net model with a VGG16⁵⁷ backbone on Sentinel-2 imagery. A total of 12,000 PV installations were manually annotated to build a globally representative training dataset. To improve detection across diverse environments, the model was trained on PV sites situated in deserts, croplands, mountains, floating systems and built environments. Further details on model architecture, training data preparation and deployment are provided in Supplementary Note 1.

By integrating the OSM records, published PV inventories and detections from our custom model, we ultimately identified 326,423 potential PV polygons globally.

PV panel segmentation with SAM

Estimating PV generation from satellite imagery requires accurately extracting the panel area that receives solar irradiance. This, in turn, depends on the precise extent of the PV panels themselves, excluding non-panel features and empty spaces often included within the facility polygons derived above. Including these non-PV areas leads to overestimates of panel surface area. Applying uniform scaling factors to correct such errors is unreliable because these ratios are neither constant nor reliably correlated with location or facility size.

The SAM³⁵ is a foundation model trained on 11 million images and 1.1 billion masks. Designed to generalize effectively across diverse image datasets, it can be adapted to new imagery and unfamiliar objects with little or no additional training. It generalizes to new imagery using prompting techniques, where user-provided inputs (such as text descriptions or spatial information) guide the model's segmentation. SAM consists of an image encoder that computes embeddings, a prompt encoder that processes user inputs and a lightweight mask decoder that generates segmentation masks. SAM has been applied across diverse domains, including three-dimensional object segmentation, medical imaging and crop disease identification⁵⁸. Here, SAM was used to segment PV arrays from pre-identified locations on satellite images, using spatial prompts in the form of bounding boxes and individual points to initiate the segmentation process (Supplementary Fig. 3).

Red, green and blue (RGB) composites were generated from Sentinel-2's 12-band imagery. At this local scale, PV panels and their immediate surroundings are readily distinguishable in an RGB composite. Areas of persistent cloud cover were filled by incorporating PlanetScope RGB imagery in those areas. With daily global coverage at a spatial resolution of 3 m, PlanetScope data offered more opportunities for cloud-free views and improved segmentation accuracy, especially for smaller PV arrays. High-resolution Google imagery was used where it was available and up to date.

Spatial prompts included bounding boxes and point annotations placed on PV panels. Segmented PV facilities were manually reviewed to ensure accurate boundaries. False detections and non-PV features were corrected using additional prompts if necessary. This two-stage approach ultimately produced segmented PV polygons for all sites previously identified. To consolidate closely located polygons into coherent PV facilities, those within 50 m of one another were merged. This final aggregation resulted in a comprehensive global database of 140,945 distinct PV facilities.

PV installation time from time series classification

To determine the installation time of each PV facility, we applied a time series classification approach using the sktime machine learning framework⁵⁹. Monthly composites of Sentinel-2 top of atmosphere reflectance from 2017 to 2024 were generated, and for each PV polygon, the mean reflectance across all pixels and spectral bands was extracted. This resulted in a multivariate time series representing the temporal reflectance signature of each facility. A manually labelled training dataset was assembled using PV sites with known installation times, identified from high-resolution satellite imagery. A supervised classifier

was trained to detect installation events on the basis of changes in reflectance patterns over time. The model was optimized using a large and diverse training set to distinguish true PV installation signals from common pre-installation changes, including grading, land clearing and temporary construction activities. The trained model was applied to all detected PV polygons to estimate installation time. As full data were not available until 2017, all PV facilities installed before 2017 were assigned an installation time of 2017. Example reflectance trajectories used to support classification are shown in Supplementary Fig. 4, with corresponding visual examples of detected facility expansion shown in Supplementary Fig. 5.

Estimation of PV power-generation potential

The theoretical electricity generation of each identified PV facility was estimated using the following equation:

$$E = I \times A \times k \times C, \quad (1)$$

where E is the potential alternating current (a.c.) power output (kW), I is the available solar irradiance at the PV surface (kW m^{-2}), A is the effective PV panel area (m^2) derived from our database, k is the coefficient accounting for system-level energy losses and C is the module conversion efficiency from solar radiation to electricity.

The parameter I depends on both atmospheric conditions and the facility's geographic orientation and tilt. Atmospheric conditions, including aerosol loading and cloud cover, can alter the balance of direct and diffuse solar radiation that reach the Earth's surface, influencing the global horizontal irradiance (GHI). The orientation and tilt angle of PV panels are optimized to maximize solar exposure over seasonal and diurnal cycles.

GHI was estimated using MERRA-2⁶⁰, the latest atmospheric reanalysis produced by NASA's Global Modeling and Assimilation Office. MERRA-2 provides an hourly, globally gridded dataset ($0.5^\circ \times 0.625^\circ$) of atmospheric and surface variables, including radiation, temperature, relative humidity and wind speeds. MERRA-2 assimilates aerosol measurements from spaceborne observations, representing their interactions with other physical processes. For this analysis, the surface incoming shortwave flux (SWGDN) from MERRA-2 M2T1NXRAD data products was used. SWGDN represents GHI, including both direct and diffuse components.

For a PV facility with fixed orientation and tilt angle, the optimal tilt was estimated from its latitude⁶¹. To translate GHI on a horizontal surface into the solar irradiance available on a tilted panel, GHI was decomposed into direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI)⁶²

$$\text{GHI} = \text{DNI} \times \cos \theta + \text{DHI}, \quad (2)$$

where θ is the solar zenith angle. The diffuse fraction (DHI/GHI) is determined using the ratio of global to extraterrestrial irradiance on a horizontal plane, as implemented in the open-source pvlib library⁶³. Using GHI components and PV geometry, irradiance on optimally tilted panels was derived using pvlib.

The effective panel area A of a PV depends on both the overall facility footprint and the spacing between individual panel arrays. This relationship is commonly expressed using a ground coverage ratio, defined as the ratio of panel-covered area to total facility land area: $A = A_L \times r$, where A_L is the total land area and r is the ground coverage ratio. Applying this calculation directly to PV polygons derived from traditional machine learning methods can be problematic, as satellite-based detections often overestimate panel areas by including non-PV features. By contrast, this two-stage search and extraction approach substantially improves detection accuracy by excluding non-PV elements during the extraction phase. This improved detection enables the use of a uniform ground coverage ratio r globally to estimate the active panel area.

The coefficient k , which accounts for system-level losses, follows a model validated by Saxena et al.⁶⁴ on a large sample of operational PV panels. This factor includes losses from suboptimal PV orientation (0.5%), mismatch in power points (0.3%), d.c. cabling (2%), a.c. cabling (0.5%), d.c.-to-a.c. conversion (2.2%), module downtime (0.5%), soiling by dust and snow (3.5%), and transformer inefficiencies (0.9%). Collectively, these factors amount to a total loss of 10.09%.

The conversion efficiency C represents the fraction of incident solar energy converted into electricity by the solar cells. For conventional silicon-based technologies, this efficiency typically ranges between 18% and 22%. A fixed value of 20% was used in this study.

To isolate the direct climate impacts that reduce PV power generation, MERRA-2's SWGDN estimated under clear-sky and clean-sky conditions is used. SWGDNCLR represents shortwave flux under clear-sky (cloud-free) conditions. Under such conditions, one can estimate facility-level PV power output solely for clear-sky scenarios:

$$E_{\text{clear}} = \text{SWGDNCLR} \times A \times k \times C. \quad (3)$$

Although MERRA-2 does not directly provide incoming shortwave flux under clean-sky (aerosol-free) conditions, it can be approximated by scaling the total SWGDN with the ratio of clean-sky net shortwave flux (SWGNTCLN) to full-sky net shortwave flux (SWGNT). This approach removes aerosol contributions and yields E_{clean} , the estimated PV output in the absence of aerosols

$$E_{\text{clean}} = \frac{\text{SWGNTCLN}}{\text{SWGNT}} \times \text{SWGDN} \times A \times k \times C. \quad (4)$$

By comparing the baseline output E (from all-sky conditions) with E_{clear} (no clouds) and E_{clean} (no aerosols), cloud and aerosol impacts can be separately quantified. Finally, the total climate-induced loss L can be computed using the following equation:

$$L = \left(\frac{E_{\text{clean}} - E}{E} + \frac{E_{\text{clear}} - E}{E} \right) \times 100\%. \quad (5)$$

To illustrate the structure of the resulting dataset, Supplementary Fig. 9 shows histograms of estimated annual generation per facility in 2023 for the globe and major regions (China, the USA, Europe, including the UK, India and the rest of the world).

Calculation of PV-effective AOD and gas concentrations

To evaluate the impact of atmospheric pollution on PV generation, aerosol conditions were quantified by calculating effective AOD for each PV facility. Major anthropogenic aerosol emissions originate from fossil-fuel combustion, particularly from coal-fired and other thermal power plants, industrial processes, vehicular transport and residential heating, as well as from biomass burning. Natural sources include desert dust, sea salt and volcanic activity⁶⁵. AOD was obtained from the MERRA-2 reanalysis (M2T1NXAER collection), which provides hourly AOD at 550 nm on a $0.625^\circ \times 0.5^\circ$ grid. The accuracy of the MERRA-2 AOD product has been extensively validated in previous global and regional studies⁶⁶. These evaluations demonstrate close agreement with AERONET and MODIS observations, with typical correlations of 0.7–0.9 and root mean square error values below 0.2, confirming that MERRA-2 AOD provides reliable and stable estimates suitable for global and regional analyses. To further assess consistency in our study, we performed an effective-AOD trend analysis for all PV facilities in China using the 1-km Multi-Angle Implementation of Atmospheric Correction (MAIAC) AOD product⁶⁷ (Extended Data Fig. 2). The MAIAC-based results show a comparable declining trend to that derived from MERRA-2, with a reduction rate of -0.0085 yr^{-1} ($-2.9\% \text{ yr}^{-1}$) during 2013–2023. This difference can be partly attributed to MAIAC's finer spatial resolution, which better captures local aerosol gradients around

coal plants and aligns with the PV–coal distance distribution shown in Supplementary Fig. 8, where most PV facilities are located 20–30 km from the nearest coal power plant, below the MERRA-2 grid spacing. These findings indicate that the AOD decline and its implications for PV performance are robust across independent datasets.

To represent trace-gas pollutants, we further derived tropospheric NO₂ and PBL SO₂ column densities from the spaceborne Ozone Monitoring Instrument (OMI). For NO₂, we used the OMI_MINDS_NO2d product and extracted the ‘ColumnAmountNO2TropCloudScreened’ variable, which represents daily tropospheric vertical column densities (molecules per square centimetre) filtered for high-quality observations with cloud fractions below 0.3 and solar zenith angles under 85°. For SO₂, we used the OMSO2e product, which provides daily global estimates of SO₂ column density in the PBL at 0.25° × 0.25° resolution.

We computed effective pollutant values to reflect their contribution to generation reductions at each PV facility. Two weighting factors were applied: (1) clear-sky irradiance weighting, which assigns greater weight to pollutants present during high-irradiance conditions (for example, daytime and summer), and (2) PV area weighting, which accounts for the relative size of each facility. The effective pollutant value was computed as

$$P_{\text{eff}} = \frac{\sum_i (P_i \times I_i \times A_i)}{\sum_i (I_i \times A_i)}, \quad (6)$$

where P_i is the pollutant value (AOD, tropospheric NO₂ or PBL SO₂) in pixel i , I_i is the clear-sky irradiance and A_i is the PV panel area.

Aerosol source attribution using GEOS-Chem

Aerosol source attribution was carried out using GEOS-Chem version 14.6.3⁶⁸ in the aerosol-only configuration, which simulates the transport and removal of aerosol species from anthropogenic and natural sources (sulfate, nitrate, ammonium, black carbon, organic carbon, dust and sea salt) using archived oxidant fields. Anthropogenic emissions used in the GEOS-Chem simulations are taken from the Community Emissions Data System (CEDS) inventory (version 2025-04), which provides an internally consistent global emissions dataset widely used in chemistry-transport modelling. China-specific inventories such as the Multi-Resolution Emission Inventory for China (MEIC) are also available and may offer more detailed representations of recent emission changes associated with national clean-air policies⁶⁹. Simulations were conducted at 0.5° × 0.625° horizontal resolution with 47 vertical layers extending to 0.01 hPa and were driven by MERRA-2 meteorological fields provided at the same native resolution (0.5° × 0.625°, 72 levels). The model was run for the year 2023, following a 6-month spin-up for the global (2° × 2.5°) simulation and an additional 6-month spin-up for the nested China simulation (0.5° × 0.625°). Lateral boundary conditions for the nested domain were updated every 3 h from the corresponding 2° × 2.5° global run.

Two categories of simulations were performed (Supplementary Fig. 10). The base simulation included all anthropogenic and natural emissions and represents total atmospheric aerosol loading. To isolate the contribution of coal-related and other energy-sector emissions, a set of sector-specific sensitivity simulations was conducted in which emissions from individual power-sector sources were activated within China while all other emissions remained switched off. These simulations used identical emission inventories and model configurations for the global and nested domains to ensure consistency of sources across spatial scales (Supplementary Table 6).

To relate the GEOS-Chem attribution to PV energy losses, monthly baseline AOD and energy-sector AOD were computed for all MERRA-2 grid cells in China for 2023. The monthly fraction of total AOD attributable to energy-sector emissions was then aggregated across all PV facilities using a weighted mean, where the weight for each facility equals its aerosol-induced energy reduction for that month. This weighting

accounts for site-specific capacity, monthly irradiance and effective AOD, providing a direct link between sector-attributed aerosol loading and PV energy losses.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The Sentinel-2 data used in this study are freely available from the European Space Agency through the Copernicus programme and can be accessed via the Copernicus Data Space Ecosystem (<https://dataspace.copernicus.eu>). The MERRA-2 reanalysis data (including M2T1NXRAD for radiation and M2T1NXAER for aerosols) are publicly available from the NASA Goddard Earth Sciences Data and Information Services Center (<https://disc.gsfc.nasa.gov/>). The OMI data products for NO₂ (OMI_MINDS_NO2d) and SO₂ (OMSO2e) are also publicly available from the NASA Goddard Earth Sciences Data and Information Services Center. The Global Coal Plant Tracker database is maintained by Global Energy Monitor and is publicly available for download (<https://globalenergymonitor.org/projects/global-coal-plant-tracker/>). The primary dataset of global, facility-level PV generation and losses developed and analysed in this study is available via Zenodo at <https://doi.org/10.5281/zenodo.18794230> (ref. 70).

Code availability

The code used in this study is publicly available via GitHub at <https://github.com/ray-climate/global-pv-generation-loss-dataset>. The version corresponding to this publication (version 1.0) is available via Zenodo at <https://doi.org/10.5281/zenodo.18844891> (ref. 71).

References

1. Adoption of the Paris Agreement (UNFCCC, 2015); <https://unfccc.int/resource/docs/2015/cop21/eng/L09r01.pdf>
2. Net Zero by 2050: A Roadmap for the Global Energy Sector (IRENA, 2021); <https://www.iea.org/reports/net-zero-by-2050>
3. Renewable Energy Highlights (IRENA, 2024).
4. Winemiller, K. O. et al. Balancing hydropower and biodiversity in the Amazon, Congo, and Mekong. *Science* **351**, 128–129 (2016).
5. Sharmin, T., Khan, N. R., Akram, M. S. & Ehsan, M. M. A state-of-the-art review on geothermal energy extraction, utilization, and improvement strategies: conventional, hybridized, and enhanced geothermal systems. *Int. J. Thermofluids* **18**, 100323 (2023).
6. Shetty, C. & Priyam, A. A review on tidal energy technologies. *Mater. Today Proc.* **56**, 2774–2779 (2022).
7. Wu, G. C. et al. Avoiding ecosystem and social impacts of hydropower, wind, and solar in Southern Africa's low-carbon electricity system. *Nat. Commun.* **15**, 1083 (2024).
8. Machin, A. & Márquez, F. Advancements in photovoltaic cell materials: silicon, organic, and perovskite solar cells. *Materials* **17**, 1165 (2024).
9. Hijjawi, U., Lakshminarayana, S., Xu, T., Fierro, G. P. M. & Rahman, M. A review of automated solar photovoltaic defect detection systems: approaches, challenges, and future orientations. *Sol. Energy* **266**, 112186 (2023).
10. Kruitwagen, L. et al. A global inventory of photovoltaic solar energy generating units. *Nature* **598**, 604–610 (2021).
11. Renewable Power Generation Costs (IRENA, 2023).
12. Kishore, T. S., Kumar, P. U. & Ippili, V. Review of global sustainable solar energy policies: significance and impact. *Innov. Green Dev.* **4**, 100224 (2025).
13. Renewables 2023 (IEA, 2023).
14. Zhang, Z. et al. Worldwide rooftop photovoltaic electricity generation may mitigate global warming. *Nat. Clim. Change* **15**, 393–402 (2025).

15. Qin, Y. et al. Amplified positive effects on air quality, health, and renewable energy under China's carbon neutral target. *Nat. Geosci.* **17**, 411–418 (2024).
16. Lowe, R. J. & Drummond, P. Solar, wind and logistic substitution in global energy supply to 2050 —barriers and implications. *Renew. Sustain. Energy Rev.* **153**, 111720 (2022).
17. Rather, K. N. & Mahalik, M. K. Investigating the assumption of perfect displacement for global energy transition: panel evidence from 73 economies. *Clean Technol. Environ. Policy* **26**, 2739–2752 (2024).
18. Pehl, M. et al. Understanding future emissions from low-carbon power systems by integration of life-cycle assessment and integrated energy modelling. *Nat. Energy* **2**, 939–945 (2017).
19. Chen, S. et al. Deploying solar photovoltaic energy first in carbon-intensive regions brings gigatons more carbon mitigations to 2060. *Commun. Earth Environ.* **4**, 369 (2023).
20. Siler-Evans, K., Azevedo, I. L., Morgan, M. G. & Apt, J. Regional variations in the health, environmental, and climate benefits of wind and solar generation. *Proc. Natl Acad. Sci. USA* **110**, 11768–11773 (2013).
21. Millstein, D., Wiser, R., Bolinger, M. & Barbose, G. The climate and air-quality benefits of wind and solar power in the United States. *Nat. Energy* **2**, 17134 (2017).
22. Vandyck, T. et al. Air quality co-benefits for human health and agriculture counterbalance costs to meet Paris Agreement pledges. *Nat. Commun.* **9**, 4939 (2018).
23. York, R. & Bell, S. E. Energy transitions or additions? Why a transition from fossil fuels requires more than the growth of renewable energy. *Energy Res. Soc. Sci.* **51**, 40–43 (2019).
24. Mahdavi, P., Ross, M. L. & Simoni, E. Fossil fuel subsidy reforms have become more fragile. *Nat. Clim. Change* **15**, 569–574 (2025).
25. Xu, Y. China's fossil-fuel challenge—how to build a bridge to renewables. *Nature* **643**, 907–910 (2025).
26. Li, X., Wagner, F., Peng, W., Yang, J. & Mauzerall, D. L. Reduction of solar photovoltaic resources due to air pollution in China. *Proc. Natl Acad. Sci. USA* **114**, 11867–11872 (2017).
27. Sweerts, B. et al. Estimation of losses in solar energy production from air pollution in China since 1960 using surface radiation data. *Nat. Energy* **4**, 657–663 (2019).
28. Li, X., Mauzerall, D. L. & Bergin, M. H. Global reduction of solar power generation efficiency due to aerosols and panel soiling. *Nat. Sustain.* **3**, 720–727 (2020).
29. Yang, J., Yi, B., Wang, S., Liu, Y. & Li, Y. Diverse cloud and aerosol impacts on solar photovoltaic potential in southern China and northern India. *Sci. Rep.* **12**, 19671 (2022).
30. Wu, W. et al. Assessment of the ecological and environmental effects of large-scale photovoltaic development in desert areas. *Sci. Rep.* **14**, 22456 (2024).
31. Klingler, M., Ameli, N., Rickman, J. & Schmidt, J. Large-scale green grabbing for wind and solar photovoltaic development in Brazil. *Nat. Sustain.* **7**, 747–757 (2024).
32. Ortiz, A. et al. An artificial intelligence dataset for solar energy locations in India. *Sci. Data* **9**, 497 (2022).
33. Feng, Q. et al. A 10-m national-scale map of ground-mounted photovoltaic power stations in China of 2020. *Sci. Data* **11**, 198 (2024).
34. Chen, Y., Zhou, J., Ge, Y. & Dong, J. Uncovering the rapid expansion of photovoltaic power plants in China from 2010 to 2022 using satellite data and deep learning. *Remote Sens. Environ.* **305**, 114100 (2024).
35. Kirillov, A. et al. Segment anything. Preprint at <https://arxiv.org/abs/2304.02643> (2023).
36. An, Z. et al. Severe haze in northern China: a synergy of anthropogenic emissions and atmospheric processes. *Proc. Natl Acad. Sci. USA* **116**, 8657–8666 (2019).
37. Wang, Z. & Fan, W. Economic and environmental impacts of photovoltaic power with the declining subsidy rate in China. *Environ. Impact Assess. Rev.* **87**, 106535 (2021).
38. Hoang, A. T. et al. Impacts of COVID-19 pandemic on the global energy system and the shift progress to renewable energy: opportunities, challenges, and policy implications. *Energy Policy* **154**, 112322 (2021).
39. Corwin, K. A. et al. Solar energy resource availability under extreme and historical wildfire smoke conditions. *Nat. Commun.* **16**, 245 (2025).
40. Zhang, Q. & Chen, Y. Environmental regulation, coal de-capacity, and PM2.5 in China. *Sci. Rep.* **15**, 7785 (2025).
41. Richter, A., Burrows, J. P., Nüß, H., Granier, C. & Niemeier, U. Increase in tropospheric nitrogen dioxide over China observed from space. *Nature* **437**, 129–132 (2005).
42. Ma, Q. et al. Impacts of coal burning on ambient PM2.5 pollution in China. *Atmos. Chem. Phys.* **17**, 4477–4491 (2017).
43. *Coal in China* (IEA, 2025); <https://www.iea.org/countries/china/coal>
44. Tang, L. et al. Substantial emission reductions from Chinese power plants after the introduction of ultra-low emissions standards. *Nat. Energy* **4**, 929–938 (2019).
45. Cui, R. Y. et al. A plant-by-plant strategy for high-ambition coal power phaseout in China. *Nat. Commun.* **12**, 1468 (2021).
46. China's construction of new coal power plants reached 10-year high in 2024. *Carbon Brief* <https://www.carbonbrief.org/chinas-construction-of-new-coal-power-plants-reached-10-year-high-in-2024/> (accessed 31 May 2025).
47. Global Coal Plant Tracker. *Global Energy Monitor* (2025).
48. Wartenberg, D. Multivariate spatial correlation: a method for exploratory geographical analysis. *Geogr. Anal.* **17**, 263–283 (1985).
49. Li, A. et al. Global photovoltaic solar panel dataset from 2019 to 2022. *Sci. Data* **12**, 637 (2025).
50. Virtanen, A. et al. High sensitivity of cloud formation to aerosol changes. *Nat. Geosci.* **18**, 289–295 (2025).
51. Wagner, G. et al. Energy policy: push renewables to spur carbon pricing. *Nature* **525**, 27–29 (2015).
52. Yin, J., Molini, A. & Porporato, A. Impacts of solar intermittency on future photovoltaic reliability. *Nat. Commun.* **11**, 4781 (2020).
53. *Energy and AI* (IEA, 2025).
54. Howe, C. China to keep building coal plants through 2027, state planner says. *Reuters* <https://www.reuters.com/sustainability/climate-energy/china-keep-building-coal-plants-through-2027-state-planner-says-2025-04-14/> (14 April 2025).
55. China and India are holding up global shift from coal power: GEM. *Bloomberg* <https://www.bloomberg.com/news/articles/2025-04-03/china-and-india-are-holding-up-global-shift-from-coal-power-gem> (3 April 2025).
56. Jin, Y., Peng, W. & Urpelainen, J. An ultra-low emission coal power fleet for cleaner but not hotter air. *Environ. Res. Lett.* **15**, 091002 (2020).
57. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. Preprint at <https://arxiv.org/abs/1409.1556> (2014).
58. Ma, J. et al. Segment anything in medical images. *Nat. Commun.* **15**, 654 (2024).
59. Löning, M. et al. sktime: a unified interface for machine learning with time series. Preprint at <https://arxiv.org/abs/1909.07872> (2019).
60. Gelaro, R. et al. The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Clim.* **30**, 5419–5454 (2017).
61. Landau, C. R. Optimum tilt of solar panels. *Solarpaneltilt.com* <https://www.solarpaneltilt.com/> (2017).

62. A. & Duffie, J. A. Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation. *Sol. Energy* **28**, 293–302 (1982).
63. Anderson, K. S. et al. pvlb python: 2023 project update. *J. Open Source Softw.* **8**, 5994 (2023).
64. Saxena, A., Brown, C., Arneth, A. & Rounsevell, M. Modelling the global photovoltaic potential on land and its sensitivity to climate change. *Environ. Res. Lett.* **18**, 104017 (2023).
65. IPCC. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2021).
66. Gueymard, C. A. & Yang, D. Worldwide validation of CAMS and MERRA-2 reanalysis aerosol optical depth products using 15 years of AERONET observations. *Atmos. Environ.* **225**, 117216 (2020).
67. Lyapustin, A., & Wang, Y. MODIS/Terra+Aqua Land Aerosol Optical Depth Daily L2G Global 1km SIN Grid V061 (dataset). NASA land processes distributed active archive center (2022); <https://doi.org/10.5067/MODIS/MCD19A2.061>
68. The international GEOS-Chem user community. geoschem/GCClassic: GCClassic 14.6.3 (14.6.3). Zenodo <https://doi.org/10.5281/zenodo.16541817> (2025).
69. Zheng, B. et al. Mapping anthropogenic emissions in China at 1km spatial resolution and its application in air quality modeling. *Sci. Bull.* **66**, 612–620 (2021).
70. Song, R. & Feng, Y. Global facility-level solar photovoltaic inventory with energy generation and loss estimates. Zenodo <https://doi.org/10.5281/zenodo.18794230> (2026).
71. Song, R. et al. Coal plants persist as a large barrier to the global solar energy transition—dataset and examples. Zenodo <https://doi.org/10.5281/zenodo.18844891> (2026).

Acknowledgements

This work was undertaken on JASMIN, the UK's collaborative data analysis environment (<https://www.jasmin.ac.uk>). Planet Labs provided the high-resolution PlanetScope imagery through their Education and Research Program. R.S., A.C.P. and R.G.G. were partly supported by the UK Natural Environment Research Council (NERC) through the National Centre for Earth Observation (grant no. NE/R016518/1). F.Y. was supported by the NERC and BBSRC AgZero+ research programme (grant no. NE/W005050/1) and the NERC CPEO programme (grant no. NE/X006328/1).

Author contributions

R.S. conceived the study, developed the primary methodology, performed the main data analysis and visualization, and wrote the

original paper. R.S. and F.Y. jointly developed the global PV database. B.S. performed the GEOS-Chem simulations and contributed to the data analysis. A.C.P., J.-P.M. and R.G.G. contributed to the development of the analytical framework and conducted additional analysis. C.H. contributed to the conceptualization of the study and provided input on the policy implications of the findings. All authors contributed to reviewing and editing the final paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41893-026-01836-5>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41893-026-01836-5>.

Correspondence and requests for materials should be addressed to Rui Song, Jan-Peter Muller or Roy G. Grainger.

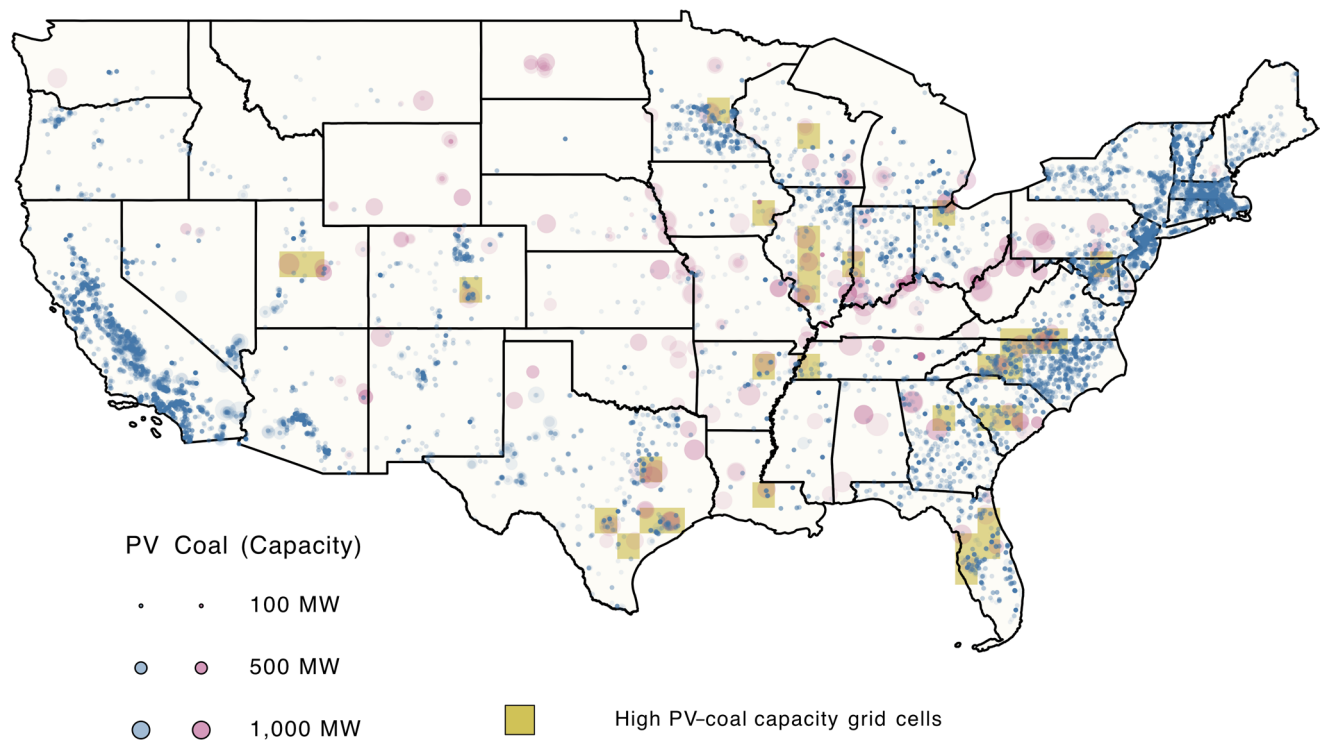
Peer review information *Nature Sustainability* thanks Quentin Paletta and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

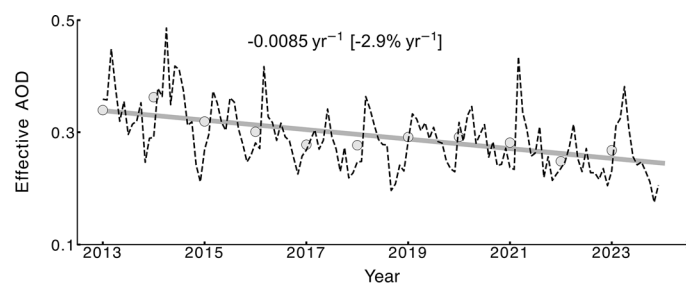
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026



Extended Data Fig. 1 | Limited co-location of coal and solar capacity in the United States. Spatial distribution of PV installations (blue markers) and coal-fired power plants (purple markers) across $1.0^\circ \times 1.0^\circ$ grid cells. In contrast to China, the small number of yellow grid cells illustrates the limited co-location of these energy sources in the United States.



Extended Data Fig. 2 | Effective-AOD trends derived from Multi-Angle Implementation of Atmospheric Correction (MAIAC) data for PV facilities in China. Temporal trends of effective aerosol optical depth (AOD) from 2013 to 2023 derived using the 1-km MAIAC product for all mapped PV facilities in China.

The MAIAC results show a declining trend in effective AOD, with a reduction rate of -0.0085 yr^{-1} ($-2.9\% \text{ yr}^{-1}$), consistent with the pattern obtained from the MERRA-2 dataset reported in the main text.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	Data collection was performed using a combination of open-source and custom code. Sentinel-2 imagery was processed using both the Copernicus Data Space Ecosystem and Google Earth Engine. PV facility detection used a custom U-Net machine learning model implemented in Python. Panel segmentation was conducted using the Segment Anything Model (SAM) version 1.0 from Meta AI. The primary dataset of global, facility-level PV generation and losses developed and analysed in this study, is publicly archived on Zenodo at https://doi.org/10.5281/zenodo.18794230 .
Data analysis	Data analysis was conducted using custom Python code developed by the authors and executed on JASMIN, the UK's collaborative data analysis environment (https://www.jasmin.ac.uk). Machine learning models were used for PV facility detection and installation date identification. A U-Net model with a VGG16 backbone was implemented using TensorFlow (v2.15.0) and trained on JASMIN's Orchid GPU cluster. Time-series classification was performed using sktime and scikit-learn (v1.4.1). Additional analysis included photovoltaic energy modeling using pvlib (v0.9.5), atmospheric data processing with xarray (v2023.1.0) and netCDF4 (v1.7.0). Data handling and visualization used pandas, NumPy, matplotlib, and GeoPandas. The code used in this study is publicly available on GitHub (https://github.com/ray-climate/global-pv-generation-loss-dataset). The version corresponding to this publication (v1.0) is archived on Zenodo at https://doi.org/10.5281/zenodo.18844891 .

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All datasets used in this study are publicly available. Sentinel-2 satellite imagery is freely available from the European Space Agency through the Copernicus programme (<https://dataspace.copernicus.eu>). MERRA-2 atmospheric reanalysis data (including M2T1NXRAD and M2T1NXAER) are available from the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) (<https://disc.gsfc.nasa.gov>). Ozone Monitoring Instrument (OMI) data products for nitrogen dioxide (OMI_MINDS_NO2d) and sulphur dioxide (OMSO2e) are also available from NASA GES DISC. Coal plant data were obtained from the Global Coal Plant Tracker by Global Energy Monitor (<https://globalenergymonitor.org/projects/global-coal-plant-tracker>). The primary dataset of global, facility-level PV generation and losses developed and analysed in this study, is publicly archived on Zenodo at <https://doi.org/10.5281/zenodo.18794230>. There are no restrictions on data availability.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Reporting on race, ethnicity, or other socially relevant groupings

Population characteristics

Recruitment

Ethics oversight

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

The global energy transition is predicated on solar photovoltaic (PV) power displacing fossil fuels to deliver projected climate and air quality benefits. However, the effectiveness of this process is limited if solar capacity merely supplements coal. This creates a self-defeating dynamic in the energy transition, a direct, physical drag that is currently unaccounted for in sustainability and climate mitigation models. To quantify this constraint, the study uses satellite observations and machine learning to develop a global, facility-level dataset of PV energy generation and atmospheric losses for 140,944 PV facilities. Facilities are grouped by country and region in a hierarchical design. A Loss-to-Growth ratio is calculated to assess the extent to which annual PV energy losses offset energy added by new PV capacity. The analysis includes regional trends in aerosol-induced PV losses from 2013 to 2023, and attribution to coal-related pollution based on aerosol composition, trends in pollutant levels, and spatial correlations between coal plant capacity and PV performance losses.

Research sample

The research sample includes 140,944 PV facilities and all coal-fired power plants operating globally. PV facilities were identified using a combination of global and regional inventories, OpenStreetMap records, and a custom machine learning model applied to Sentinel-2 satellite imagery. Coal plant data were obtained from the Global Coal Plant Tracker, maintained by Global Energy Monitor, which provides facility-level locations and generation capacity. The PV sample includes spatial footprints, installation dates, and annual estimates of energy generation and losses. These data were also aggregated at the country level for regional analysis. These samples support analysis of regional energy trends, spatial co-location between coal and solar infrastructure, and attribution of aerosol-induced PV performance losses to coal-related pollution.

Sampling strategy

The sample includes all available PV facilities and coal-fired power plants identified through existing inventories and satellite-based

Sampling strategy	detection. The aim was comprehensive global coverage rather than statistical sampling, and the resulting sample size is sufficient for the spatial and temporal analyses conducted in this study.
Data collection	We generated PV facility data primarily using machine learning applied to Sentinel-2 satellite imagery, complemented by global and regional inventories and OpenStreetMap records. Panel footprints were extracted using the Segment Anything Model (SAM), followed by manual verification. We obtained coal plant data from the Global Coal Plant Tracker maintained by Global Energy Monitor. Energy generation and atmospheric loss estimates were derived using the Modern-Era Retrospective Analysis for Research and Applications (MERRA-2) reanalysis. Pollutant concentrations, including sulphur dioxide and nitrogen dioxide, were obtained from the Ozone Monitoring Instrument (OMI) satellite dataset.
Timing and spatial scale	The PV facilities included in this study were detected and dated between 2017 and 2024. The analysis covers global spatial extent, with results aggregated at the facility, national, and regional levels.
Data exclusions	No data were excluded from the analysis.
Reproducibility	Results can be reproduced using publicly available satellite and atmospheric reanalysis data, and the global facility-level PV energy generation and losses data we have made publicly available.
Randomization	This study is based on satellite observations and does not involve group allocation. Randomization is not applicable.
Blinding	Blinding is not applicable.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A