World Health Organization

WHO ambient air quality database

April 2022 update

Methodology

Methods

The database includes annual mean concentrations of PM_{10}, $PM_{2.5}$ and NO_2 based on ground measurements of these pollutants. It provides an average for a city or town as a whole rather than at individual stations. Most of the measurements were made between 2010 and 2019.

Data sources

The primary sources of data were official reports of countries sent to WHO upon request, official national and subnational reports and national and subnational websites that contain measurements of PM_{10} or $PM_{2.5}$ and ground measurements compiled in the framework of the Global Burden of Disease project *(18)*. For NO_2, ground measurements compiled for research by Larkin et al. in 2017 *(19)* were obtained. Measurements reported by the following regional networks were also used: Clean Air for Asia *(20)*, the Air quality e-reporting database of the European Environment Agency *(21)* for Europe and the AirNow Programme from the United States embassies and consulates *(22)*. If such official data were not available, values from peer-reviewed journals were used.

Types of data

Annual mean concentrations of particulate matter (PM_{10} or $PM_{2.5}$) and NO_2 derived from daily stationary measurements or data that could be aggregated into annual means, were used. In the absence of annual means, measurements over a more limited part of the year were used exceptionally to derivate the annual mean, if the different seasons were represented.

In order to present air quality data that represent human exposure, we used mainly urban measurements, comprising urban background, residential areas, commercial and mixed areas or rural areas and industrial areas close to urban settlements. Only data from stationary measurements, as opposed to mobile stations, were included. Air quality stations that covered particular “hot spots” and exclusively industrial areas were not included in the analysis, as such measurements often represent areas with the highest exposure and not mean population exposure. “Hot spots” were either designated as such in the original reports or were qualified as such because they were, e.g., near exceptionally busy roads. It should be noted that the omission of these measurements, might, however, have resulted in underestimates of the mean air pollution in a city.

When data from various sources were available for an urban area, only the latest, most reliable sources were used. For locations for which no new data were available, data from the previous version of the database were used.

We could not retrieve or use all the publicly available data of interest, because they were not in one of the four languages selected for the search (i.e., English, French, Portuguese and Spanish) or they provided incomplete information (such as the reference year or station coordinates). Data were used as presented in the original sources. The numbers of monitors cited do not necessarily correspond to the number of operational stations in a city but to the number of stations used to derive the mean.

Search strategy

When official reporting from countries to WHO were not available, we screened the websites of national ministries of the environment and health and statistics offices for publicly available data

The web searches were conducted with the terms “air quality”, “air pollution”, “suspended particles”, “monitoring”, “PM10”, “PM2.5” and “NO2”. The languages chosen were English, French, Portuguese and Spanish.

Only measurements up to 2019 were included, although some late searches included 2020.

Data processing and reporting

When they were available, means for cities and towns reported in the original sources were included. When a mean was not provided, data from the eligible monitoring station in the city or town were averaged. As monitoring stations may be placed in locations that do not represent the level of background pollution, aggregation of their data may not necessarily represent mean air pollution in a city. This risk was partly mitigated by excluding data from monitoring stations located in hot spots, as stated above.

The population data used for weighting and for estimating the proportion of the population covered were derived from United Nations Population Statistics *(23)* (when available) for all the human settlements covered or census data from national statistics offices *(24)*. When no population data were available, the median for a specific sector of the population was retrieved manually from the latter.

The temporal coverage represents the number of days per year covered by measurements and any other range provided in the original sources. When data from several monitoring stations in one city or town were available, the average temporal coverage was used as the overall average. Although information on temporal coverage was not always available, the reporting agencies often set a threshold for the number of days covered before reporting the measurements from a station or used it to estimate the city mean.