Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

First close insight into global daily gapless 1 km $PM_{2.5}$ pollution, driving factors, and health impact

Jing Wei (weijing@umd.edu)
University of Maryland
Zhanqing Li (zhanqing@umd.edu)
University of Maryland
Alexei Lyapustin
NASA Goddard Space Flight Center
Jun Wang
University of Iowa
Oleg Dubovik
Université de Lille
Joel Schwartz
Harvard TH Chan School of Public Health
Lin Sun
Shandong University of Science and Technology
Chi Li
Washington University in St. Louis, St. Louis
Song Liu
Southern University of Science and Technology
Tong Zhu
Peking University

Research Article

Keywords: Fine particulate matter, daily 1 km resolution, machine learning, big data, mortality burden

Posted Date: February 28th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2626358/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Fine particulate matter ($PM_{2.5}$) is the largest environmental risk factor impacting human health. While $PM_{2.5}$ has been measured widely across the world, there has been no high-resolution and gapless global $PM_{2.5}$ data on a daily scale. We generate a global daily $PM_{2.5}$ concentration at 1 km resolution using satellite gap-filled aerosol products and machine learning. Daily $PM_{2.5}$ retrievals agreed well with ground measurements, with sample-, space-, and time-based cross-validated correlations of 0.93, 0.89, and 0.88, respectively. This enables us to unprecedentedly monitor the day-to-day variations of $PM_{2.5}$, exposure risk, and mortality burden around the globe. More than 96% of the days exceeded the World Health Organization (WHO) recommended daily air quality guidelines (AQG) level (15 μg m^{-3}) in 2020, and 99% of populated areas were exposed to $PM_{2.5}$ risk at least one day; in particular, the proportions are 91% and 64% similarly in 7 and 30 days, respectively. The annual population-weighted mean $PM_{2.5}$ concentration was 27.6 μg m^{-3} (~5.5 times higher than the WHO annual AQG level of 5 μg m^{-3}), resulting in estimated premature deaths of ~4.2 million people and accounting for ~6.6% of total global deaths. Substantial differences are noted in many parts of the world between 2019 and 2020 associated with widespread episodes of wildfires or the COVID-19 shutdowns. The overall air quality in 2020 was significantly better than in 2019 in more than 70% of major cities. The global population-weighted mean $PM_{2.5}$ decreased by ~5.1%, and the associated number of premature deaths dropped by 56,700.

Significance Statement

Unprecedented global daily 1-km gapless $PM_{2.5}$ concentrations were generated using machine learning integrating multi-source data from monitors, satellites, reanalyses, and models. This enables us to assess daily population-risk exposure to $PM_{2.5}$ at any location. We show that ~99%, 91%, and 64% of global populated areas exceeded the WHO short-term air quality guidelines level of 15 μg m^{-3} at least 1, 7, and 30 days in 2020, respectively. The differences rapidly grow with expanding the exposure periods between developed and developing countries. Contrasting patterns were seen between the western and eastern Hemispheres in changes of $PM_{2.5}$ and associated mortality driven by widespread wildfires and COVID-19 lockdown. Global $PM_{2.5}$ declined by ~5%, saving estimated ~56,700 lives in 2020 relative to 2019.

Introduction

Ambient air pollution poses a major global environmental concern, ranked as the fourth highest risk factor for human health, causing nearly 6.7 million global deaths in 2019 in the Global Burden of Disease (GBD) Study (1). Exposure to $PM_{2.5}$ (particulate matter with aerodynamic diameters ≤ 2.5 μm) pollution is associated with a variety of circulatory and respiratory diseases (2–5) and contributed to 4 million deaths globally in 2020. $PM_{2.5}$ emitted from different sources may have different health risks, especially ultrafine particles from urban traffic or combustion, including black carbon-containing particles from wildfire smoke (6–9). Besides long-term effects, short-term exposure to high levels of pollution is also a focal

point in environmental health studies (10–14). With the accumulating evidence of the health effect of $PM_{2.5}$, the World Health Organization (WHO) recently renewed its short-term and long-term air quality guidelines and aimed at continual reduction in air pollution in 2021 (15). Despite a steady increase in $PM_{2.5}$ monitoring capabilities over the last decade, many people still live far from monitors, and the lack of daily global high-resolution $PM_{2.5}$ data has been hindering our ability to assess air pollution exposure and health effects in any part of the world.

Regional $PM_{2.5}$ estimates have been extensively studied from the satellite perspective using retrieved aerosol optical depth (AOD) products (16–19), but only a handful of studies have been conducted on the global $PM_{2.5}$ scale. Lary et al. (20) trained an ensemble machine-learning model to estimate daily $PM_{2.5}$ in the early years of the Sea-viewing WideField-of-view Sensor and Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) Deep Blue AOD products at a 10-km resolution. van Donkelaar et al. (21) first derived global daily $PM_{2.5}$ from MODIS and Multiangle Imaging Spectroradiometer (MISR) AOD products via a vertical conversion factor determined from the 3-D Goddard Earth Observing System (GEOS)-Chem chemical transport model at a 10 km resolution. Although the spatial resolution has subsequently improved to 1 km by including MODIS Multi-Angle Implementation of Atmospheric Correction (MAIAC) AOD product (22) combing the Geographically Weighted Regression (GWR) and Chemical Transport Model (CTM), global $PM_{2.5}$ estimates are only available on annual (23, 24) or monthly (25) scales.

Monitoring global daily surface $PM_{2.5}$ from space at a fine resolution is highly sought but a very challenging task given that $PM_{2.5}$ air pollution can vary dynamically with space and time over land and thus exhibits strong spatiotemporal heterogeneities (26). Potential factors affecting $PM_{2.5}$ are complex and diverse: they are sensitive not only to pollution emission profiles, the density of the human population, and meteorological conditions, but also to regionally variant geographic location and terrain conditions (27). In addition, the uneven distribution of ground-based observation stations around the world, dense in developed countries but sparse or missing in most developing countries, leads to great challenges in spatial extrapolation (17). Lastly, ubiquitous clouds, when aerosols cannot be measured from space, result in similarly ubiquitous gaps (17, 28), which seriously hinders global observation of air pollution at daily frequency, especially under severe haze and heavy smoke conditions of particular concern to the public.

What is the local, regional, and global $PM_{2.5}$ concentration on a certain day? What are the daily exposure to $PM_{2.5}$ pollution and its associated impact on public health? How much is the impact of the Covid episode on local to global scales? Here, we address these questions by deriving global 1-km-resolution surface $PM_{2.5}$ amounts with full coverage at an unprecedented daily scale from 1 January 2019 to 31 December 2020 via machine learning (ML) using big data when the whole world experienced the most dramatic environmental changes in its modern history. A state-of-the-art ML allows for the integration of virtually all information pertaining to $PM_{2.5}$ from ground measurements, satellite remote sensing products, model simulation products of aerosols and gaseous pollutants from the CTM, as well as

population density, topography, atmospheric reanalysis of meteorological fields, pollutant emission inventory, and of the regional economic level of regional development. In particular, our ML-extended model takes advantage of model simulations and accounts for the spatiotemporal autocorrelation and differences in air pollutants. It first solves the problem of missing data in satellite aerosol products over cloudy and snow/ice surfaces and then uses machine learning to retrieve $PM_{2.5}$ from AOD accounting for aerosol physical mechanisms and chemical composition. This leads to spatially complete (i.e., gap-filled) and more accurate $PM_{2.5}$ information for assessments of exposure risk and health burden.

Results

Performance in AOD gap-filling

The spatial coverage of daily MAIAC AOD products is usually higher in the mid-latitudes (> 60%) but much lower (< 30%) near the equator due to frequent cloud contamination and at high latitudes because of consistent snow/ice cover. The average AOD spatial coverage over land is only ~ 36% (Figure S1). Filling in both satellite scanning gaps and missing values in MAIAC AOD retrievals with the developed machine learning model has significantly increased the availability of daily data to 100%, enabling us to obtain AOD information at each 1 km grid cell on any day (Figure S2). AOD is well reconstructed not only in clean air regions like North America and most of Europe but also over highly polluted and cloudy regions, including India and eastern China (outlined by red rectangles), with a higher fraction of missing retrievals in MAIAC AOD product. Our daily gap-filled AODs agree reasonably well with AErosol RObotic NETwork (AERONET) ground-based AOD measurements with an average correlation (R) of 0.73 and normalized root-mean-square error (NRMSE) of 0.96 from 2019 to 2020. The accuracy of reconstructed AOD is slightly lower than that of clear-sky AOD retrievals (R = 0.74, NRMSE = 0.88), while the retrievable $PM_{2.5}$ data samples are substantially increased by 2.8 times (N = 2,401,776) outweighs slight accuracy loss.

Performance in $PM_{2.5}$ retrieval

Our model can estimate daily $PM_{2.5}$ concentrations at most ground-based monitoring stations and predict daily $PM_{2.5}$ levels at locations where or when observations are not available, as confirmed by close agreements with out-of-sample, out-of-station, and out-of-day 10-fold cross-validated correlations (CV-R) greater than 0.7 at ~ 88%, 83%, and 80% of the stations, and low uncertainties with NRMSE less than 0.6 at ~ 87%, 79%, and 80% of the stations (Fig. 1a-f), respectively. In addition, our model also works well for most continents or sub-regions, especially over East Asia (CV-R = 0.85–0.93, NRMSE = 0.34–0.47) and Europe (CV-R = 0.84–0.89, NRMSE = 0.45–0.55) (Table S1). In general, the estimation and prediction accuracy (CV-R) of daily $PM_{2.5}$ retrievals over land are 0.93, 0.89, and 0.88, and the average uncertainties (NRMSE) are 0.49, 0.59, and 0.61 at the overall, spatial, and temporal scales (Fig. 1g-i), respectively. In particular, the accuracy has continuously increased with decreasing uncertainty on monthly (CV-R = 0.93–0.98, NRMSE = 0.24–0.4) and annual (CV-R = 0.96–0.99, NRMSE = 0.14–0.24) time scales over land

(Figure S3). Similar conclusions can also be drawn at the continental and regional scales (Table S1), assuring the data quality for assessing $PM_{2.5}$ pollution exposure and health burden.

Day-to-day gapless $PM_{2.5}$ variations

To portray the global variation pattern, Fig. 2 presents our 1 km $PM_{2.5}$ retrieval on an individual day, 8 October 2022. In general, the spatial patterns of satellite-derived $PM_{2.5}$ retrieval were highly consistent with the ground measurements regionally (Fig. 2a-h). Air quality was very good in Canada, the eastern United States (US), Europe, and Australia (Fig. 2c, d, f). By contrast, severe $PM_{2.5}$ pollution is observed in the Beijing-Tianjin-Hebei region in eastern China, along the Indo-Gangetic Plain in northern India (Fig. 2g, h), and in Zambia in southern Africa, especially around respective metropolis (i.e., Beijing, New Delhi, and Lusaka), where high population density and anthropogenic emissions are strongly linked. Extremely high $PM_{2.5}$ levels (> 100 μg m^{-3}) were observed in the western US (especially in California) and in central South America (Fig. 2a, b), caused mainly by severe wildfires emitting a large amount of smoke into the air. High $PM_{2.5}$ concentrations were also found in the Sahara in northern Africa (Fig. 2e) and the Taklimakan desert in northwestern China related to dust storms. Without gap-filling, the daily $PM_{2.5}$ map looks spotty (Figure S4 vs. Figure 2), making it harder to discern the spatial patterns of $PM_{2.5}$ pollution, especially on small scales, and increasing the probability of missing the high-pollution events.

The gap-filled data allows us to more accurately capture $PM_{2.5}$ pollution with strong day-to-day variations throughout the year. While the global and national $PM_{2.5}$ loadings may be misleading as pollution is inherently both regional and transboundary, these numbers still represent the severity of the air pollution problem at the national level and the disparity between developed and developing worlds (Fig. 3). It also reflects many phenomena, for instance, shallow planetary boundary layer (PBL) resulting in ubiquitous pollution maximum during northern winter. Daily data also help identify the development of extreme high-pollution events, where examples of haze episodes, wildfires, and dust storms in Pakistan, Brazil, and Nigeria are outlined by red dashed ellipses in Fig. 3g-i. Such data are particularly useful for continuous space-time monitoring of regional severe haze episodes in eastern China and South Asia (Figures S5 and S6). They also can well represent the dynamics of smoke emissions at different stages of wildfires, from initial ignition, to spreading and intensification, and to final suppression. Figures S7 and S8 give two examples of the wildfires in western North America on 20–29 September 2020 and eastern Australia's bushfires from 29 December 2019 to 9 January 2020.

Daily risk exposure to $PM_{2.5}$ pollution

Health effects, as well as health exposure guidelines, are generally characterized as long-term and short-term or acute (29, 30). While long-term exposure is generally well characterized by annual $PM_{2.5}$ (23), our novel daily global $PM_{2.5}$ data allow us to introduce the short-term exposure risk as the number (or percentage, %) of days in one year when $PM_{2.5}$ exceeds the WHO recommended short-term air quality

guidelines (15) over a given location. Alternatively, for every day, we can evaluate the proportion of the global populated area (using the 1 km × 1 km gridded population data of the world) with $PM_{2.5}$ pollution exceeding the recommended standards. Figure 4 shows the global assessment of the population $PM_{2.5}$ daily exposure risk at a fine 1-km^2 spatial scale in terms of WHO's four short-term interim targets (S-IT1 – S-IT4 with daily $PM_{2.5}$ = 75, 50, 35 and 25 μg m^{-3}) and short-term air quality guidelines (S-AQG) level of 15 μg m^{-3}. In general, except for densely populated areas like Indo-Gangetic and North China Plains, the relative number of days with severe pollution above the two S-IT1 and S-IT2 targets were less than 20% in most areas in 2020 (Fig. 4a, b). The proportions are much smaller in many developed countries in the North America (Fig. 3d) and Europe. With the continuous promotion of interim targets 3 and 4, both the covered areas and the respective exposure risk to a moderate $PM_{2.5}$ pollution increase (Fig. 4c, d). In terms of the S-AQG level (Fig. 4e), nearly the entire Earth's populated areas were exposed to $PM_{2.5}$ risks, including some developed countries like Italy (A), Poland and Slovakia, and mega-cities like Los Angeles in USA (B), with more than 40% of days above the S-AQG level. High daily exposure risks were concentrated in developing countries: In Pakistan, Bangladesh, India, and China (Fig. 3c, e), all days of the year 2020 exceeded the S-AQG level. Similar conditions were also observed in many big cities or localities, including Mexico City in Mexico (C), Región Metropolitana de Santiago in Chile (D), Sao Paulo in Brazil (E), Gauteng in South Africa (F), Tehran in Iran (G), Krung Thep in Thailand (H), Ho Chi Minh in Viet Nam (I), and Pyongyang in North Korea (J).

With the exception of a few days, global population-weighted $PM_{2.5}$ satisfied the S-IT1 and S-IT2 targets, but 21% and 52% of the days failed to meet the S-IT3 and S-IT4 targets, respectively, and ~ 97% of the days were above the S-AQG level in 2020 (Fig. 3a). Daily $PM_{2.5}$ of developed countries have achieved all short-term interim targets, yet 19% of the days exceeded the S-AQG (Fig. 3b). By contrast, although all the days of $PM_{2.5}$ in developing countries met the S-IT1 target, 7%, 31%, and 59% of the days failed to meet the other three targets, while not a single day has met the S-AQG level (Fig. 3c). In addition, ~ 35%, 55%, 72%, 90%, and 99% of Earth's populated areas exceeded the WHO S-IT1 through S-IT4 targets and S-AQG level, respectively, at least once in 2020. The respective proportions are 15%, 30%, 42%, 64%, and 91% for 7 days, and 5%, 14%, 23%, 35%, and 64% for 30 days in a year (Fig. 4). Significant differences exist between developed and developing countries: while the exposure areas for at least 1 day above S-AQG are comparable (98% vs. 100%), the difference rapidly widens with higher exposure risks, e.g., S-IT4 (85% vs. 95%), S-IT3 (60% vs. 84%), S-IT2 (41% vs. 70%), and S-IT1 (21% vs. 50%). This gap rapidly grows with expanding the exposure period to 7 and 30 days, emphasizing the greater short-term exposure risk in middle- or low-income counties. Arguably, when the exposure days are related to wildfire or agricultural biomass burning and come in succession rather than randomly during the year, they potentially may lead to more adverse health effects (8, 31, 32).

$PM_{2.5}$ abundance and driving factors

Regional contrasts in $PM_{2.5}$ patterns

Figure 5 shows the spatial distribution of global annual $PM_{2.5}$ (a) in 2020 and the difference in 2019 (b). $PM_{2.5}$ was generally low over vast regions and most developed countries in the Americas, Europe, and Oceania, while high pollution was found in major developing countries like northern India and eastern China, especially the northern winter (Figure S9c). Satellite-derived 1 km $PM_{2.5}$ data enable us to monitor fine-scale $PM_{2.5}$ pollution at the city level. We reveal $PM_{2.5}$ hot spots in many large cities around the world (Figure S10), including Los Angeles, USA (A), Monterrey and Mexico City, Mexico (B and C), Po Valley, Italy (D), Katowice, Poland (E), Ankara, Turkey (F), Moscow, Russia (G), Sao Paulo and Rio de Janeiro, Brazil (H and I), Bandung and Jakarta, Indonesia (J-K), Johannesburg, South Africa (L), Baghdad, Iraq (M), and Tehran, Tehran (N). Besides, expected high $PM_{2.5}$ loadings were also found in major deserts, including Sahara and Arabian peninsula, Taklimakan, Gobi, Thar, and others (Figure S10i-iv), especially in northern spring and winter when more frequent dust storms occur (Figure S9a, d), but we did not consider the large mineral dust aerosol sources in part because they represent natural predominantly coarse-mode aerosols. El Nino year 2020 stands alone, resulting in large regional contrast differences: stronger than usual wildfire activity (Figure S11) in the western part of North America (33–35) and biomass burning in the Amazon basin (36–38) created elevated levels and positive $PM_{2.5}$ anomalies over most of South America, especially in the dry fire seasons (Figures S12). Specifically, monthly $PM_{2.5}$ was 82% and 113% higher in the western US (September) and central South America (October) for the month with the most wildfire records in 2020 compared to 2019, while southeastern Australia saw a 79% $PM_{2.5}$ increase in December 2019 compared to late 2020 (Figure S13a-c).

COVID-19 impacts on $PM_{2.5}$ pollution

The above-mentioned seasonal effects, limited in their geographical extent, are superimposed on the longer-term near-global-scale reduction of pollution associated with the COVID-19 pandemic. The latter transpires as negative $PM_{2.5}$ anomalies seen in approximately 60% of global populated areas (Fig. 5b), especially during the first half of 2020 (Figure S12). Larger reductions were observed in the Eastern Hemisphere (separated by the meridian at 20°W and 160°E), e.g., most countries in Asia, Europe, equatorial and southern Africa, and the Middle East, leading to an overall 5.7% decrease in $PM_{2.5}$ due to greatly declined major pollutants discharged (39). We quantified $PM_{2.5}$ reductions from COVID-19 lockdowns, comparing 2020 with 2019 for the short-term control periods, which were different for different countries. Significant improvements in air quality were observed in China (23 January to 14 March) (40), India (25 March to 31 May) (41), and South Africa (May) (42), especially in major populated provinces or states including Shandong (↓31.7%), Uttar Pradesh (↓32.9%), and Gauteng (↓52.4%) (Figure S13d-f), attributing to the strict lockdown measures in early 2020 (39). Similar but smaller $PM_{2.5}$ reductions are found in eastern Europe (April) and the eastern US (15 March to 30 April) (Figure S13g, h) in agreement with (43, 44). COVID-19-related $PM_{2.5}$ reductions may have been offset by meteorological conditions in some regions in southwestern Europe and the southeastern US with a lack of shutdown enforcement (43, 44). Overall, the global population-weighted $PM_{2.5}$ in 2020 decreased by ~ 5.1% relative to 2019, and population-weighted $PM_{2.5}$ concentrations (within 20 × 20 km^2 around city centers) at more

than 70% of the global major cities (defined as urban agglomerations with populations greater than 300,000 in 2018, reported by the World Urbanization Prospects) also decreased, indicating global air quality has improved overall mainly due to the COVID-19 related measures.

City-, country-, and global-scale $PM_{2.5}$ level

Overall, among ~ 1860 major cities, 90% and 10% of the top-twenty most polluted cities were from India and Pakistan, of which Delhi, India, Lahore, Pakistan, and Lucknow (capital of Uttar Pradesh, India's most populous state) were the top three with annual population-weighted $PM_{2.5}$ concentrations of 96.6 μg m^{-3}, 95.3 μg m^{-3}, and 83.6 μg m^{-3} in 2020, respectively (Table S2). In 2020, the top-twenty countries with the highest $PM_{2.5}$ in the world came from North Africa (50%), the Middle East (25%), and South and East Asia (25%), among which Pakistan (50.1 μg m^{-3}), Bangladesh (47.9 μg m^{-3}), and Niger (46.9 μg m^{-3}) top the list (Table S2). To date, no country has reached the recommended long-term AQG (L-AQG: annual $PM_{2.5}$ = 5 μg m^{-3}) level, while ~ 71%, 43%, 13%, and 6% of the countries failed to reach the recommended long-term interim targets (L-IT4 – L-IT1 with annual $PM_{2.5}$ = 10, 15, 25 and 35 μg m^{-3}), respectively. The global population-weighted mean $PM_{2.5}$ concentration is ~ 27.6 μg m^{-3}, and the numbers are 29.1 and 13.1 μg m^{-3} for developing and developed countries, which are ~ 5.5, 5.8, and 2.6 times higher than the L-AQG level. Nearly 96%, 48%, 29%, 12%, and 6% of the global populated area exceeded the L-AQG level and four L-IT4 to L-IT1 targets, respectively (Fig. 5).

Mortality burden attributed to $PM_{2.5}$ pollution

Using this gapless $PM_{2.5}$ data, we estimate the number of premature deaths attributed to both daily and annual $PM_{2.5}$ exposure at each 1-km^2 grid in populated areas (Figure S14), using October 8, 2020, as an example. The high mortality estimates are concentrated in areas of high population density, which go hand in hand with high pollution emission rates from demands for transportation, energy, industry, and infrastructure. The zoomed-in areas show metropolitan regions of Los Angeles (A), Chicago (B), New York (C), USA, Madrid, Spain (D), Berlin, Germany (E), Île-de-France, France (F), and Moskovskaya Oblast, Russia (G), and Pearl River Delta (H), China. The total number of estimated global premature deaths was 14.5 [95% confidence interval (CI): 10.2, 18.6] thousand on this day. We further calculate the accumulated premature deaths attributed to short-term $PM_{2.5}$ exposure: intense wildfires in the US (September), Brazil (October), and Australia (December) led to additional deaths of 4.9 (95% CI: 3.4, 6.3), 3.6 (95% CI: 2.5, 4.5), and 0.7 (95% CI: 0.5, 0.9) thousand people between 2019 and 2020, respectively. The decrease in $PM_{2.5}$ pollution during COVID-19 lockdowns, however, can be associated with saving lives of 58.4 (95% CI: 44.0, 69.5), 64.7 (95% CI: 48.1, 78.0), and 2.5 (95% CI: 1.8, 3.1) thousand people in China, India, and South Africa, respectively.

In terms of the annual mortality burden (Fig. 3c), China, India, and Indonesia were the top three countries with the highest number of annual premature deaths in 2020, i.e., 1242.4 (95% CI: 947.0, 1525.3), 899.3 (95% CI: 709.7, 107.3), and 129.3 (95% CI: 88.5, 173.5) thousand, respectively. The total number of

premature deaths associated with $PM_{2.5}$ pollution around the world was ~ 4.2 (95% CI: 3.1, 5.4) million, accounting for ~ 6.6% of the total number of deaths in 2020. Compared to 2019, wildfire smoke-related $PM_{2.5}$ emissions increased the number of premature deaths by ~ 17.1 (95% CI: 12.3, 21.4) thousand in the Western Hemisphere, especially in the US [↑6.3 (95% CI: 4.3, 8.1) thousand] and Brazil [↑4.2 (95% CI: 3.3, 5.0) thousand] and some countries in central Africa, especially in the Democratic Republic of Congo [↑2.6 (95% CI: 2.0, 3.0) thousand]. The number of premature deaths decreased in most countries in the Eastern Hemisphere [average = ↓73.8 (95% CI: 52.5, 84.4) thousand], especially in China [↓59.3 (95% CI: 50.8, 61.6) thousand] benefiting from reduced anthropogenic emissions (Fig. 3d). Globally, the number of $PM_{2.5}$-related premature deaths had decreased by ~ 56.7 (95% CI: 31.2, 72.1)] thousand as a result of the improved air quality due to COVID-19, which, itself, accounts for 3.3 million deaths in 2020 according to WHO (45).

Discussion

Effects of AOD gap-filling

Spatiotemporal continuity in $PM_{2.5}$ data is critical for both air pollution exposure and health burden assessments. However, satellite AOD retrievals are not performed over cloudy and snow/ice-covered scenes and may not be available for some heavy pollution episodes when they are masked as clouds (27). From a long-term perspective, missing AOD creates large gaps at high latitudes during northern winter and spring, where air quality is generally good and, if unaccounted for, would lead to overestimations of annual $PM_{2.5}$ (Figure S15). Global air pollution is overestimated at low latitudes but tends to be underestimated in the middle latitudes, especially in areas with high population densities like eastern China and the US. In general, the annual global population-weighted $PM_{2.5}$ is overestimated by ~ 9% in 2020 if gap-filling is not applied. The short-term difference, e.g., for the daily population-weighted $PM_{2.5}$, can be as high as 19–79% (average = 54%). Some daily trends could even be reversed at the beginning or end of the year (outlined by black dotted line boxes in Figure S16), which is important for years with an increase in mega-fires (e.g., El Nino years). Similarly, there was a large impact on the mortality burden, e.g., without AOD gap-filling, the global estimated premature deaths were underestimated by ~ 10% (95% CI: 8%, 13%) in 2020, and the number was underestimated by ~ 67% (95% CI: 66%, 68%) during the short-term COVID-19 lockdown in China.

Comparison with other global-scale studies

Only a few studies have estimated global $PM_{2.5}$ (Table S3), for example, chemical transport models have simulated the spatiotemporal variations of global $PM_{2.5}$ at high-frequency (1 to 3 hourly) [e.g., Modern-Era Retrospective analysis for Research and Applications, version 2 and Copernicus Atmosphere Monitoring Service (CAMS)], but the spatial resolution is much coarser (> 0.25^{o}), which usually differ considerably from ground observations partially for using much fewer observational data (46, 47). An attempt has been made to correct CAMS $PM_{2.5}$ simulations by using more ground observations through

deep learning, with a small improvement (CV-R = 0.58–0.61) (46). A high-quality $PM_{2.5}$ dataset at the annual scale (CV-R = 0.95) was generated by combining the Bayesian hierarchical model and the GEOS-Chem chemical transport model (48). Many studies have demonstrated that satellite remote sensing products of AOD provide more pertinent information to improve the estimation of $PM_{2.5}$ concentrations retrievals (19–21, 23, 49). Using satellite observations, combined with GEOS-Chem (21) and ensemble machine-learning (20) models, the estimation accuracy (CV-R = 0.51–0.77) and spatial resolution (10 km) of daily $PM_{2.5}$ estimates were improved. Later, 1 km fine-scale $PM_{2.5}$ maps with improved correlations (R = 0.90–0.96) were generated using the GWR model but only for annual (23) and monthly (25) mean concentrations. In comparison with these estimates made with the chemical transport model, statistical regression, and machine-learning models, our product is unique for having the highest resolution of 1 km on a daily basis, gap-free, and highest accuracy to date (CV-R = 0.93–0.99). Additionally, we compared our data set with the only publicly available monthly Global $PM_{2.5}$ Estimates (V5.GL.02) data set. While they bear similar spatial distribution patterns, our monthly estimates match more closely (e.g., R = 0.99, slope = 0.96) than the above-mentioned dataset (e.g., R = 0.91, slope = 0.88) (25) with ground measurements from 2019 to 2020 (Figure S17).

Conclusions

Taking advantage of 1-km high-quality MAIAC AOD products and spatially complete GEOS Composition Forecast (GEOS-CF) AOD simulations, unlike previous studies, we made the gap filling of global daily AODs; in combination with GEOS-CF $PM_{2.5}$ simulations, we then, for the first time, generated global daily gapless 1 km $PM_{2.5}$ data. These operations were performed under the veil of big data, including additional inputs of worldwide ground-based observations, surface- and population-related satellite remote sensing products, meteorological reanalysis, and emission inventories, all driven by a developed high-performance ensemble learning model, which considers the spatiotemporal heterogeneities of air pollutants during the modeling to improve the predictive ability, robustness, and stability of the model. This arduous task was realized on a cloud-computing platform (Google Earth Engine).

The generated dataset enables us to monitor global daily $PM_{2.5}$ distributions and variations, and investigate the effects of short-term $PM_{2.5}$ pollution on public health (50–52), especially at the finer city level (53). Independent cross-validated results demonstrated that our global $PM_{2.5}$ estimates and predictions agree well with ground measurements at varying spatiotemporal scales. Using this reliable and fine-resolution dataset, we investigated the global $PM_{2.5}$ change, exposure risk, and associated mortality burden on a 1 km grid from both short-term and long-term exposure perspectives, also its unique value in seeing the dynamics of the impact of the COVID-19 pandemic on $PM_{2.5}$ pollution in terms of the spatial and temporal variations as it is spreading around the world. Global $PM_{2.5}$ and its related premature deaths decreased on the whole in 2020, presumably associated with COVID-19 lockdown measures, while the spike in forest mega-fires in 2020 has reversed the trend in some countries in the Western Hemisphere (e.g., the US, Brazil, and Mexico). In 2020, global fractions of the Earth populated

area with daily $PM_{2.5}$ exposure exceeding WHO recommended short-term thresholds of 15, 25, 37.5, 50, and 75 μg m^{-3} at least on one day was 99%, 90%, 72%, 55%, and 35%, respectively. Thus, achieving the objectives of the WHO air quality guidelines globally remains a challenge that requires a joint international effort. Having demonstrated the unique merits of our model and global $PM_{2.5}$ product, we need to mention limitations caused by unevenly distributed and scarce observations in developing world regions, particularly in Africa and South America. $PM_{2.5}$ predictions in these regions can only be truly tested when new observations become available. $PM_{2.5}$ measurements from low-cost sensors may help in these regions, but their data quality remains uncertain (54, 55).

Materials And Methods

$PM_{2.5}$ measurements

Hourly $PM_{2.5}$ measurements in 2019 and 2020 were collected at more than ~ 9300 ground-based monitoring stations around the world (Figure S18). Sources of the data include OpenAQ, the China National Environmental Monitoring Centre, the US Environmental Protection Agency, the Canadian National Air Pollution Surveillance Program, and the European Air Quality e-Reporting. Days having at least eight valid hourly $PM_{2.5}$ measurements were first identified, whose values were averaged to obtain daily means at each monitoring site and used for independent training and validation of machine-learning models.

AOD data

The MAIAC AOD product (MCD19A2) (22) has the highest spatial resolution of 1 km among all MODIS operational aerosol products over land with high accuracy (56, 57). The data were selected as the main predictor for estimating $PM_{2.5}$ concentrations. Terra (~ 10:30 AM overpass time) and Aqua (~ 13:30 PM overpass time) MAIAC daily 1-km AOD retrievals at 550 nm with recommended quality assurances (58) from 2019 to 2020 were used. Spatially complete AOD assimilations (550 nm) from the GEOS Forward Processing (FP) system every 3 hours at a horizontal resolution of 0.25° × 0.3125° were employed to fill satellite AOD gaps. The global AERONET network provides instantaneous AOD measurements made at ~ 500 stations over land every 15 min (Figure S18). Only the recommended high-quality (Level 2.0) AOD data (59) at 550 nm interpolated by the quadratic polynomial fitting method (60) were used to validate the satellite gap-filled AODs.

Auxiliary data

First, hourly $PM_{2.5}$ simulations from the GEOS-CF system at a horizontal resolution of 0.25° × 0.25° were employed (61). Considering the secondary formation via chemical reaction, four main precursors of $PM_{2.5}$, i.e., ammonia, nitrogen oxides, sulfur dioxide, and volatile organic compounds, are also involved from the CAMS global high-resolution (~ 0.1° × 0.1°, monthly) emission inventory (62). Meteorological variables and their vertical profiles affect air pollution, including temperature, humidity, wind, pressure,

precipitation, evaporation, and boundary-layer height. Therefore, these meteorological fields were collected from hourly ERA5-Land (~ 0.1° × 0.1°) and ERA5 global (~ 0.25° × 0.25°) reanalysis datasets (63, 64). In addition, population density, economic level, land cover, and terrain changes impacting air pollution were also considered, which are directly or indirectly represented by the highly relevant and available satellite remote sensing products, i.e., global high-resolution annual WorldPOP population (1 km resolution) (65), monthly Visible Infrared Imaging Radiometer Suite (VIIRS) nighttime lights (500 m resolution) (66), monthly MOD13A3 normalized difference vegetation index (NDVI; 1 km resolution), and Shuttle Radar Topography Mission digital elevation model (DEM; 90 m resolution) products, respectively. In total, 18 independent variables were included in this study (Table S4), and the links for all used data sources are provided in Text S1. Here, all variables are transformed to uniform 0.01° × 0.01° grids used for subsequent air pollution modeling.

Air pollution modelling

Here, a tree-based ensemble-learning extremely randomized trees (extra-trees) (67), which has unique advantages with stronger randomness and an anti-interference ability with reference to other similar types of models of superior performance (68), was adopted for modelling air pollutants. Previous studies have illustrated that the model performance can be significantly improved if the spatiotemporal heterogeneity of air pollution is considered during modeling (69). Therefore, for global modeling, different from our previous studies (19, 68), we developed a new 4-Dimensional Space-Time Extra-Trees (4D-STET) model by introducing Euclidean spherical space and triangular spiral time into the original machine-learning model to better describe autocorrelations and differences in spatial locations (e.g., different hemispheres) and temporal series (e.g., seasonal cycles) (Text S2).

The framework includes two steps: 1) for satellite AOD gap-filling, Terra and Aqua MAIAC AOD retrievals were first combined via linear regression conversion models to minimize the biases caused by different observation times and to expand the spatial coverage (58, 70). Available MAIAC AOD retrievals were then used as true values to train the 4D-STET model together with the main GEOS-FP AOD simulations and potentially influencing factors, i.e., spatially continuous meteorological fields (i.e., boundary-layer height, temperature, humidity, wind, and pressure), land cover, and elevation, as well as spatiotemporal terms. Missing satellite retrievals can thus be computed and scanning gaps filled to generate gapless AODs on a daily basis (Text S3); 2) For surface $PM_{2.5}$ estimation, their measurements were regarded as targets, and satellite gap-filled AODs, GEOS-CF $PM_{2.5}$ simulations, and CAMS pollutant emissions were regarded as main predictors. Auxiliary variables affecting $PM_{2.5}$ pollution, including all meteorological fields (i.e., boundary-layer height, temperature, humidity, wind, pressure, precipitation, and evaporation), NDVI, DEM, population, and spatiotemporal terms, were input to the 4D-STET model to construct a robust AOD-$PM_{2.5}$ relationship (Text S3). Models were trained for 2019 (sample size = ~ 1,155,063) and 2020 (sample size = ~ 1,246,713) separately with sufficient data samples.

Model validation

Evaluation of global AOD gap-filling was performed independently against the AERONET AOD measurements, while $PM_{2.5}$ surveillance was done at three independently sample-based (out-of-bag), space-based (out-of-station), and time-based (out-of-day) levels using the widely used 10-fold cross-validation method (71), i.e., all data samples, stations, and days were randomly divided into ten folds, of which nine (90%) were used for modeling and one (10%) for validation, respectively. This process was run ten times, in turn, to get the final accuracy after averaging. These methods ensure that the training and test data are independent at overall, spatial, and temporal scales, which are used to evaluate the overall accuracy of the model and its predictive ability in areas and on days without any available observations, respectively (58).

Exposure risk assessment

Population-health exposure risks related to ambient $PM_{2.5}$ pollution were assessed according to the new global air quality guidelines updated by the WHO in 2021 (15). Short-term exposure risks were calculated by counting the proportion of days in a year with daily $PM_{2.5}$ concentrations exceeding the WHO-recommended short-term AQG level (daily $PM_{2.5}$ = 15 μg m^{-3}) and four interim targets, i.e., IT4 (daily $PM_{2.5}$ = 25 μg m^{-3}), IT3 (daily $PM_{2.5}$ = 37.5 μg m^{-3}), IT2 (daily $PM_{2.5}$ = 50 μg m^{-3}), and IT1 (daily $PM_{2.5}$ = 75 μg m^{-3}). Long-term exposure risks were calculated by counting the proportion of areas with annual $PM_{2.5}$ concentrations exceeding the WHO-recommended long-term AQG level (annual $PM_{2.5}$ = 5 μg m^{-3}) and interim targets: IT4 (annual $PM_{2.5}$ = 10 μg m^{-3}), IT3 (annual $PM_{2.5}$ = 15 μg m^{-3}), IT2 (annual $PM_{2.5}$ = 25 μg m^{-3}), and IT1 (annual $PM_{2.5}$ = 35 μg m^{-3}).

Mortality burden estimation

Long-term mortality burden attributable to ambient $PM_{2.5}$ pollution is estimated globally at a 1-km resolution by adopting the exposure-response function collected from the updated 2019 Global Burden of Disease (GBD) study (1) that integrated the annual $PM_{2.5}$ exposure level, population distribution, relative risk, and baseline mortality (Text S4). Annual premature deaths for adults (ages > 25 years old) are the sum of mortality from six diseases, including chronic obstructive pulmonary disease, diabetes (Type 2), ischemic heart disease, lung cancer, lower respiratory infection, and stroke. The population density was provided by the 1-km WorldPOP product, while the age-populated proportion, relative risk, and baseline mortality for each age-specific disease were collected from the 2019 GBD study (1). The health impact of exposure to short-term $PM_{2.5}$ pollution is attracting more and more public attention, but unlike long-term studies, no unified exposure-response function is available (10–13). The short-term mortality burden attributable to daily $PM_{2.5}$ exposure was thus assessed at a 1-km resolution by employing a different linear exposure-response function determined by a meta-analysis from a recent review study (14), e.g., the relative risk for mortality from all causes is 1.0065 (95% CI: 1.0044–1.0086) with an increase in daily $PM_{2.5}$ concentration per 10 μg/m^3 (Text S4).

Declarations

Data availability

The global daily gap-filled high-resolution (1 km) and high-quality $PM_{2.5}$ (GlobalHigh$PM_{2.5}$) dataset generated in this study is available at https://doi.org/10.5281/zenodo.6449740, which will be publicly available once this paper is published.

Acknowledgements

This work was in part supported by the NASA Applied Science Program (80NSSC21K1980) and NASA MODIS algorithm maintenance funding. The authors greatly thank Huanbi Yue from the Ocean University of China for helping calculate the mortality burden, Lorena Castro from the University of Iowa for collecting and processing the OpenAQ data, and Maureen Cribb from the University of Maryland for editing and polishing the paper.

Competing interests: The authors declare no competing interests.

References

1. Murray CJL, *et al.* (2020) 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(10258):1223–1249.
2. Southerland VA, *et al.* (2022) Global urban temporal trends in fine particulate matter (PM2·5) and attributable health burdens: estimates from global datasets. The Lancet Planetary Health 6(2):e139-e146.
3. Bowe B, *et al.* (2018) The 2016 global and national burden of diabetes mellitus attributable to PM2·5 air pollution. The Lancet Planetary Health 2(7):e301-e312.
4. Pope CA, *et al.* (2011) Lung Cancer and Cardiovascular Disease Mortality Associated with Ambient Air Pollution and Cigarette Smoke: Shape of the exposure-response Relationships. 119(11):1616–1621.
5. Turner MC, *et al.* (2020) Outdoor air pollution and cancer: An overview of the current evidence and public health recommendations. 70(6):460–479.
6. Schraufnagel DE (2020) The health effects of ultrafine particles. Experimental & Molecular Medicine 52(3):311–317.
7. Sinharay R, *et al.* (2018) Respiratory and cardiovascular responses to walking down a traffic-polluted road compared with walking in a traffic-free area in participants aged 60 years and older with chronic lung or heart disease and age-matched healthy controls: a randomised, crossover study. The Lancet 391(10118):339–349.
8. Aguilera R, Corringham T, Gershunov A, & Benmarhnia T (2021) Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California. Nature Communications 12(1):1493.

9. Chen G, *et al.* (2021) Mortality risk attributable to wildfire-related PM2.5 pollution: a global time series study in 749 locations. The Lancet Planetary Health 5(9):e579-e587.
10. Chen R, *et al.* (2017) Fine Particulate Air Pollution and Daily Mortality. A Nationwide Analysis in 272 Chinese Cities. American journal of respiratory and critical care medicine 196(1):73–81.
11. Di Q, *et al.* (2017) Association of Short-term Exposure to Air Pollution With Mortality in Older Adults. Jama 318(24):2446–2456.
12. Kloog I, Ridgway B, Koutrakis P, Coull BA, & Schwartz JD (2013) Long- and Short-Term Exposure to PM2.5 and Mortality: Using Novel Exposure Models. Epidemiology 24(4).
13. Liu C, *et al.* (2019) Ambient Particulate Air Pollution and Daily Mortality in 652 Cities. 381(8):705–715.
14. Orellano P, Reynoso J, Quaranta N, Bardach A, & Ciapponi A (2020) Short-term exposure to particulate matter (PM10 and PM2.5), nitrogen dioxide (NO2), and ozone (O3) and all-cause and cause-specific mortality: Systematic review and meta-analysis. Environment International 142:105876.
15. WHO (2021) WHO global air quality guidelines. Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. *Geneva: World Health Organization*:Licence: CC BY-NC-SA 3.0 IGO.
16. Li Y, *et al.* (2021) Satellite Remote Sensing for Estimating PM2.5 and Its Components. Current Pollution Reports 7(1):72–87.
17. Shin M, *et al.* (2020) Estimating ground-level particulate matter concentrations using satellite-based data: a review. GIScience & Remote Sensing 57(2):174–189.
18. Xu X, Zhang C, & Liang Y (2021) Review of satellite-driven statistical models PM2.5 concentration estimation with comprehensive information. Atmospheric Environment 256:118302.
19. Wei J, *et al.* (2021) Reconstructing 1-km-resolution high-quality PM2.5 data records from 2000 to 2018 in China: spatiotemporal variations and policy implications. Remote Sensing of Environment 252:112136.
20. Lary DJ, *et al.* (2014) Estimating the global abundance of ground level presence of particulate matter (PM2.5). Geospatial health 8(3):S611-630.
21. Donkelaar Av, *et al.* (2010) Global Estimates of Ambient Fine Particulate Matter Concentrations from Satellite-Based Aerosol Optical Depth: Development and Application. 118(6):847–855.
22. Lyapustin A, Wang Y, Korkin S, & Huang D (2018) MODIS Collection 6 MAIAC algorithm. Atmos. Meas. Tech. 11(10):5741–5765.
23. van Donkelaar A, *et al.* (2016) Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors. Environmental Science & Technology 50(7):3762–3772.
24. Hammer MS, *et al.* (2020) Global Estimates and Long-Term Trends of Fine Particulate Matter Concentrations (1998–2018). Environ Sci Technol 54(13):7879–7890.

25. van Donkelaar A, *et al.* (2021) Monthly Global Estimates of Fine Particulate Matter and Their Uncertainty. Environ Sci Technol 55(22):15287–15300.
26. Wei J, *et al.* (2022) Ground-Level NO2 Surveillance from Space Across China for High Resolution Using Interpretable Spatiotemporally Weighted Artificial Intelligence. *Environmental Science & Technology*.
27. Liang FC, *et al.* (2020) The 17-y spatiotemporal trend of PM2.5 and its mortality burden in China. Proceedings of the National Academy of Sciences of the United States of America 117(41):25601–25608.
28. Christopher SA & Gupta P (2010) Satellite Remote Sensing of Particulate Matter Air Quality: The Cloud-Cover Problem. Journal of the Air & Waste Management Association 60(5):596–602.
29. Shi L, *et al.* (2016) Low-Concentration PM < sub > 2.5</sub > and Mortality: Estimating Acute and Chronic Effects in a Population-Based Study. 124(1):46–52.
30. EPA (2020) Review of the National Ambient Air Quality Standards for Particulate Matter. *Federal Register* 85(244).
31. Sigsgaard T, *et al.* (2015) Health impacts of anthropogenic biomass burning in the developed world. 46(6):1577–1588.
32. Karanasiou A, *et al.* (2021) Short-term health effects from outdoor exposure to biomass burning emissions: A review. Science of The Total Environment 781:146739.
33. Ren L, *et al.* (2022) Widespread wildfires over the western United States in 2020 linked to emissions reductions during COVID-19. n/a(n/a):e2022GL099308.
34. Findlay A (2022) Pantanal fires. Nature Climate Change 12(2):118–118.
35. Zhou X, *et al.* (2021) Excess of COVID-19 cases and deaths due to fine particulate matter exposure during the 2020 wildfires in the United States. 7(33):eabi8789.
36. van der Velde IR, *et al.* (2021) Vast CO2 release from Australian fires in 2019–2020 constrained by satellite. Nature 597(7876):366–369.
37. Ward M, *et al.* (2020) Impact of 2019–2020 mega-fires on Australian fauna habitat. Nature Ecology & Evolution 4(10):1321–1326.
38. Godfree RC, *et al.* (2021) Implications of the 2019–2020 megafires for the biogeography and conservation of Australian vegetation. Nature Communications 12(1):1023.
39. Cooper MJ, *et al.* (2022) Global fine-scale changes in ambient NO2 during COVID-19 lockdowns. Nature 601(7893):380–387.
40. Liu F, *et al.* (2020) Abrupt decline in tropospheric nitrogen dioxide over China after the outbreak of COVID-19. 6(28):eabc2992.
41. Soni P (2021) Effects of COVID-19 lockdown phases in India: an atmospheric perspective. Environment, development and sustainability 23(8):12044–12055.
42. Haider N, *et al.* (2020) Lockdown measures in response to COVID-19 in nine sub-Saharan African countries. 5(10):e003319.

43. Hammer MS, *et al.* (2021) Effects of COVID-19 lockdowns on fine particulate matter concentrations. 7(26):eabg7670.
44. Goldberg DL, *et al.* (2020) Disentangling the Impact of the COVID-19 Lockdowns on Urban NO2 From Natural Variability. 47(17):e2020GL089269.
45. Wang H, *et al.* (2022) Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. The Lancet 399(10334):1513–1536.
46. Xiao Y, Wang Y, Yuan Q, He J, & Zhang L (2022) Generating a long-term (2003 – 2020) hourly 0.25° global PM2.5 dataset via spatiotemporal downscaling of CAMS with deep learning (DeepCAMS). Science of The Total Environment 848:157747.
47. Jin C, Wang Y, Li T, & Yuan Q (2022) Global validation and hybrid calibration of CAMS and MERRA-2 PM2.5 reanalysis products based on OpenAQ platform. Atmospheric Environment 274:118972.
48. Shaddick G, *et al.* (2018) Data Integration for the Assessment of Population Exposure to Ambient Air Pollution for Global Burden of Disease Assessment. Environmental Science & Technology 52(16):9069–9078.
49. Park Y, *et al.* (2020) Estimating PM2.5 concentration of the conterminous United States via interpretable convolutional neural networks. Environmental Pollution 256.
50. Wei Y, *et al.* (2019) Short term exposure to fine particulate matter and hospital admission risks and costs in the Medicare population: time stratified, case crossover study. 367:l6258.
51. Xie Y, *et al.* (2021) Short-Term Ambient Particulate Air Pollution and Hospitalization Expenditures of Cause-Specific Cardiorespiratory Diseases in China: A Multicity Analysis. The Lancet Regional Health – Western Pacific 15.
52. Giani P, *et al.* (2020) Short-term and long-term health impacts of air pollution reductions from COVID-19 lockdowns in China and Europe: a modelling study. The Lancet Planetary Health 4(10):e474-e482.
53. Southerland VA, *et al.* (2022) Global urban temporal trends in fine particulate matter (PM2.5) and attributable health burdens: estimates from global datasets. The Lancet Planetary Health 6(2):e139-e146.
54. Bi J, *et al.* (2022) Publicly available low-cost sensor measurements for PM2.5 exposure modeling: Guidance for monitor deployment and data selection. Environment International 158:106897.
55. Bi J, Wildani A, Chang HH, & Liu Y (2020) Incorporating Low-Cost Sensor Measurements into High-Resolution PM2.5 Modeling at a Large Spatial Scale. Environmental Science & Technology 54(4):2152–2162.
56. Qin W, *et al.* (2021) MODIS high-resolution MAIAC aerosol product: Global validation and analysis. Atmospheric Environment 264:118684.
57. Liu N, *et al.* (2019) Evaluation and comparison of multiangle implementation of the atmospheric correction algorithm, Dark Target, and Deep Blue aerosol products over China. Atmos. Chem. Phys. 19(12):8243–8268.

58. Wei J, *et al.* (2019) Estimating 1-km-resolution PM2.5 concentrations across China using the space-time random forest approach. Remote Sensing of Environment 231.
59. Giles DM, *et al.* (2019) Advancements in the Aerosol Robotic Network (AERONET) Version 3 database – automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. Atmos. Meas. Tech. 12(1):169–209.
60. Wei J, *et al.* (2020) MODIS Collection 6.1 3 km resolution aerosol optical depth product: global evaluation and uncertainty analysis. Atmospheric Environment 240:117768.
61. Keller CA, *et al.* (2021) Description of the NASA GEOS Composition Forecast Modeling System GEOS-CF v1.0. 13(4):e2020MS002413.
62. Crippa M, *et al.* (2018) Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4.3.2. Earth Syst. Sci. Data 10(4):1987–2013.
63. Hersbach H, *et al.* (2020) The ERA5 global reanalysis. 146(730):1999–2049.
64. Muñoz-Sabater J, *et al.* (2021) ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13(9):4349–4383.
65. Lloyd CT, *et al.* (2019) Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. Big earth data 3(2):108–139.
66. Elvidge CD, Zhizhin M, Ghosh T, Hsu F-C, & Taneja J (2021) Annual Time Series of Global VIIRS Nighttime Lights Derived from Monthly Averages: 2012 to 2019. 13(5):922.
67. Geurts P, Ernst D, & Wehenkel L (2006) Extremely Randomized Trees. 36:3–42.
68. Wei J, *et al.* (2021) Himawari-8-derived diurnal variations in ground-level PM2.5 pollution across China using the fast space-time Light Gradient Boosting Machine (LightGBM). Atmos. Chem. Phys. 21(10):7863–7880.
69. Wei J, *et al.* (2019) Estimating 1-km-resolution PM2.5 concentrations across China using the space-time random forest approach. Remote Sensing of Environment 231:111221.
70. Hu X, *et al.* (2014) Estimating ground-level PM2.5 concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model. Remote Sensing of Environment 140:220–232.
71. Rodriguez JD, Perez A, & Lozano JA (2010) Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence 32(3):569–575.

Supplementary Tables

Supplementary Tables S1-S4 are not available with this version

Supplementary Figures

Supplementary Figures S1-S18 are not available with this version

Figures

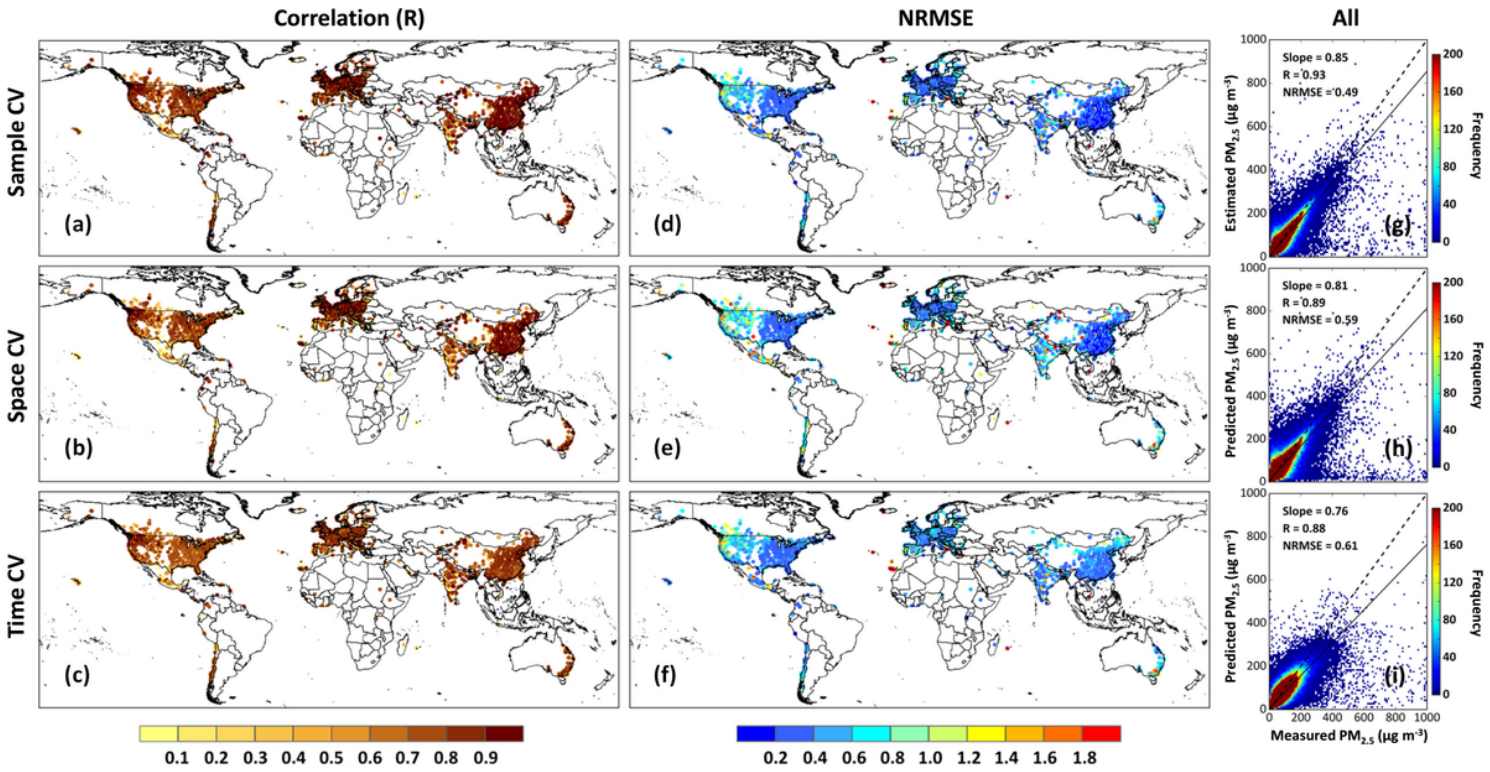


Figure 1

Spatial distributions of sample-based (top row), space-based (middle row), and time-based (bottom row) cross-validated (CV) (a-c) correlation (R) and (d-f) uncertainty (NRMSE) of daily $PM_{2.5}$ estimates and predictions (μg m^{-3}) against ground-based measurements (μg m^{-3}) at each monitoring station (only those with at least ten data samples were plotted), and (g-i) all monitoring stations from 2019 to 2020 over land (number of samples = 2,401,776).

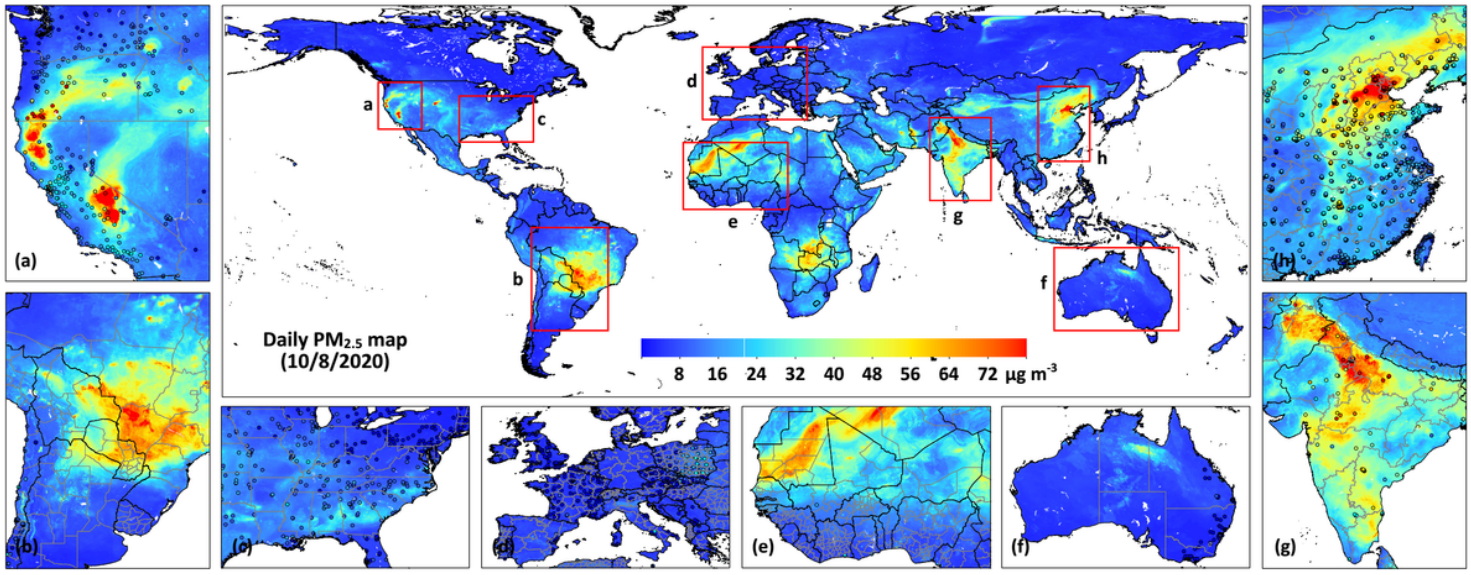


Figure 2

Satellite-derived global $PM_{2.5}$ concentrations (unit: μg m^{-3}) on 8 October 2020 at 1 km spatial resolution and zoomed-in maps (outlined by red rectangles) showing ground-based $PM_{2.5}$ measurements (colored dots) over the (a) western United States, (b) central South America, (c) eastern United States, (d) Europe, (e) northwestern Africa, (f) Australia, (g) eastern China, and (h) India. Thin black lines represent country boundaries or shorelines, and gray lines represent the state or provincial boundaries.

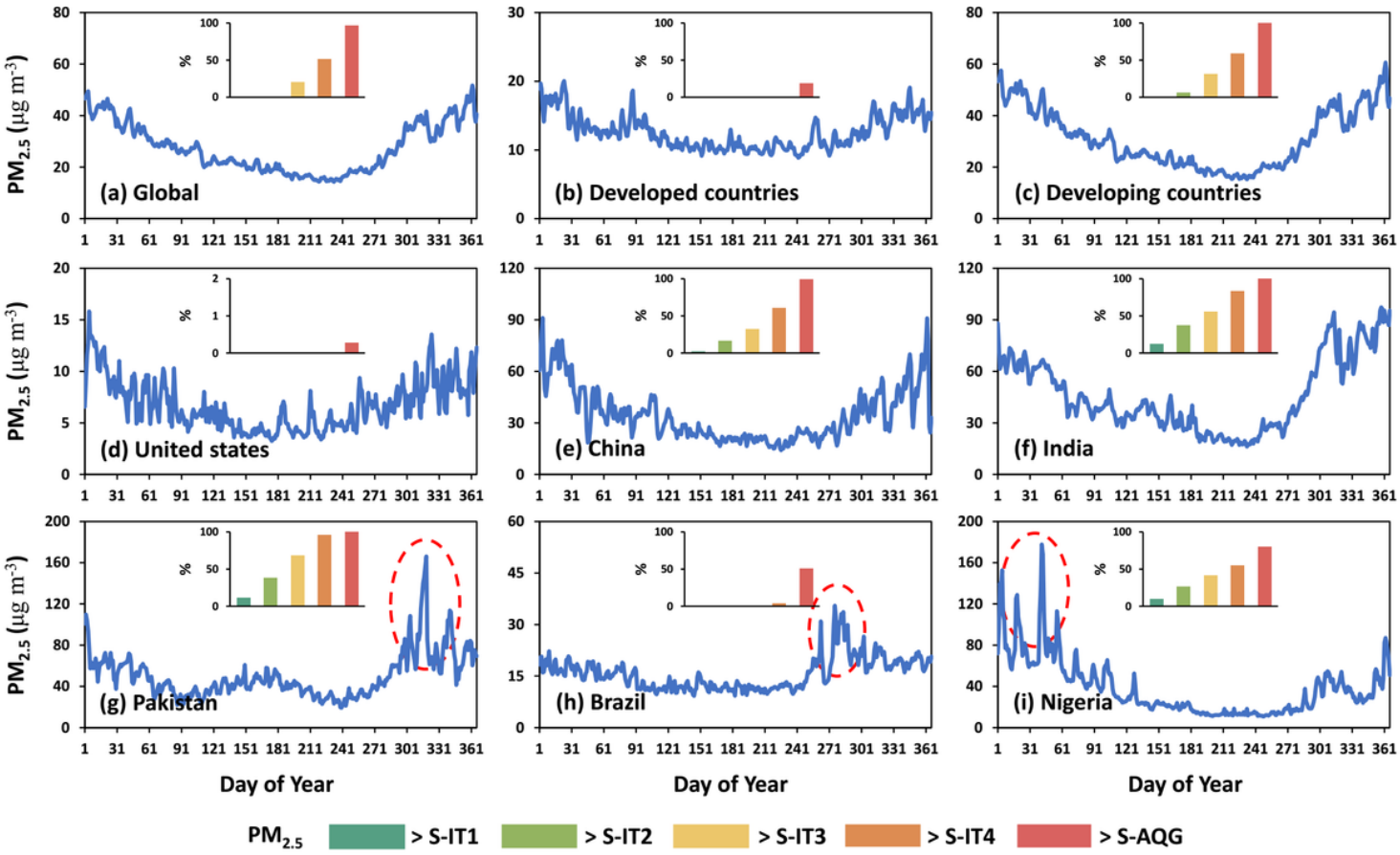


Figure 3

Time series of daily population-weighted $PM_{2.5}$ concentrations (unit: μg m^{-3}) as a function of the day of year (DOY) in 2020 for the global world, developed and developing countries, and seven selected countries, where red dashed ellipses outline days with anomalously heavy pollution. Inserted bar charts count the percentages of the days (unit: %) exceeding the World Health Organization (WHO) recommended short-term four interim targets [S-IT1(daily $PM_{2.5}$ > 75 μg m^{-3}), S-IT2 (daily $PM_{2.5}$ > 50 μg m^{-3}), S-IT3 (daily $PM_{2.5}$ > 37.5 μg m^{-3}), and S-IT4 (daily $PM_{2.5}$ > 25 μg m^{-3})] and air quality guidelines (S-AQG) level (daily $PM_{2.5}$ > 15 μg m^{-3}).

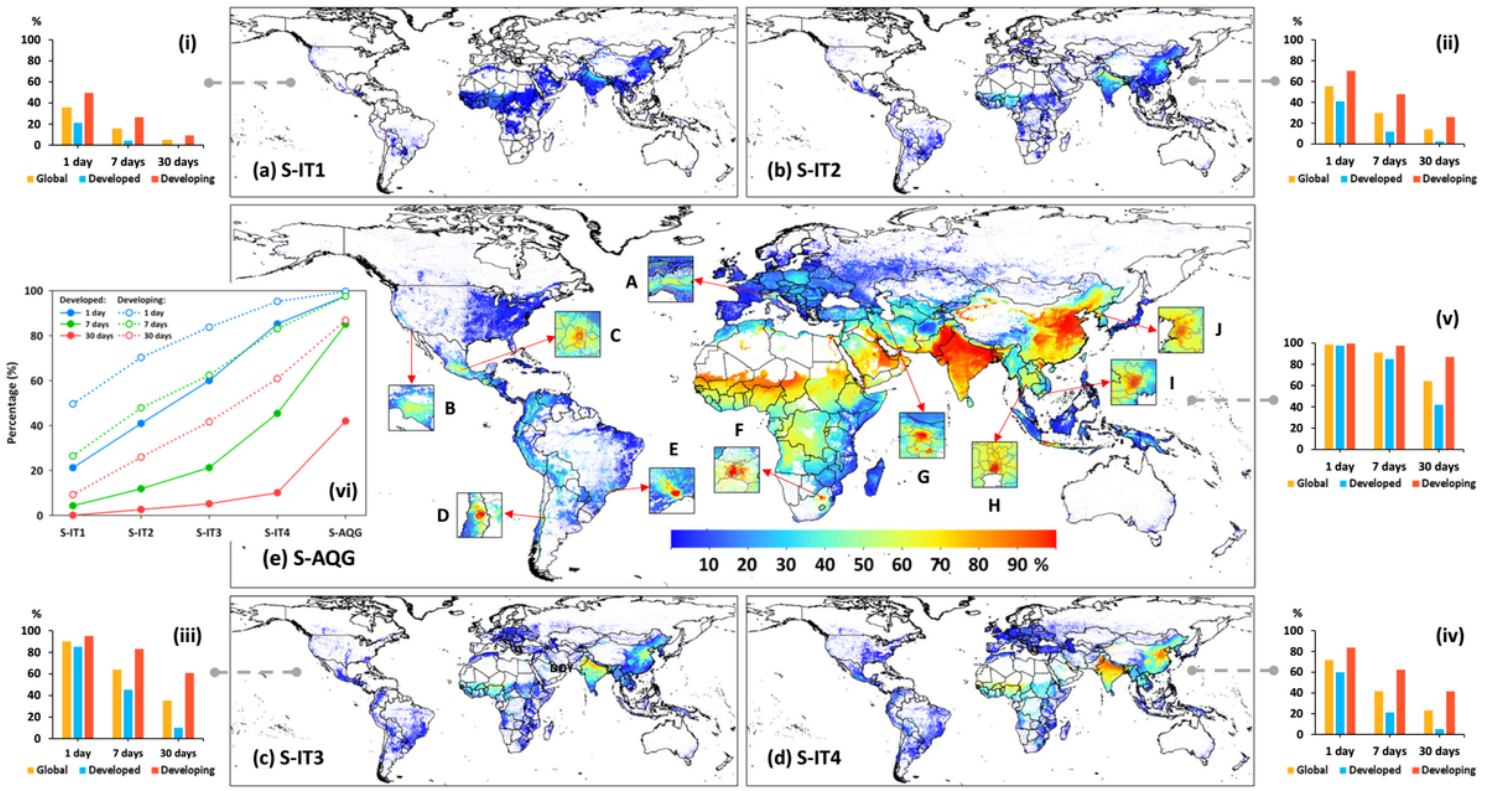


Figure 4

Spatial distribution of global daily exposure risks with the percentages (unit: %) of days exceeding the World Health Organization (WHO) recommended four short-term interim targets: (a) S-IT1 (daily $PM_{2.5}$ > 75 μg m^{-3}), (b) S-IT2 (daily $PM_{2.5}$ > 50 μg m^{-3}), (c) S-IT3 (daily $PM_{2.5}$ > 37.5 μg m^{-3}), (d) S-IT4 (daily $PM_{2.5}$ > 25 μg m^{-3}), and short-term air quality guidelines (S-AQG) level (daily $PM_{2.5}$ > 15 μg m^{-3}) at each 1-km^2 grid in areas with population density > 1 per km^2 in 2020. Rectangular insets zoom in on the hot spots with high $PM_{2.5}$ exposure risk in mega-cities. The (i-v) bar charts count the percentages (unit: %) of the populated areas that have at least 1, 7, or 30 days exceeding the WHO daily standards for the global, developed, and developing world, and the (vi) line chart plots the percentages (unit: %) of the populated areas a function of WHO daily standards for developed and developing countries (i.e., solid and dashed lines), and for exposure periods of 1 day, 7 days, and 30 days (different colors) in 2020.

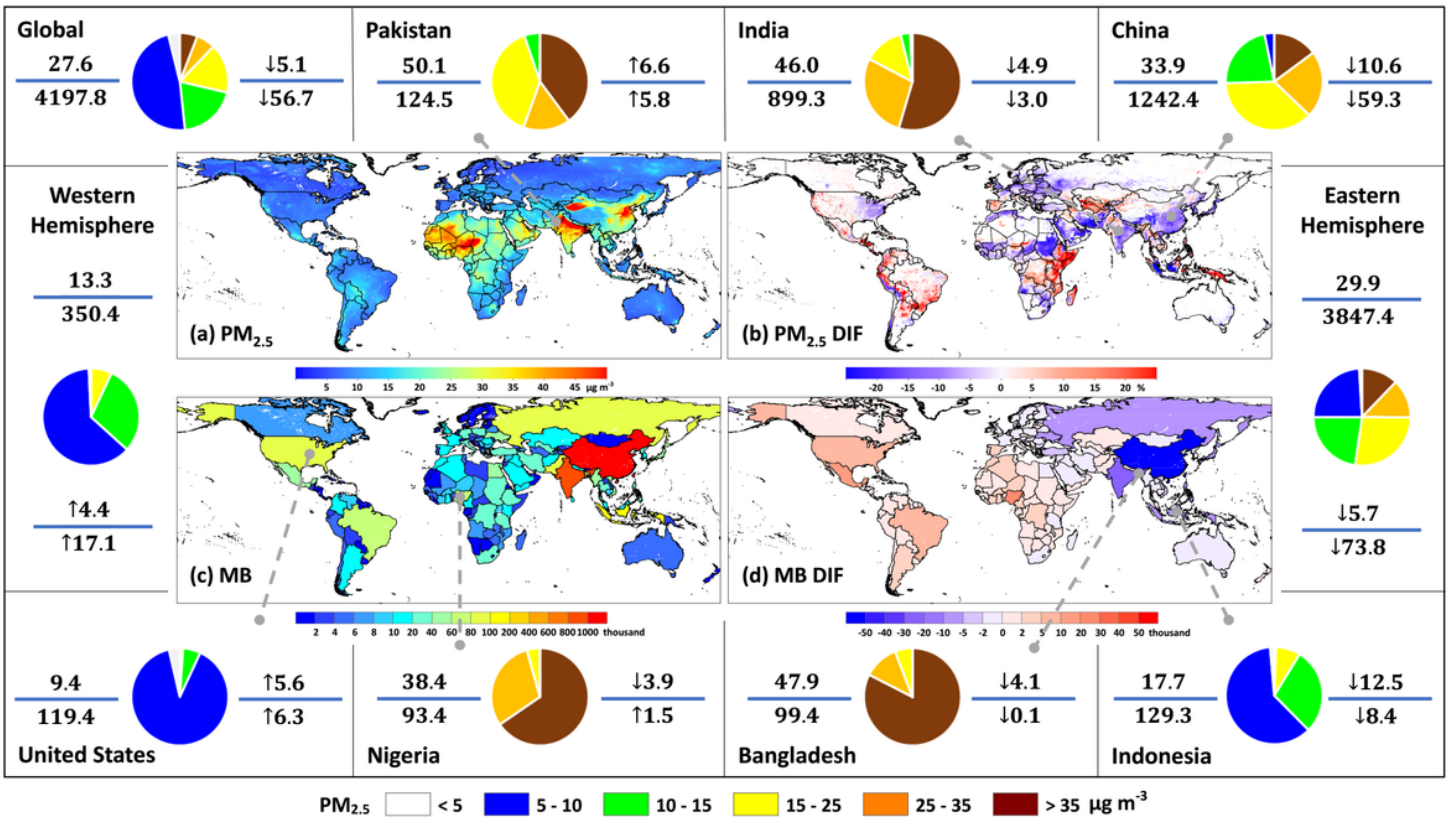


Figure 5

Spatial distributions of (a) global 1-km-resolution annual mean $PM_{2.5}$ concentration (unit: μg m^{-3}) in 2020 and (b) relative difference (unit: %) from 2019 in areas with population density > 1 per km^2; (c) global mortality burden (MB, unit: thousand) in 2020 and (d) difference (unit: thousand) from 2019 for each country. Surrounding panels show statistical results of ambient $PM_{2.5}$ levels and associated mortality burdens for the whole world, Western and Eastern Hemispheres, and the top seven countries, where the left-hand group of numbers indicates annual population-weighted $PM_{2.5}$ concentrations (top number, unit: μg m^{-3}) and total premature deaths (bottom number, unit: thousand) in 2020, and the right-hand group of numbers indicate differences in population-weighted $PM_{2.5}$ (unit: μg m^{-3}) and differences in the number of deaths (unit: thousand) from 2019. Pie charts in the middle illustrate the proportions of populated-area exposures to different $PM_{2.5}$-polluted levels according to the World Health Organization's recommended long-term air quality guidelines level and interim targets.