The Innovation, Volume *4*

Supplemental Information

Antagonism between ambient ozone increase and urbanization-oriented population migration on Chinese cardiopulmonary mortality

Haitong Zhe Sun, Junchao Zhao, Xiang Liu, Minghao Qiu, Huizhong Shen, Serge Guillas, Chiara Giorio, Zosia Staniaszek, Pei Yu, Michelle W.L. Wan, Man Mei Chim, Kim Robin van Daalen, Yilin Li, Zhenze Liu, Mingtao Xia, Shengxian Ke, Haifan Zhao, Haikun Wang, Kebin He, Huan Liu, Yuming Guo, and Alexander T. Archibald

SUPPLEMENTARY MATERIALS

Antagonism between ambient ozone increasing and urbanization-oriented population migration on Chinese cardiopulmonary mortality

Haitong Zhe Sun[1,2,3⊥], Junchao Zhao[4⊥], Xiang Liu[5], Minghao Qiu[6], Huizhong Shen[7], Serge Guillas[8,9], Chiara Giorio[1], Zosia Staniaszek[1], Pei Yu[10], Michelle W.L. Wan[1], Man Mei Chim[1], Kim Robin van Daalen[11,12,13], Yilin Li[1], Zhenze Liu[14], Mingtao Xia[15], Shengxian Ke[16], Haifan Zhao[17], Haikun Wang[5], Kebin He[4], Huan Liu[4*], Yuming Guo[10*], and Alexander T. Archibald[1,18*]

1 Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK
2 Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ, UK
3 Department of Environmental Health and Engineering, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA
4 State Key Joint Laboratory of ESPC, State Environmental Protection Key Laboratory of Sources and Control of Air Pollution Complex, School of Environment, Tsinghua University, Beijing 100084, China
5 School of Atmospheric Sciences, Nanjing University, Nanjing 210023, China
6 Department of Earth System Science, Stanford University, Stanford, CA 94305, USA
7 School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China
8 Department of Statistical Science, University College London, London WC1E 6BT, UK
9 The Alan Turing Institute, London NW1 2DB, UK
10 School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC 3004, Australia
11 British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK
12 Heart and Lung Research Institute, University of Cambridge, Cambridge CB2 0BD, UK
13 Barcelona Supercomputing Center, Department of Earth Sciences, Barcelona 08034, Spain
14 School of Environmental Science and Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China
15 Department of Mathematics, University of California, Los Angeles, California 90095, USA
16 State Key Laboratory of New Ceramics and Fine Processing, Key Laboratory of Advanced Materials of Ministry of Education, School of Materials Science and Engineering, Tsinghua University, Beijing, 100084, China
17 Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK
18 National Centre for Atmospheric Science, Cambridge CB2 1EW, UK

⊥HZS and JZ contributed equally.
*Correspondence to: Alexander T. Archibald, Yuming Guo, and Huan Liu.

56 pages with 8 methodological notes, 17 tables, 20 figures, and 2 sections of listed contents for Supplementary Materials

Updated on 27 September 2023

CONTENTS

Supplementary Methods

Supplementary Tables

Supplementary Figures

Supplementary Contents

SUPPLEMENTARY METHODS

Method S1 | Multi-model Fusion and Downscaling

The initial version of ambient O_3 concentration dataset developed by space-time Bayesian neural network downscaler (BayNNDv1) followed two major steps: i) multi-model fusion[1], and ii) urban-rural distinguished downscaling[2]. During multi-model fusion, a total of 10 CMIP6 historical simulations were selected as inputs for 1990–2014, and 6 SSP2-RCP4.5 scenario projections for 2015–2019[2]. The imbalanced model numbers between the 2 phases (Phase 1: 1990–2014, Phase 2: 2015–2019) introduced additional heterogeneities. The cross-scenario divergences were way lower than the cross-model discrepancies, and thus we replaced SSP2-RCP4.5 with SSP3-RCP7.0 to reach homogeneity with the maximum number of models between the two Phases. We fused 8 coupled earth system models with interactive chemistry as i) BCC-ESM1, ii) CESM2-WACCM, iii) EC-Earth3-AerChem, iv) GFDL-ESM4, v) GISS-E2-1, vi) MRI-ESM2-0, vii) UKESM1-0-LL, and viii) CCMI, an average of 2 earlier generation atmospheric models, CESM1-WACCM and CMAM[3-10]. All the involved CMIP6 model simulation outputs are downloaded from Earth System Grid Federation repository platform: https://esgf-node.llnl.gov/search/cmip6.

Following the established methodology with replacement of data sources and adding *in situ* observations during 2014–2019 provided by China National Environmental Monitoring Centre (CNEMC), we improved the accuracy of BayNNDv1. The optimised product BayNNDv2 is of higher global overall accuracy R^2=0.91, RMSE=4.5 ppb for urban, and R^2=0.89, RMSE=5.2 ppb for rural sites.

Method S2 | Phased Data Fusion

As the base ambient O_3 products were of different temporal coverage, time-period phased data fusion was conducted. For *Phase I* (Roman numerals were used here to avoid confusion with the aforementioned Phase 1) during 1990–2002, fusion with calibration were conducted on BayNNDv2 and M^3-BME. Due to the lack of systematic observations in China during this period, we trained the supervised deep learning model merely based on the observation archives from Tropospheric Ozone Assessment Report (TOAR) project[11], and predicted the ambient O_3 for China assisted with geographic and sociodemographic features as a compromised choice. For *Phase II* of 2003–2012, BayNNDv2, M^3-BME and CEML were blended after unification into 1/8°×1/8° spatial resolution. Still, no Chinese localised observations were involved, but satellite-based remote-sensing measurements were included to increase the reliability in capturing the spatiotemporal pattern. For *Phase III* of 2013–2019, we mixed all four base databases nested in China territory, supervised by *in situ* observations from China National Environmental Monitoring Centre (CNEMC). The urban-rural distinguishment was inherited from BayNNDv2, and data fusions were performed for urban and rural concentrations separately.

All ground-level site-based observations were aggregated into 1/8°×1/8° as supervised training labels. The fusion processes can be expressed as follows:

Phase I:
$$O_3^{urban} = f(BayNND^{urban}, M^3\text{–}BME, s_1, s_2, s_3, t_1, t_2, t_3),$$
$$O_3^{rural} = f(BayNND^{rural}, M^3\text{–}BME, s_1, s_2, s_3, t_1, t_2, t_3),$$

Phase II:
$$O_3^{urban} = f(BayNND^{urban}, M^3\text{–}BME, CEML, s_1, s_2, s_3, t_1, t_2, t_3),$$
$$O_3^{rural} = f(BayNND^{rural}, M^3\text{–}BME, CEML, s_1, s_2, s_3, t_1, t_2, t_3),$$

Phase III:
$$O_3^{urban} = f(BayNND^{urban}, M^3\text{–}BME, CEML, TAP, s_1, s_2, s_3, t_1, t_2, t_3),$$
$$O_3^{rural} = f(BayNND^{rural}, M^3\text{–}BME, CEML, TAP, s_1, s_2, s_3, t_1, t_2, t_3),$$

where f stands for the trained elastic net linear regressor, s_1, s_2, s_3 refer to the spatial geometric coordinates, and t_1, t_2, t_3 are temporal periodical and sequential indicators as listed below[12]. Cross-validation test results and overall performance evaluations were summarised in Table S6.

$$
\begin{aligned}
s_1 &= cos\left(2\pi\frac{longitude}{360}\right) cos\left(2\pi\frac{latitude}{180}\right),\\
s_2 &= cos\left(2\pi\frac{longitude}{360}\right) sin\left(2\pi\frac{latitude}{180}\right),\\
s_3 &= sin\left(2\pi\frac{longitude}{360}\right),\\
t_1 &= cos\left(2\pi\frac{month}{12}\right),\\
t_2 &= sin\left(2\pi\frac{month}{12}\right),\\
t_3 &= \frac{month}{360}.
\end{aligned}
$$

It should be furtherly noted that the BayNNDv2 urban-rural downscaled dataset was treated fully as a core basis dataset, and then 3 other well-developed datasets (M^3-BME, CEML and TAP) were fused using elastic net regressor rather than being incorporated as auxiliary predictors for Bayesian neural network downscaler. We selected such design for the purpose of maintaining the temporal homogeneity, as the elastic net regressor would "*respect*" the source dataset *closest* to the labels for supervision (i.e. observations), and regard the other two datasets as a strategy of "*belt and braces* (*double insurance*)" in case the Bayesian neural network "*missed*" any information that had been captured by M^3-BME, CEML or TAP. The elastic net regressor (instead of other base machine learners like random forest or gradient boosting decision tree) would not substantially destroy the spatiotemporal pattern of the very input dataset closet to the observations, and tune with the rest input datasets if necessary. It can effectively avoid causing "*fractures*" in the "*junction*" year of different datasets (e.g. CEML starts from 2003, and hence 2003 is a junction year that the temporal fracture will be inclined to occur). When calculating the importance features of *Phase III* (2013–2019), the core dataset BayNNDv2 occupied 96.8% and 94.1% weights for urban and rural predictions, respectively, justifying the necessity and credibility of long-term global-scale space-time integrated training.

Method S3 | Detailed specification of Chinese administrative divisions

We used 7-division scheme in this study. This scheme of Chinese Administrative Geographical Division considers geography, history, culture, and ethnicity into comprehensively. The municipalities directly under Chinese Central Government and Autonomous Regions are all of provincial executive level. *Northeast China* includes 3 provinces: Heilongjiang, Jilin, and Liaoning. *North China* includes 3 provinces: Hebei, Shanxi, Inner Mongolia Autonomous Region; and 2 direct-administered municipalities: Beijing and Tianjin. *East China* includes 7 provinces: Shandong, Jiangsu, Anhui, Zhejiang, Jiangxi, Fujian, and Taiwan; and a direct-administered municipality: Shanghai. *Central China* includes 3 provinces: Henan, Hubei, and Hunan. *South China* includes 3 provinces: Guangxi Zhuang Autonomous Region, Guangdong, and Hainan; and 2 Special Administrative Regions (SAR): Hong Kong SAR and Macao SAR. *Southwest China* includes 4 provinces: Tibet Autonomous Region, Yunnan, Sichuan, and Guizhou; and a direct-administered municipality: Chongqing. *Northwest China* includes 5 provinces: Xinjiang Uygur Autonomous Region, Qinghai, Gansu, Ningxia Hui Autonomous Region, and Shaanxi. *Jing-Jin-Ji* (JJJ) urban agglomeration consists of Beijing, Tianjin, 11 prefecture-level cities (Shijiazhuang, Baoding, Tangshan, Langfang, Qinhuangdao, Zhangjiakou, Chengde, Cangzhou, Hengshui, Xingtai, Handan) in Hebei Province, and Anyang in Henan Province. "Ji" ("冀", pronounced as jì) is the ancient name of Hebei Province. Some schools abbreviate the megalopolis as BTH (Beijing, Tianjin, and Hebei). *Cheng-Yu* (CY) urban agglomeration consists of Sichuan Province (excluding Liangshan, Panzhihua, Aba, Ganzi, Guangyuan, Bazhong) and Chongqing (excluding Qianjiang, Pengshui, Youyang, Xiushan, Chengkou, Wushan, Wuxi, Fengjie). The alternative historical name of Chongqing is "Yu" ("渝", pronounced as yú), and hence for the phonological harmony, Chengdu-Chongqing district is more commonly shorted as Cheng-Yu rather than Cheng-Chong. *Yangtze River Delta* (YRD) urban agglomeration consists of Jiangsu Province, Anhui Province, Zhejiang Province, and Shanghai. The China *Greater Bay Area* (GBA) circumscribes Hong Kong SAR, Macao SAR, and 9 prefecture-level cities in Guangdong Province (Guangzhou, Shenzhen, Foshan, Dongguan, Zhongshan, Jiangmen, Huizhou, Zhuhai, Zhaoqing), which is alternatively entitled as the Guangdong–Hong Kong–Macao Greater Bay Area. The 9 cities in Guangdong Province are collectively named as Pearl River Delta (PRD) Economic Zone. The 7 Chinese administrative divisions and 4 megalopolises were mapped in Figure S1.

Method S4 | Identification of mortality causes

Meta-analyses were performed on the extracted cohort-based epidemiological evidences (e.g. hazard ratio, HR) relevant to long-term O_3 exposure from systematic review updated until October 2022. All reported mortality causes were included for meta-analysis extended from the latest relevant systematic reviews[13,14], and the causes with pooled positive relative risks were considered for mortality estimation. Applying the Hunter-Schmidt estimator, 6 mortality causes (might not be mutually exclusive due to hierarchically overlapping) were identified to be of positive relative risks: non-accidental causes, chronic respiratory diseases, chronic obstructive pulmonary disease, cardiovascular diseases, ischaemic heart diseases, and congestive heart failure, as plotted in Supplementary Figures 3–7, and potential publication biases were tested by trim-and-fill method (Figure S8).

In terms of mortality estimation, the non-accidental cause mortalities were narrowed to mortalities of non-communicable diseases (NCDs), as it is reasonable to assume the non-accidental causes other than NCDs (e.g. communicable, maternal, neonatal, and nutritional diseases, injuries, suicide and homicide, etc.) are of no association with ambient O_3 exposure. In addition, mortality estimations in this study did not include the congestive heart failure which was not listed as an individual mortality cause in the GBD 2019 Study[15]. Therefore, further explorations on the nonlinear exposure-response relationships and excess mortality estimations only involve i) NCDs, ii) chronic respiratory diseases (CRDs), iii) chronic obstructive pulmonary disease (COPD), iv) cardiovascular diseases (CVDs), and v) ischaemic heart disease (IHD).

Method S5 | Construction of exposure-response curve

As the exposure-response association strengths may not necessarily follow the linear pattern, curved trends were explored by meta-regression enhanced by exposure-range resampling for the sake of more accurate risk estimation[14,16,17]. Most of the pre-existing studies were conducted on the North American and European countries where ambient O_3 pollution has been effectively constrained in the past decades, and thus the averaged exposure levels of the cohort participants were lower than the Chinese population. Under this circumstance, multi-cause mortality relative risks for Chinese residents estimated by conventional meta-regression method would rely on exposure extrapolation, leading to high uncertainties. To address this issue, exposure-range resampling would make full use of the literature-reported population exposure levels rather than the study-specific averaged exposure concentrations, so as to cover the exposure range as wide as possible and thus increase the estimation robustness.

The concentration-response curves were adopted in priority if reported in the literature. We queried the authors of the published studies providing the non-linear concentration-response relationships for the detailed values of the curves; and as for the studies we did not receive responses by October 2022, we recovered the values directly from the figure by mean of geometric measurement in Microsoft Visio. If the original studies did not explore the concentration-response trends, linear relative risk models were assumed across the reported exposure range, with the theoretical minimum risk exposure level (TMREL) presumed to be a random value uniformly distributed between the minimum and lowest 5th percentile following a previous study[18]. The statistical approach to reproduce the lowest 5th percentile was described in a prior systematic review[14], and the resampled/imputed distribution statistics were listed in Table S15. The estimated concentration-response curves for mortality risks of NCDs, chronic respiratory diseases, COPD, cardiovascular diseases, and ischaemic heart disease were presented in Figure S9.

The exposure range resampling reproduced the exposure level (OSDMA8 in ppb) by every 1 ppb increment between the literature-reported minimum and maximum exposure level as $\boldsymbol{x}$. In linear-model presumed relative risk recovering, for each resampled exposure concentration $\boldsymbol{x}$, the corresponding effective exposure "dose" $\Delta\boldsymbol{x}$ is defined as

$$\Delta\boldsymbol{x} = ReLU\{\boldsymbol{x} - \boldsymbol{TMREL}\},$$

where ReLU is the rectified linear unit choosing the greater value between 0 and $\boldsymbol{x} - \boldsymbol{TMREL}$. Given the reported risk association (i.e. HR) with 95% confidence interval (CI) as HR_{LB} to HR_{UB} by every Δy incremental exposure, the relative risk with 95% CI at exposure concentration $\boldsymbol{x}$ can be calculated as

$$\boldsymbol{HR}_{\boldsymbol{x}} = e^{lnHR\cdot\Delta x/\Delta y};$$

$$\boldsymbol{HR}_{LB,\boldsymbol{x}} = e^{lnHR_{LB}\cdot\Delta x/\Delta y};$$

$$\boldsymbol{HR}_{UB,\boldsymbol{x}} = e^{lnHR_{UB}\cdot\Delta x/\Delta y}.$$

Following the procedures illustrated above leads to an exposure-response sequence for each study that did not report the concentration-response curve; the fully resampled sequences undergo MR-BRT multi-study pooling with the literature-reported exposure-response curves.

Several previous studies have provided estimations on O_3-associated excess COPD mortality. Taking 2017 as an example, the GBD report estimated the COPD mortality as 113 (95% Uncertainty Interval, UI: 53–178) thousand[17], which is lower than our results (183, 95% UI: 125–245 thousand), as GBD applied undersized RR values[19]. Yin et al. reported 178 (95% UI: 68–286) thousand COPD deaths attributable to O_3 exposure[20], which is more coherent with our result in terms of central estimate. This is because the RR value they used (RR=1.040, 95% CI: 1.013–1.067) from a single cohort study[21] is close to the multi-study pooled RR by our meta-analysis (RR=1.056, 95% CI: 1.029–1.084); but their result is still of greater estimation uncertainty. Contrarily, Malley et al. used oversized risk association strength (RR=1.12, 95% CI: 1.08–1.16) and reported 316 (95% UI: 230–403) thousand respiratory deaths for 2010[22], which is substantially higher than our estimates (179, 95% UI: 122–241 thousand). The unneglectable cross-study divergences and great estimation uncertainties reveal the insufficiency of epidemiological evidences. Furthermore, leaving out cardiovascular mortality risks leads to dubious conservative overall estimations. We consider cardiopulmonary beyond respiratory mortality for the first time and thus provide an *aggressive* estimation to update the literature.

Method S6 | Construction procedures of gridded population dataset

The step-by-step procedures to construct the gridded Chinese population dataset are illustrated in the flowchart (Figure S17), in which the rounded rectangles indicate procedural semi-manufactures, rectangles refer to the initial input and final output datasets, and the number-marked arrows represent operations.

Starting point: UN WPP-adjusted GPWv4. The Gridded Population of the World with adjustment from United Nation World Population Prospects[23] (version 4.11) was set as the footstone, as it is the latest global population distribution product with the finest spatial resolution (30''×30'') and densest temporal coverage (2000–2019).

Step 1: Spatial re-gridding. The spatial resolution of finally enhanced ambient O_3 concentration dataset with urban-rural distinguishment is 1/8°×1/8°, based on which the population exposure levels were assessed. By averaging the 15×15 adjacent grids (1/8°=30"×15), the raw 30"×30" dataset was re-gridded into 1/8°×1/8°.

Step 2: Temporal extrapolation. The GPWv4 dataset covers 20 studied years: 2000–2019. For each re-gridded 1/8°×1/8° cell, the restricted cubic spline regression model with 3 knots was applied to the cell-level population against year, so as to extrapolate the temporal coverage onto the complete study years: 1990–2019, following previous studies[1,24].

Step 3: China localisation. The global longitude-latitude-based grids were geographically projected onto the map of China provided by the Ministry of Natural Resources of People's Republic of China, and all grids belonging to China territory were extracted for further processing. Geographical mapping and administrative division projection (i.e. country, provinces, and prefecture-level cities) were performed in QGIS (version 3.26.10).

Step 4: Urban-rural distinguishment. The Population Dynamics dataset (version 1.01), identifying urban and rural population counts for each 1/8°×1/8° grid[25], was extrapolated onto 30 consecutive years by mean of restricted cubic spline model[2], based on which the urban and rural population fractions were calculated. The cell-specific fractions were then multiplied onto the 30-year extrapolated GPWv4 China gridded population dataset (i.e. procedural semi-manufactures of Step 3), to update the urban-rural distinguished population distribution. The consensus has been widely accepted that GPWv4 datasets reporting 20 consecutive years were more reliable than interpolated data products.

Step 5: Urban-rural calibration. The China Statistical Yearbook series reported the numbers of urban and rural residents for each year, with which the estimated values were linearly aligned. For an instance, if the predicted total count of urban residents of a certain province (Step 4) was Pop_{pred} while the factual count provided by the China Statistical Yearbook was Pop_{stat}, the urban population count *for each grid* was then multiplied by a coefficient of Pop_{pred}/Pop_{stat} . Province-level calibrations were performed for 2005–2019 in accordance with the Yearbook precision, while nation-level calibrations were conducted for 1990–2004 as a compromise given the data unavailability.

Step 6: Age group specification. Fractions of population aged above 25 were retrieved from GBD Population Estimates 1950–2019,[26] and the China Statistical Yearbook series 2004–2019[27]. Values provided by the China Statistical Yearbook series were adopted in priority for 2004–2019, while for the earlier years 1990–2003 when the China Statistical Yearbook did not archive the population pyramid, the GBD Population Estimates were used as a compromise. The nation-level year-specific fractions were multiplied onto each grid to identify the counts of population age ≥25. After this step, the enhanced gridded population age ≥25 differentiated with urban and rural residence was used as the capstone dataset for main analysis.

Step 7: Gender group specification. Genders were furtherly specified for sensitivity analysis. Province-level gender proportions for 2000–2019 and nation-level gender proportions for 1990–1999 were obtained from the China Statistical Yearbook series 1990–2019[27]. The province- or nation-level male and female proportions were uniformly applied onto each grid circumscribed inside the corresponding province or the whole country territory, respectively.

Step 8: Dataset encapsulation. After all the aforementioned data processing, the gridded population was structured into the meta-dataset: 1/8°×1/8° spatial resolution; yearly resolved spanning 1990–2019; each grid encapsulating 4 population counts as: i) urban male age ≥25, ii) urban female age ≥25, iii) rural male age ≥25, and iv) rural female age ≥25.

Method S7 | Sensitivity analyses

Long-term ambient O_3 tracking covers earlier years beyond the satellite-based remote sensing measurements or chemical reanalysis (i.e. 1990–2002), indicating predictions would merely relied on the CMIP6 numerical simulations for this period.

We therefore extended a sensitivity analysis for the first-stage space-time Bayesian neural network-based data assimilation during 2003–2019 under two scenarios, as fusing eight CMIP6 models with (ScA) and without (ScB) a machine-learning-calibrated remote-sensing measurements and chemical reanalysis outputs[28], assisted with over 40 auxiliary features[2].

We then evaluated the accuracies of 10-fold cross-validation tests by random split (70% dataset matched with observations), external validation tests (the rest 30%), and overall fitting, as summarized in Table S16. We compared the developed ambient O_3 datasets under the two scenarios by coefficient of variation (CoV): standard deviation divided by the arithmetic mean. We concluded that the deep-learning-based prediction accuracies by solely using CMIP6 simulations were as competitive as fusing additional measurements, and no substantial discrepancies were observed between ScA and ScB (CoV=1.0%, spatiotemporal 5–95th%ile: 0.1–2.8%).

We furtherly split the full dataset manually for cross-validation tests under ScB, maintaining the temporal coherence: i) 2003–2012 for training and 2013–2019 for testing; ii) 2003–2007 and 2015–2019 for training and 2008–2014 for testing; and iii) 2010–2019 for training and 2003–2009 for testing. All three temporally staged cross-validation tests had revealed good performances (R^2=0.90, 0.92, 0.92; RMSE=2.86, 2.71, 2.70 ppb, respectively for the three tests). The constrained cross-scenario divergences and stable temporal generalizability verified the credibility of model-based ambient O_3 tracking in the earlier years.

Parallel with the curved risk model, the *linear risk model* was adopted for attributable mortality estimation as reference, which assumed that relative risks change linearly with the exposure level $\boldsymbol{x}$ following

$$RR_x = e^{lnRR\cdot\frac{\Delta x}{\Delta y}}$$

where RR is the multi-study pooled value scaled in each Δy incremental exposure, and Δx is the effective dose above the TMREL.

We performed a series of further sensitivity analyses on the estimation for 2017 as an example. The exposure-response relationships might be the major source of estimation uncertainty, and thus we applied the multi-study pooled RRs onto the simplest log-linear model parallel to the curved model as presented in the main results. The threshold (also known as TMREL or low-concentration cut-off) for long-term O_3 exposure-associated mortality risk was also contentious, and thence we tested several values as directed in literature: i) the global lowest 5th percentile PWE in 2017 by BayNND, 42.6 ppb (Scenario 1, Sc1); ii) the 30-year global lowest 5th percentile PWE by BayNND, 40.8 ppb (Sc2); and iii) the maximum of literature-reported lowest 5th percentile exposure levels from studies included for meta-analysis, 44.0 ppb (Sc3). We used the grid-averaged ambient O_3 concentrations to quantify population exposure, supposing the ambient O_3 exposure levels were not distinguished for urban and rural environments, as Sc4. Gender-specified mortality metrics other than the gender-standardized estimates reported by IHME, were used as Sc5. Province-specific mortality metrics for 2017 provided by China CDC were applied as Sc6[29]. In Sc7, we replaced the O_3 tracking dataset with M^3-BME solely, which was used in the GBD 2019 study. In Sc8, we adopted cardiovascular mortality risk association (RR=1.227, 95% CI: 1.108–1.359, *p-value*=0.79) pooled from 2 cohort studies on Chinese population reporting higher RRs[30,31].

Estimations for excess deaths differentiating the designed schemes were summarized in Table S17. The cross-scheme discrepancies were constrained not to exceed 10%, and therefore sensitivity analyses validated the robustness of our mortality estimations, verified the coherence of the data sources, and justified the rationality of innovations in our study design.

Method S8 | Cross-validation for spatiotemporal generalizability

Since China lacked systematic ground-level measurements in earlier years before 2013, and the observation sites deployed in urban and rural environments were disproportional. We therefore decided to train the model at global scale with sufficient supervision by observations, and conducted strengthened rigorous cross-validation tests on the spatiotemporal extrapolation reliability to verify the generalizability of the deep learning downscaling algorithm. Besides the cross-validation and external validation tests by random split, we extended region-clustered cross-validation tests on spatial extrapolation capability (cvs_1: training on North America, testing on Europe; cvs_2: training on Europe, testing on North America; cvs_3: training on North America and Europe, testing on Asia; and cvs_4: training on locations outside China, testing on China), and staged cross-validation tests on global-scale temporal generalization (cvt_1: training on 1990–2013, testing on 2014–2019; cvt_2: training on 1990–2007 and 2014–2019, testing on 2008–2013; cvt_3: training on 1990–2001 and 2008–2019, testing on 2002–2007; cvt_4: training on 1990–1995 and 2002–2019, testing on 1996–2001; cvt_5: training on 1996–2019, testing on 1990–1995) for the second-stage urban-rural differentiated downscaling. Spatiotemporal generalizability tests are summarized in Table S7.

Supplementary Tables

Table S1 | Province-level average of ambient ozone concentrations in 1990 and 2019.

Urban, rural and population-weighted exposure (PWE) concentrations are scaled as 6-month (April to September) ozone-season daily 8-hour maximum average (OSDMA8) in ppb for either year. Statistics include the regional median and spatial 5-95th percentile range. Hong Kong SAR and Macao SAR have realised full urbanisation before 1990, and thus rural concentrations are not considered.

Region	Year 1990 Urban		Year 1990 Rural		Year 1990 PWE		Year 2019 Urban		Year 2019 Rural		Year 2019 PWE	
Nationwide	40.2	(20.7–48.7)	54.2	(44.2–62.8)	49.0	(39.1–57.2)	59.5	(46.1–91.9)	67.9	(56.0–93.2)	63.3	(52.4–87.3)
Northeast China	34.6	(31.4–43.4)	47.6	(40.6–58.3)	44.1	(36.8–53.8)	49.0	(40.2–74.5)	59.7	(47.0–78.5)	55.6	(43.8–69.9)
Heilongjiang	32.5	(30.3–36.7)	46.9	(39.3–48.1)	42.3	(36.4–44.5)	39.4	(34.0–48.5)	49.6	(43.9–56.5)	43.4	(37.9–51.6)
Jilin	36.3	(31.8–42.1)	48.7	(46.9–58.3)	44.9	(42.1–53.2)	42.9	(41.6–59.7)	56.4	(50.5–67.1)	48.5	(45.3–62.8)
Liaoning	40.7	(39.2–44.4)	56.5	(54.4–58.6)	51.0	(49.1–53.7)	61.8	(54.1–76.5)	67.1	(63.4–81.0)	63.5	(57.0–77.9)
North China	38.2	(30.9–45.5)	51.1	(45.5–59.3)	46.5	(41.6–53.4)	58.3	(42.1–93.2)	65.7	(54.3–95.7)	61.4	(50.0–87.1)
Inner Mongolia	37.9	(30.5–40.6)	49.3	(45.5–53.6)	45.9	(41.1–49.8)	54.7	(41.8–86.6)	59.8	(52.7–87.5)	56.6	(45.8–86.9)
Beijing	44.6	(40.0–45.3)	56.2	(54.0–58.1)	50.4	(47.0–51.7)	96.1	(84.7–96.1)	96.5	(89.8–96.5)	96.2	(85.4–96.2)
Tianjin	45.3	(45.1–45.3)	58.1	(57.6–58.1)	52.4	(52.0–52.4)	87.2	(87.2–90.2)	90.5	(90.5–92.0)	87.8	(87.8–90.5)
Hebei	44.4	(40.0–45.5)	56.2	(54.0–59.8)	53.4	(50.7–56.4)	89.8	(87.1–99.0)	91.8	(89.2–96.5)	90.7	(88.0–98.0)
Shanxi	38.4	(35.5–52.2)	57.5	(52.0–66.7)	52.5	(47.8–63.0)	90.3	(80.2–92.0)	91.6	(87.9–96.3)	90.8	(83.3–93.7)
East China	37.1	(16.3–44.9)	52.2	(37.9–56.5)	46.2	(36.0–54.6)	65.1	(49.7–90.8)	71.3	(62.2–96.4)	67.9	(55.6–91.8)
Shandong	43.8	(39.9–47.4)	56.1	(54.6–60.2)	52.7	(50.5–56.7)	79.2	(75.3–96.5)	88.3	(85.2–101.0)	82.7	(79.1–98.3)
Jiangsu	38.4	(37.7–44.9)	53.4	(50.9–56.5)	48.7	(46.8–52.9)	75.1	(63.0–85.7)	82.5	(70.7–91.6)	77.3	(65.3–87.5)
Shanghai	38.0	(38.0–38.0)	50.9	(50.9–50.9)	44.2	(44.2–44.2)	63.0	(63.0–63.0)	70.7	(70.7–70.7)	63.9	(63.9–63.9)
Anhui	38.4	(29.6–41.3)	53.2	(47.0–56.5)	49.9	(43.1–53.1)	78.2	(50.8–83.0)	86.6	(68.4–91.6)	81.9	(58.6–86.8)
Jiangxi	29.4	(20.3–31.2)	43.9	(37.9–46.5)	39.6	(32.8–41.7)	51.7	(49.7–65.1)	65.1	(62.2–71.3)	57.4	(55.0–67.7)
Zhejiang	36.6	(21.4–40.5)	50.9	(39.2–56.5)	46.1	(33.2–51.1)	61.3	(50.9–85.7)	67.5	(63.1–91.6)	63.2	(54.6–87.5)
Fujian	21.4	(20.1–38.6)	44.5	(43.5–57.2)	37.9	(36.8–51.9)	49.7	(48.1–56.4)	63.1	(61.1–68.1)	54.2	(52.5–60.3)
Taiwan	37.4	(37.4–37.5)	54.7	(50.7–54.7)	44.4	(42.8–44.5)	56.6	(53.1–56.6)	68.8	(65.3–68.8)	59.3	(55.7–59.3)
Central China	40.0	(26.1–69.2)	52.7	(44.8–70.0)	48.1	(38.9–56.6)	61.5	(49.8–86.4)	67.5	(60.5–87.6)	64.5	(54.0–83.9)
Henan	51.1	(42.4–57.5)	60.6	(54.2–70.0)	58.7	(51.8–67.5)	76.8	(62.6–82.6)	83.0	(66.8–87.3)	79.7	(64.5–84.8)
Hubei	47.6	(27.0–55.2)	53.5	(43.9–60.0)	52.0	(39.5–58.7)	62.6	(51.8–86.4)	68.9	(62.9–87.6)	65.0	(56.1–86.8)
Hunan	36.6	(25.9–40.0)	48.4	(42.1–52.7)	45.6	(38.2–49.7)	50.1	(49.6–61.2)	62.3	(58.0–67.5)	55.3	(53.1–63.9)
South China	32.3	(18.2–55.8)	47.3	(43.2–56.9)	41.3	(25.1–50.5)	57.5	(52.5–66.5)	63.1	(59.3–69.8)	60.2	(51.6–66.7)
Guangxi	32.2	(26.9–55.8)	46.8	(44.9–56.9)	43.7	(41.1–56.7)	57.5	(53.7–60.5)	63.1	(59.3–69.4)	60.3	(56.4–64.9)
Guangdong	32.3	(18.2–56.0)	49.6	(43.2–59.6)	43.2	(33.9–58.3)	60.7	(52.5–66.5)	66.9	(61.9–69.8)	62.5	(55.2–67.4)
Hainan	35.0	(35.0–35.5)	54.8	(53.7–54.8)	49.3	(48.5–49.4)	51.4	(51.4–58.1)	60.3	(60.3–63.0)	55.0	(55.0–60.1)
Hong Kong	33.6	(33.2–33.9)	-		33.6	(33.2–33.9)	52.5	(52.1–53.2)	-		52.5	(52.1–53.2)
Macao	32.3	(32.3–32.3)	-		32.3	(32.3–32.3)	66.5	(66.5–66.5)	-		66.5	(66.5–66.5)
Northwest China	38.4	(32.9–46.0)	50.9	(42.6–58.6)	48.9	(39.2–56.5)	51.1	(42.0–62.1)	59.8	(51.8–69.4)	57.3	(47.4–67.6)
Xinjiang	38.6	(32.9–46.4)	50.9	(42.6–60.8)	48.1	(40.4–57.5)	48.8	(41.7–61.1)	58.1	(50.1–71.7)	53.3	(45.7–66.2)
Qinghai	40.2	(37.0–42.9)	50.0	(47.2–54.2)	47.7	(44.8–51.5)	56.2	(46.4–62.4)	59.7	(52.9–67.5)	57.7	(49.3–64.7)
Gansu	37.4	(30.4–40.4)	50.3	(44.3–53.8)	47.9	(41.7–51.3)	52.8	(46.4–56.6)	63.8	(51.8–69.4)	58.5	(49.2–63.2)
Ningxia	37.5	(30.4–38.0)	51.2	(44.3–51.5)	47.6	(40.7–47.8)	51.7	(47.5–56.5)	64.2	(59.7–65.7)	56.7	(52.4–60.2)
Shaanxi	33.5	(21.1–38.9)	51.2	(42.3–53.9)	47.0	(37.3–50.3)	50.4	(47.5–82.5)	61.8	(59.7–84.2)	55.0	(52.4–83.2)
Southwest China	36.7	(18.8–41.9)	50.3	(40.2–54.5)	44.9	(33.8–50.0)	56.0	(47.1–64.9)	64.1	(59.0–68.6)	58.9	(51.8–64.3)
Tibet	38.9	(35.8–43.4)	51.6	(47.3–57.5)	49.5	(45.3–55.1)	62.2	(49.3–67.4)	63.8	(59.0–68.9)	63.3	(55.9–68.4)
Sichuan	32.1	(12.0–38.0)	49.4	(35.3–53.2)	45.8	(30.5–50.0)	53.9	(43.7–58.3)	64.1	(59.2–67.2)	58.6	(50.9–62.4)
Chongqing	22.6	(12.0–26.8)	42.3	(35.3–47.4)	36.7	(28.7–41.6)	52.7	(50.8–56.5)	64.7	(61.9–66.9)	56.7	(54.5–59.9)
Guizhou	24.4	(14.6–29.9)	43.1	(38.4–48.4)	40.0	(34.5–45.3)	51.4	(48.8–56.0)	62.0	(57.7–67.6)	56.8	(53.3–61.9)
Yunnan	32.1	(22.6–35.7)	47.7	(43.2–51.3)	44.8	(39.4–48.4)	52.2	(47.9–61.3)	64.0	(59.0–66.7)	58.2	(53.6–64.1)

Table S2 | Regional and nationwide 1990 mortality metrics associated with ozone exposure.

Excess cardiopulmonary mortalities are defined as the total deaths caused from COPD and all-type cardiovascular diseases. Three mortality metrics are considered as i) number of excess deaths in thousand, ii) mortality rate per 100 000, and iii) years of life lost (YLLs) in million years. Estimates are summarised by median with 95% uncertainty intervals from 1000-time Monte Carlo bootstrap.

Region	Excess Deaths (thousand)			Mortality Rates (per 100 000)			YLLs (million years)		
	Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
Northeast China	4.0 (2.5 to 5.4)	17.5 (11.4 to 24.1)	21.5 (13.9 to 29.6)	17.9 (11.5 to 24.8)	28.5 (18.4 to 39.2)	26.2 (16.9 to 36.1)	0.44 (0.28 to 0.61)	0.61 (0.39 to 0.84)	1.05 (0.66 to 1.44)
North China	14.3 (9.3 to 19.9)	29.0 (18.8 to 39.8)	43.3 (28.1 to 59.6)	23.1 (14.9 to 32.0)	34.0 (22.0 to 46.7)	31.0 (20.0 to 42.6)	0.57 (0.36 to 0.79)	0.73 (0.47 to 1.01)	1.29 (0.81 to 1.77)
East China	41.0 (26.3 to 56.5)	67.1 (43.4 to 91.9)	107.8 (69.6 to 148.2)	22.5 (14.5 to 31.0)	34.2 (22.1 to 46.9)	30.3 (19.6 to 41.7)	0.56 (0.34 to 0.75)	0.74 (0.47 to 1.01)	1.29 (0.81 to 1.76)
Central China	18.6 (12.1 to 25.8)	37.5 (24.3 to 51.7)	56.1 (36.3 to 77.4)	22.0 (14.2 to 30.3)	29.6 (19.1 to 40.8)	27.6 (17.8 to 38.0)	0.54 (0.34 to 0.74)	0.64 (0.41 to 0.88)	1.17 (0.73 to 1.61)
South China	1.5 (0.8 to 2.0)	13.9 (9.0 to 19.4)	15.5 (9.9 to 21.5)	2.3 (1.5 to 3.2)	21.3 (13.7 to 29.5)	15.1 (9.7 to 20.9)	0.05 (0.03 to 0.08)	0.47 (0.29 to 0.64)	0.54 (0.34 to 0.74)
Northwest China	2.6 (1.7 to 3.6)	18.6 (11.9 to 25.5)	21.2 (13.6 to 29.2)	14.5 (9.3 to 20.1)	29.5 (19.1 to 40.7)	27.6 (17.8 to 38.0)	0.35 (0.21 to 0.49)	0.64 (0.41 to 0.88)	0.99 (0.63 to 1.38)
Southwest China	2.3 (1.5 to 3.1)	31.6 (20.3 to 43.6)	34.0 (21.9 to 46.9)	4.4 (2.8 to 6.1)	23.5 (15.2 to 32.5)	20.2 (13.0 to 28.0)	0.11 (0.06 to 0.15)	0.51 (0.32 to 0.69)	0.62 (0.40 to 0.86)
Nationwide	84.2 (54.3 to 116.3)	215.3 (139.1 to 296.1)	299.5 (193.3 to 412.4)	17.5 (11.3 to 24.2)	29.4 (19.0 to 40.4)	26.3 (17.0 to 36.2)	2.62 (1.62 to 3.61)	4.34 (2.75 to 5.95)	6.95 (4.37 to 9.56)

Table S3 | Historical 30-year regional and nationwide ozone-associated mortality trends.

Longitudinal trends scaled in decadal average change rates are calculated by log-linear meta-regression maximum likelihood estimator from the annually resolved values with 95% confident intervals (CIs). When estimated trend approaches 0, an additional decimal place is reserved.

Region	Excess Deaths (thousand dec^{-1})			Mortality Rates (per 100 000 dec^{-1})			YLLs (million years dec^{-1})		
	Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
Northeast China	1.9 (1.3 to 2.8)	-0.07 (-0.09 to -0.05)	1.8 (1.2 to 2.7)	0.6 (0.4 to 0.7)	-0.2 (-0.4 to -0.1)	-0.7 (-1.0 to -0.7)	-0.019 (-0.028 to -0.011)	-0.041 (-0.054 to -0.027)	-0.060 (-0.077 to -0.036)
North China	7.7 (5.0 to 11.9)	-1.5 (-2.2 to -1.0)	6.2 (4.0 to 9.7)	2.6 (1.7 to 3.4)	1.8 (1.1 to 2.3)	1.0 (0.6 to 1.2)	0.012 (0.009 to 0.015)	-0.010 (-0.018 to -0.004)	0.002 (-0.001 to 0.005)
East China	18.1 (11.1 to 29.5)	-8.8 (-12.1 to -6.4)	9.3 (4.7 to 17.4)	1.2 (0.8 to 1.5)	0.3 (0.2 to 0.4)	-0.6 (-1.0 to -0.3)	-0.016 (-0.019 to -0.013)	-0.039 (-0.053 to -0.027)	-0.055 (-0.068 to -0.044)
Central China	9.9 (6.4 to 15.3)	-1.8 (-2.5 to -1.3)	8.1 (5.1 to 12.7)	1.9 (1.3 to 2.5)	1.6 (1.0 to 2.1)	0.9 (0.5 to 1.1)	0.001 (-0.001 to 0.003)	-0.009 (-0.014 to -0.005)	-0.008 (-0.012 to -0.005)
South China	4.4 (2.8 to 6.9)	0.17 (0.11 to 0.26)	4.6 (2.9 to 7.3)	4.2 (2.7 to 5.9)	1.2 (0.8 to 1.6)	1.2 (0.8 to 1.6)	0.075 (0.050 to 0.109)	-0.006 (-0.009 to -0.004)	0.069 (0.044 to 0.096)
Northwest China	0.9 (0.6 to 1.4)	-0.8 (-1.1 to -0.6)	0.08 (0.01 to 0.28)	0.2 (0.1 to 0.3)	-0.4 (-0.6 to -0.3)	-1.3 (-1.8 to -0.8)	-0.018 (-0.028 to -0.008)	-0.046 (-0.064 to -0.032)	-0.063 (-0.092 to -0.041)
Southwest China	4.1 (2.7 to 6.3)	-1.2 (-1.7 to -0.8)	2.9 (1.9 to 4.6)	3.5 (2.2 to 4.8)	1.6 (1.0 to 2.1)	0.5 (0.3 to 0.7)	0.062 (0.040 to 0.086)	-0.008 (-0.013 to -0.003)	0.053 (0.033 to 0.080)
Nationwide	47.1 (30.4 to 64.2)	-13.9 (-19.4 to -9.1)	33.2 (21.3 to 44.8)	2.1 (1.4 to 2.8)	0.7 (0.4 to 0.9)	0.2 (0.1 to 0.3)	0.104 (0.074 to 0.138)	-0.162 (-0.210 to -0.117)	-0.059 (-0.087 to -0.035)

Table S4 | 30-year multi-cause cross-sectional baseline mortality rates of Chinese population.

Mortality rates (per 100 000) of 5 causes (NCDs, non-communicable diseases; CRDs, chronic respiratory diseases; COPD, chronic obstructive pulmonary disease; CVDs, cardiovascular diseases; IHD, ischaemic heart disease) are retrieved from the IHME GBD 2019 result portal (https://vizhub.healthdata.org/gbd-results), with 95% uncertainty intervals.

Year	NCDs	CRDs	COPD	CVDs	IHD
1990	954.5 (856.2, 1049.4)	215.1 (157.8, 241.2)	206.1 (151.1, 231.2)	396.0 (353.5, 443.6)	100.0 (88.3, 111.9)
1991	940.0 (854.0, 1032.0)	211.9 (157.0, 236.6)	203.2 (149.9, 226.9)	388.3 (348.9, 437.3)	99.1 (88.6, 111.4)
1992	925.0 (840.6, 1015.4)	208.8 (154.4, 233.6)	200.4 (147.6, 224.5)	381.3 (341.7, 427.2)	97.7 (87.5, 108.7)
1993	911.1 (829.3, 993.2)	205.0 (151.6, 227.5)	196.8 (144.3, 218.7)	374.7 (340.7, 414.2)	96.5 (87.2, 107.2)
1994	891.7 (818.6, 967.0)	199.4 (147.2, 220.9)	191.5 (140.9, 211.3)	365.0 (331.9, 407.2)	94.2 (86.0, 104.3)
1995	876.4 (814.3, 945.0)	193.3 (143.7, 213.0)	185.5 (137.5, 204.7)	358.7 (329.1, 401.3)	92.8 (84.9, 104.0)
1996	866.5 (807.4, 931.4)	188.0 (139.1, 205.9)	180.5 (133.0, 198.0)	355.7 (326.6, 393.2)	92.5 (84.9, 101.9)
1997	853.5 (802.1, 912.8)	181.3 (136.6, 198.0)	174.1 (130.1, 190.8)	351.5 (326.4, 385.9)	92.1 (85.4, 100.8)
1998	846.5 (791.4, 903.2)	175.5 (134.4, 191.2)	168.6 (128.1, 183.9)	349.8 (323.4, 384.5)	92.6 (85.6, 101.3)
1999	852.9 (801.2, 906.7)	172.7 (136.5, 188.0)	165.8 (130.5, 180.4)	355.2 (328.7, 392.8)	95.1 (88.1, 104.5)
2000	869.2 (816.2, 928.7)	170.8 (136.5, 186.0)	164.0 (130.4, 178.3)	366.5 (339.6, 404.3)	100.4 (93.1, 109.9)
2001	874.7 (817.9, 940.8)	166.2 (138.3, 180.8)	159.5 (131.8, 173.8)	373.4 (345.9, 409.0)	105.6 (97.7, 115.1)
2002	883.9 (827.3, 949.0)	162.5 (135.8, 176.5)	155.9 (129.6, 169.5)	382.5 (352.3, 418.1)	112.4 (103.7, 122.6)
2003	893.4 (834.9, 953.9)	158.7 (136.6, 172.1)	152.2 (130.2, 165.2)	390.9 (362.7, 423.4)	120.2 (111.3, 129.7)
2004	908.8 (852.7, 964.6)	156.3 (137.0, 168.5)	149.8 (130.7, 161.7)	401.0 (371.9, 434.0)	128.2 (118.7, 138.5)
2005	905.0 (848.7, 960.5)	150.4 (132.6, 161.9)	144.2 (127.0, 155.2)	402.2 (373.2, 435.7)	133.1 (123.5, 143.8)
2006	878.9 (826.9, 935.4)	140.2 (126.0, 150.2)	134.4 (120.3, 143.9)	392.4 (363.2, 421.7)	134.1 (124.4, 144.5)
2007	868.2 (817.4, 920.9)	133.4 (120.3, 143.7)	127.8 (115.0, 137.7)	390.5 (362.3, 419.7)	137.0 (126.9, 146.9)
2008	873.9 (821.9, 927.0)	130.1 (116.9, 140.8)	124.7 (112.2, 134.9)	397.7 (367.1, 426.9)	142.7 (131.6, 154.0)
2009	884.9 (835.0, 942.2)	127.9 (115.7, 137.9)	122.5 (110.8, 131.9)	408.4 (378.1, 437.2)	149.9 (138.4, 161.0)
2010	896.7 (834.7, 961.8)	125.3 (113.5, 137.4)	119.9 (108.5, 131.6)	419.9 (384.8, 451.8)	158.0 (144.6, 170.8)
2011	895.2 (832.5, 961.1)	120.6 (108.2, 135.9)	115.3 (103.4, 129.5)	424.3 (386.3, 458.8)	163.0 (148.0, 177.0)
2012	882.7 (820.2, 947.7)	114.9 (104.2, 130.7)	109.7 (99.5, 125.1)	420.3 (385.4, 452.7)	163.5 (149.5, 176.2)
2013	874.6 (804.4, 941.0)	110.1 (99.2, 129.5)	105.0 (94.5, 123.9)	420.3 (382.7, 455.0)	166.3 (151.1, 181.2)
2014	870.7 (800.5, 949.1)	106.4 (95.2, 125.3)	101.5 (90.7, 119.9)	420.0 (379.6, 458.2)	167.9 (151.6, 183.2)
2015	866.7 (784.9, 948.9)	103.4 (92.5, 122.9)	98.5 (88.2, 117.8)	419.1 (378.7, 459.0)	169.1 (152.9, 186.0)
2016	876.4 (787.8, 969.9)	102.9 (89.9, 124.0)	98.1 (85.5, 118.2)	424.4 (377.3, 474.3)	171.6 (152.0, 191.3)
2017	883.9 (788.6, 980.8)	101.8 (89.6, 124.6)	97.1 (85.5, 119.0)	427.8 (378.5, 475.3)	174.3 (155.1, 194.7)
2018	894.6 (788.2, 1008.6)	102.2 (88.1, 124.1)	97.6 (84.1, 119.3)	431.3 (375.8, 488.6)	176.3 (153.9, 200.0)
2019	914.6 (800.2, 1037.7)	104.2 (89.3, 126.8)	99.7 (85.4, 121.6)	439.6 (379.3, 499.7)	179.8 (154.6, 204.5)

Table S5 | Associations between rural-urban ambient ozone difference and land cover features.

The rural-urban differences are defined as localised (i.e. within a prescribed downscaled spatial grid) rural ambient O_3 concentration minus the adjacent urban levels. Backward stepwise selection (p-value <0.20) is adopted to identify associated variables. Features with high collinearity is censored as appropriate (e.g. *emission rate of BC, aerosol optical depth at 550 nm, and surface $PM_{2.5}$ concentrations are deleted due to collinearity with emission rate of OC*). Regression coefficient β_s shows the standardised effect of each feature when controlling all the other considered factors, reported with Wald's p-value and 95% CI. The population-related features are obtained from aforementioned calibration. The emission rates of NO_X, total NMVOC, organic carbon (OC), NH_3, CO and SO_2 are retrieved from Emission Inventory developed by Peking University (PKU-Inventory)[32-42] and Multi-resolution Emission Inventory for China (MEIC)[43-49], while the emission rates of biogenic NMVOC are modelled by CESM2-WACCM (accessed from the CMIP6 repository: https://esgf-node.llnl.gov/search/cmip6). Biomass features, vegetation, and urban land occupation fractions refer to the Land Use Harmonisation database (*historical* experiment for 1990–2014 and *ssp370* experiment for 2015–2019)[50,51].

Features	β_s	p-value	95% CI
Population and urbanisation indices			
lg-transformed total population	1.832	<0.001	(1.761, 1.902)
urban population fraction	0.144	<0.001	(0.106, 0.182)
urban land occupation	0.086	0.001	(0.036, 0.136)
Emission rate			
emission rate of NO_X	−0.053	0.10	(−0.117, 0.010)
emission rate of total NMVOC	0.138	<0.001	(0.094, 0.182)
emission rate of biogenic NMVOC	0.231	<0.001	(0.193, 0.270)
emission rate of OC	1.379	<0.001	(1.286, 1.473)
emission rate of NH_3	−0.030	0.18	(−0.075, 0.014)
emission rate of CO	0.164	<0.001	(0.133, 0.195)
emission rate of SO_2	0.156	<0.001	(0.102, 0.210)
Vegetation land occupation			
C_3 annual and perennial crops	0.201	0.006	(0.057, 0.345)
C_4 annual and perennial crops	0.316	<0.001	(0.184, 0.449)
pasture	0.370	<0.001	(0.313, 0.427)
rangeland	0.826	<0.001	(0.728, 0.925)
primary forested land	0.397	<0.001	(0.349, 0.445)
primary non-forested land	0.669	<0.001	(0.583, 0.755)
secondary forested land	1.015	<0.001	(0.941, 1.090)
secondary non-forested land	0.118	<0.001	(0.075, 0.162)
Biomass features			
secondary mean age	0.184	<0.001	(0.146, 0.223)
secondary mean biomass carbon density	0.237	<0.001	(0.171, 0.302)

Interpretation: The research hypothesis to test is that "spatial pattern of the rural-urban ambient O_3 differences can be reflected by sociodemographic and geographical features in spatial statistics". Taking the variable "urban land occupation" as an example, the standardised coefficient is positive, as β_s = 0.086, 95% CI: 0.036–0.136, which means summarising from all studied cells across the 30 years, **the greater the urban land occupation is, the larger the rural–urban ambient O_3 gap will be**. This coincides with the fact that greater urban land occupations usually indicate higher emissions to form aerosols, and higher urban aerosols suppress the urban O_3 formation, finally making the rural–urban gaps greater (urban ↓, rural–urban ↑). Relevant characteristics such as urban population fraction (β=0.144, 95% CI: 0.106–0.182), and organic carbon emission (β=1.379, 95% CI: 1.286–1.473) thus also show positive partial correlations. For another example, the coefficient of C_3 annual and perennial crops is also positive as β_s = 0.201, 95% CI: 0.057–0.345. This is a typical rural indicator, meaning that larger C_3 crop vegetated land occupations usually indicate higher biogenic VOC emissions to form rural O_3, finally making the rural–urban gaps greater (rural ↑, rural–urban ↑). The other studied features can be interpreted in similar way, that emission rate of CO (β=0.164, 95% CI: 0.133–0.195), emission rate of biogenic non-methane VOCs (β=0.231, 95% CI: 0.193–0.270), and other vegetation coverage (e.g. cropland, pasture and rangeland), as rural indicators, also display positive associations with intensified rural O_3 pollution.

Table S6 | Performance evaluations of phased data fusion with urban-rural distinguishment.

Algorithm performance assessments include 10-fold cross-validation tests and full-scale overall evaluations separately for urban and rural sites for phased data fusion. Full-scale refers to model training, prediction and evaluation using full dataset. Due to heterogeneity in input data, cross-validation tests for 30-year full-length evaluation are not applicable (NA).

		Cross-validation test		Full-scale evaluation		Scale
		R^2	RMSE (ppb)	R^2	RMSE (ppb)	
Phase I	*urban*	0.84	4.2	0.93	3.2	Global
	rural	0.85	5.1	0.90	4.8	Global
Phase II	*urban*	0.88	4.2	0.94	3.6	Global
	rural	0.90	5.9	0.93	5.1	Global
Phase III	*urban*	0.82	4.9	0.91	4.2	China
	rural	0.86	7.0	0.89	5.2	China
30-year	*urban*	NA	NA	0.90	3.6	Global
	rural	NA	NA	0.93	5.0	Global

Table S7 | Evaluation of spatial and temporal extrapolation accuracy by space-time Bayesian neural network downscaler with urban-rural differentiation.

Different from classical cross-validation tests by randomly splitting the dataset, spatiotemporal generalisability validation tests manually divide the initial dataset by location or time period. Region-clustered spatial generalisability tests use observations in aggregated regions for algorithm training, and assign observations in other aggregated regions for testing, including four sub-experiments (cross-validation for spatial generalisability, cvs_1: training on North America, testing on Europe; cvs_2: training on Europe, testing on North America; cvs_3: training on North America and Europe, testing on Asia; and cvs_4: training on locations outside China, testing on China). Period-staged temporal generalisability tests treat six consecutive years as testing subset based on trainings from the rest 24-year global-scale dataset, including five sub-experiments (cross-validation for temporal generalisability, cvt_1: training on 1990–2013, testing on 2014–2019; cvt_2: training on 1990–2007 and 2014–2019, testing on 2008–2013; cvt_3: training on 1990–2001 and 2008–2019, testing on 2002–2007; cvt_4: training on 1990–1995 and 2002–2019, testing on 1996–2001; cvt_5: training on 1996–2019, testing on 1990–1995). Prediction evaluation statistics include crude R^2 and RMSE (in ppb) before 1:1 linear regression calibration, together with linear regression slope (*k*) and intercept (*b*).

	Urban				Rural			
Spatial extrapolation	R^2	RMSE (ppb)	*k*	*b*	R^2	RMSE (ppb)	*k*	*b*
cvs_1	0.89	6.3	0.89	4.14	0.88	6.7	0.93	4.43
cvs_2	0.89	6.0	0.92	4.28	0.86	7.3	0.88	3.66
cvs_3	0.85	5.1	0.85	7.15	0.85	7.9	0.82	5.01
cvs_4	0.88	4.9	0.80	9.65	0.81	6.6	0.87	2.84
Temporal extrapolation								
cvt_1	0.90	5.7	0.92	1.65	0.89	4.7	1.07	−0.51
cvt_2	0.88	5.0	0.93	1.89	0.84	5.3	1.05	−0.52
cvt_3	0.91	4.9	0.92	1.44	0.84	4.6	1.02	−0.53
cvt_4	0.87	5.1	0.91	1.67	0.84	4.4	1.02	−0.56
cvt_5	0.85	4.7	0.91	1.38	0.82	4.8	1.01	−0.29

Table S8 | Quality assessment tool for observational cohort and cross-sectional studies.

A. Was the research question or objective in this paper clearly stated?
B. Was the study population clearly specified and defined?
C. Was the participation rate of eligible persons at least 50%?
D. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?
E. Was a sample size justification, power description, or variance and effect estimates provided?
F. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?
G. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?
H. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?
I. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?
J. Was the exposure(s) assessed more than once over time?
K. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?
L. Were the outcome assessors blinded to the exposure status of participants?
M. Was loss to follow-up after baseline 20% or less?
N. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?

Source: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools.

Table S9 | Quality assessment of 29 included cohort studies for meta-analysis.

Study-specific quality assessments aim to examine the reliability of the epidemiological evidence and ensure the quality for meta-analysis. A total of 14 assessment items are considered according to the Quality Assessment Tool of Observational Cohort and Cross-Sectional Studies developed by the National Institute of Health (NIH) (Table S8), and assigned with one score for each, and the tallied scores are translated into a rating of quality. Studies scoring full marks, 14, are categorised as “Good,” 10–13 as “Fair”, and <10 as “Poor.”

Study	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Score	Ref
Abbey et al. 1999	√	√	√	√	√	√	√	√		√	√	√	√	√	Fair	52
Lipfert et al. 2006	√	√	√		√	√	√	√		√	√	√	√	√	Fair	53
Jerrett et al. 2009	√	√	√	√	√	√	√	√		√	√	√	√	√	Fair	21
Krewski et al. 2009	√	√	√	√	√		√	√		√	√	√	√	√	Fair	54
Smith et al. 2009	√	√	√		√		√	√		√	√	√	√	√	Fair	55
Lipsett et al. 2011	√	√	√		√	√	√	√		√	√	√	√	√	Fair	56
Zanobetti et al. 2011	√	√	√	√	√	√	√	√		√	√	√	√	√	Fair	57
Carey et al. 2013	√	√	√		√	√		√		√	√	√	√	√	Fair	58
Jerrett et al. 2013	√	√	√	√	√	√	√	√		√	√	√	√	√	Fair	59
Bentayeb et al. 2015	√	√	√		√	√	√	√	√	√	√	√	√	√	Fair	60
Crouse et al. 2015	√	√	√		√	√	√	√	√	√	√	√	√	√	Fair	61
Tonne et al. 2016	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	62
Turner et al. 2016	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	63
Di et al. 2017	√	√	√		√	√	√	√	√	√	√	√	√	√	Fair	64
Weichenthal et al. 2017	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	65
Cakmak et al. 2018	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	66
Hvidtfeldt et al. 2019	√	√	√		√	√		√	√	√	√	√	√	√	Fair	67
Kazemiparkouhi et al. 2019	√	√	√		√	√	√	√		√	√	√	√	√	Fair	68
Lim et al. 2019	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	69
Paul et al. 2020	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	70
Shi et al. 2021	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	71
Strak et al. 2021	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	72
Yazdi et al. 2021	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	73
Bauwelinck et al. 2022	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	74
Stafoggia et al. 2022	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	75
So et al. 2022	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	76
Liu et al. 2022	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	31
Niu et al. 2022	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	30
Yuan et al. 2022	√	√	√	√	√	√	√	√	√	√	√	√	√	√	Good	77

Table S10 | GRADE assessment for evidence of ozone-associated mortality risks of NCDs.

Domains	Assessment	Rating
Start level	All cohort studies.	High
Risk of bias	The overall risk of bias in all cohorts is low.	No change
Imprecision	All studies included report the 95% confidence interval around the best estimate of the absolute effect.	No change
Inconsistency	The values of effect sizes across the studies are inconsistent, as the point estimates are in the range of 0.816 to 1.108.	Downgrade
Indirectness	All studies include the desired population, exposures and outcomes.	No change
Publication bias	The trim-and-fill tool detects 1 study (Yuan et al. 2022) reporting significant positive publication bias, which is excluded in censored meta-analysis. The publication bias for censored meta-analysis is non-significant.	No change
Magnitude of associations	The magnitude of effect sizes is not large enough to upgrade the level of evidence.	No change
Dose-response trend	Linear dose-response relationships are assumed in all studies, and at least 4 studies after censoring (Di et al. 2017, Shi et al. 2021, Bauwelinck et al. 2022, and So et al. 2022) have checked the dose-response trends.	Upgrade
Plausible confounding towards null	Cakmak et al. 2018 reports higher RR after adjusting confounders; but 1 study out of 29 reporting plausible confounding is not sufficient for an upgrading.	No change
Overall Judgment	**High**	

Table S11 | GRADE assessment for evidence of ozone-associated mortality risks of CRDs.

Domains	Assessment	Rating
Start level	All cohort studies.	High
Risk of bias	The overall risk of bias in all cohorts is low.	No change
Imprecision	All studies included report the 95% confidence interval around the best estimate of the absolute effect.	No change
Inconsistency	The values of effect sizes across the studies are inconsistent, as the point estimates are in the range of 0.782 to 1.144	Downgrade
Indirectness	All studies include the desired population, exposures and outcomes.	No change
Publication bias	The publication bias for censored meta-analysis is non-significant.	No change
Magnitude of associations	The magnitude of effect sizes is not large enough to upgrade the level of evidence.	No change
Dose-response trend	Linear dose-response relationships are assumed in all studies, and at least 3 out of 11 censored studies (Lim et al. 2019, Bauwelinck et al. 2022, and So et al. 2022) have tested dose-response trends.	Upgrade
Plausible confounding towards null	No crude and adjusted risks are provided for each study.	No change
Overall Judgment	**High**	

Table S12 | GRADE assessment for evidence of ozone-associated mortality risks of COPD.

Domains	Assessment	Rating
Start level	All cohort studies.	High
Risk of bias	The overall risk of bias in all cohorts is low.	No change
Imprecision	All studies included report the 95% confidence interval around the best estimate of the absolute effect.	No change
Inconsistency	The values of effect sizes across the studies are inconsistent, as the point estimates are in the range of 0.746 to 1.090.	Downgrade
Indirectness	All studies include the desired population, exposures and outcomes.	No change
Publication bias	The publication bias for censored meta-analysis is non-significant.	No change
Magnitude of associations	The magnitude of effect sizes (RR=1.060, 95% CI: 1.040–1.080) can be considered to upgrade the level of evidence.	Upgrade
Dose-response trend	Linear dose-response relationships are assumed in all studies, but no studies check dose-response trends.	No change
Plausible confounding towards null	No crude and adjusted risks are provided for each study.	No change
Overall Judgment	**High**	

Table S13 | GRADE assessment for evidence of ozone-associated mortality risks of CVDs.

Domains	Assessment	Rating
Start level	All cohort studies.	High
Risk of bias	The overall risk of bias in all cohorts is low.	No change
Imprecision	All studies included report the 95% confidence interval around the best estimate of the absolute effect.	No change
Inconsistency	The values of effect sizes across the studies are inconsistent, as the point estimates are in the range of 0.831 to 1.249.	Downgrade
Indirectness	All studies include the desired population, exposures and outcomes.	No change
Publication bias	The publication bias for censored meta-analysis is non-significant.	No change
Magnitude of associations	The magnitude of effect sizes is not large enough to upgrade the level of evidence.	No change
Dose-response trend	Linear dose-response relationships are assumed in all studies, and at least 7 out of 15 studies (Lim et al. 2019, Paul et al. 2020, Strak et al. 2021, Bauwelinck et al. 2022, So et al. 2022, Liu et al. 2022, and Niu et al. 2022) have checked the dose-response trends.	Upgrade
Plausible confounding towards null	No crude and adjusted risks are provided for each study.	No change
Overall Judgment	**High**	

Table S14 | GRADE assessment for evidence of ozone-associated mortality risks of IHD.

Domains	Assessment	Rating
Start level	All cohort studies.	High
Risk of bias	The overall risk of bias in all cohorts is low.	No change
Imprecision	All studies included report the 95% confidence interval around the best estimate of the absolute effect.	No change
Inconsistency	The values of effect sizes across the studies are inconsistent, as the point estimates are in the range of 0.761 to 1.360.	Downgrade
Indirectness	All studies include the desired population, exposures and outcomes.	No change
Publication bias	The publication bias for censored meta-analysis is non-significant.	No change
Magnitude of associations	The magnitude of effect sizes is not large enough to upgrade the level of evidence.	No change
Dose-response trend	Linear dose-response relationships are assumed in all studies, and at least 3 (Strak et al. 2021, Liu et al. 2022, and Niu et al. 2022) out of 8 censored studies have considered dose-response trends.	Upgrade
Plausible confounding towards null	Cakmak et al. 2018 reports higher RR after adjusting confounders; but 1 study reporting plausible confounding is not sufficient for an upgrading.	No change
Overall Judgment	**High**	

Table S15 | Statistically resampled distributions of ozone exposure levels for each study.

The distribution features include arithmetic mean, standard deviation (SD), minimum, 5th, 25th, 50th (median), 75th, and 95th percentile, maximum, inter-quartile range (IQR), and full range, based on ozone exposure concentrations scaled by OSDMA8 metric in ppb. Values in **Bold** font represent the statistics reported by literature, while the rest indicate imputed values. Detailed resampling procedures and imputation accuracy evaluation can be found in a previous study[14].

Study	Mean	SD	Min	5%	25%	Median	75%	95%	Max	IQR	Range
Abbey et al. 1999	**50.4**	**14.9**	16.1	26.1	40.6	50.4	60.5	74.8	84.9	**23.2**	**84.9**
Lipfert et al. 2006	**80.1**	**9.7**	**36.6**	64.2	73.5	80.1	86.7	96.1	**106.6**	13.2	69.9
Jerrett et al. 2009 Krewski et al. 2009 Smith et al. 2009	50.1	12.6	**27.5**	30.0	41.6	50.1	58.5	70.7	**86.1**	17.0	**58.5**
Lipsett et al. 2011	**55.6**	**10.1**	**29.4**	39.1	48.8	55.6	62.4	72.3	**95.5**	**12.9**	**66.1**
Zanobetti et al. 2011	**45.9**	5.2	**26.6**	**40.1**	**44.0**	**48.4**	**51.1**	**52.5**	**71.2**	**6.9**	**44.7**
Carey et al. 2013	**51.0**	**2.3**	**43.8**	47.2	49.5	51.0	52.6	54.9	**62.0**	**2.9**	**18.1**
Jerrett et al. 2013	**58.3**	**16.9**	**19.8**	**33.3**	**42.5**	**58.7**	**70.5**	**85.8**	**103.2**	**28.0**	**83.5**
Bentayeb et al. 2015	**49.4**	**4.9**	20.3	25.4	45.5	**48.9**	52.1	57.0	60.2	**6.2**	39.9
Crouse et al. 2015	**39.5**	7.3	**10.7**	26.8	**34.2**	**39.0**	**44.0**	51.0	**59.9**	**9.8**	49.1
Tonne et al. 2016	**39.8**	**3.8**	30.7	33.4	**37.3**	**40.0**	**42.5**	46.4	49.0	**5.2**	18.4
Turner et al. 2016	**44.2**	**4.6**	**30.1**	36.5	41.0	44.2	47.3	51.8	**68.6**	6.2	**37.7**
Di et al. 2017	**46.3**	9.9	54.0	**36.3**	70.5	**77.1**	83.8	**55.9**	100.2	13.3	46.1
Weichenthal et al. 2017	38.1	6.6	**1.0**	**27.5**	**33.6**	**38.0**	**42.5**	**50.4**	**60.3**	**9.0**	**59.3**
Cakmak et al. 2018	**39.1**	**6.7**	**0.0**	28.1	34.6	39.1	43.6	50.1	**58.6**	9.0	**58.6**
Hvidtfeldt et al. 2019	54.7	4.9	43.5	**44.0**	51.3	**54.7**	57.8	**59.9**	65.9	6.6	22.4
Kazemiparkouhi et al. 2019	**45.1**	**5.3**	**31.0**	36.3	41.5	45.1	48.7	53.9	**65.1**	7.2	**34.1**
Lim et al. 2019	45.5	6.1	31.3	35.4	**41.4**	**45.5**	**49.7**	55.6	59.8	**8.3**	28.6
Paul et al. 2020	**46.8**	4.7	35.8	39.0	43.6	46.8	49.9	54.6	57.8	**6.4**	22.0
Shi et al. 2021	**40.2**	**4.8**	**17.9**	**30.5**	**37.5**	**40.9**	**43.3**	**47.2**	**50.0**	**5.8**	**32.1**
Strak et al. 2021	**43.5**	**4.6**	**18.5**	**36.0**	**40.1**	**44.0**	**47.3**	**49.7**	**58.9**	**7.2**	**40.4**
Yazdi et al. 2021	**41.9**	3.9	**31.9**	35.5	**39.4**	**42.5**	**44.7**	48.3	**50.0**	**5.3**	**18.1**
Bauwelinck et al. 2022	**39.5**	1.6	**19.8**	**34.9**	**38.3**	**39.5**	**40.5**	**42.6**	**46.4**	**2.2**	**26.7**
So et al. 2022	**40.9**	**2.2**	**24.9**	**36.0**	**40.1**	**41.4**	**42.2**	**43.5**	**46.9**	**2.1**	**22.0**
Liu et al. 2022	**37.4**	**1.2**	33.7	35.4	36.6	37.4	38.2	39.4	43.0	1.6	9.3
Niu et al. 2022	**45.8**	**7.3**	28.8	33.8	40.9	45.8	50.7	57.8	62.8	9.8	34.0
Yuan et al. 2022	**51.4**	9.0	**31.0**	36.7	45.4	51.4	57.4	66.1	**72.7**	12.0	**41.7**

Note: Jerrett et al. 2009 did not report the arithmetic mean and standard deviation directly. The values were derived by weighted averaging the centric concentrations of 4 exposure intervals on the populations given in Table 1 from the original literature. Zanobetti et al. 2011 did not provide the exposure distribution features directly. The quartiles were extracted from the legends in Fig. 1 of the original literature.

Methods: To reproduce the distribution, the arithmetic means and standard deviations (σ) were firstly extracted from literatures included for meta-analysis; if unavailable, the arithmetic means and standard deviations were estimated based on the reported descriptive statistics including median, first- and third-quartile, and all the other percentiles, to finally identify the parameters for presumed Gaussian normal distribution. Reported values were always treated as priority when divergences with estimations occurred. The centric level, arithmetic mean and median, were treated as exchangeable, but the arithmetic means were preferred. Theoretically, the minimum and maximum values of the distribution were not predictable, and thus 1st and 99th percentiles were used as proxies. Calculations for σ from key percentiles followed: 75th%ile = mean + 0.6745 σ, 95th%ile = mean + 1.6449 σ, and 99th%ile = mean + 2.3263 σ. If IQRs were stated, then IQR = 1.3490 σ; if the 5–95th percentile ranges were reported, then $range_{5\text{-}95}$ = 3.2898 σ; if full minimum-maximum ranges were given, then range = 4.6527 σ. If more than one distribution features were provided, IQRs were more preferred for σ estimation due to higher robustness.

Table S16 | Evaluations of accuracies of deep-learning-based data assimilation with (ScA) and without (ScB) satellite-based remote-sensing measurements and chemical reanalysis outputs.

Accuracy evaluations include coefficient of determination (R^2) and root-mean-square error (RMSE, ppb) for 10-fold cross-validation tests using 70% observation-matched dataset by random split, external validation tests using 30% dataset, and overall model fitting for the two scenarios respectively. Given systematic *in situ* observations were unavailable in earlier years of China, and CNEMC sites were allocated in urban and rural environments disproportionally, model fitting and performance evaluations are conducted on global scale.

Evaluation Metrics	ScA	ScB
Cross-validation R^2	0.883	0.882
Cross-validation RMSE (ppb)	3.887	3.876
External validation R^2	0.885	0.883
External validation RMSE (ppb)	3.879	3.868
Overall fitting R^2	0.969	0.968
Overall fitting RMSE (ppb)	2.550	2.542

Table S17 | Multi-scenario sensitivity analysis.

Sensitivity analyses are conducted on the estimation for 2017 as an example by multiple designed scenarios (Sc) beyond the main analysis. Cardiopulmonary mortality numbers are estimated for urban and rural population separately. Changes in total population mortalities (%) for different scenarios against the main analysis results are calculated. **Sc1:** Using log-linear risk model (rather than curved risk model in main analysis) with multi-study pooled RRs by random-effects meta-analysis, assuming threshold exposure level (also known as TMREL or low-concentration cut-off) as the global lowest 5th percentile PWE in 2017 by BayNNDv2 dataset (see Method S1), 42.6 ppb. **Sc2:** Using log-linear risk model assuming threshold as the 30-year global lowest 5th percentile PWE by BayNNDv2 dataset, 40.8 ppb. **Sc3:** Using log-linear risk model assuming threshold as the maximum of literature-reported lowest 5th percentile exposure levels from studies included for meta-analysis, 44.0 ppb. **Sc4:** Using grid-averaged ambient ozone concentrations to quantify population exposure (following a previous study[1]), supposing the ambient ozone concentrations are not distinguished for urban and rural environments. **Sc5:** Using gender-specified other than the gender-standardised mortality metrics provided by IHME[15] (GBD 2019 Study report). **Sc6:** Using province-specific mortality metrics for 2017 provided by China CDC[29], as the cause-specific mortality rates are proportionally converted from the estimated DALY (disability-adjusted life years) rates. **Sc7:** Using M^3-BME ambient ozone tracking data product instead of the fused one. As M^3-BME did not distinguish urban and rural ozone, urban and rural mortalities were not applicable (NA). **Sc8:** Using cardiovascular mortality linear risk association (RR=1.227, 95% CI: 1.108–1.359) pooled from two cohort studies exclusively on Chinese population[73,74].

Scenarios	Urban Mortality (thousand)	Rural Mortality (thousand)	Total Mortality (thousand)	Change (%)
Main Result	191.2 (123.6 to 260.0)	172.5 (111.4 to 234.9)	363.7 (235.0 to 495.0)	Ref.
Sc1	179.1 (113.0 to 248.9)	160.0 (100.8 to 222.6)	339.1 (213.8 to 471.5)	-6.74 (-8.99 to -4.74)
Sc2	188.5 (119.0 to 261.8)	168.3 (106.1 to 234.0)	356.8 (225.1 to 495.8)	-1.88 (-4.18 to 0.16)
Sc3	173.4 (109.3 to 241.1)	155.1 1(97.6 to 215.8)	328.5 (207.0 to 456.9)	-9.66 (-11.9 to -7.68)
Sc4	189.9 (119.9 to 263.8)	137.8 1(86.7 to 192.0)	327.8 (206.6 to 455.8)	-9.88 (-12.1 to -7.91)
Sc5	195.0 (119.3 to 275.5)	176.0 (107.5 to 248.9)	371.0 (226.8 to 524.3)	2.01 (-3.45 to 5.93)
Sc6	201.0 (127.7 to 279.9)	181.4 (115.1 to 252.9)	382.4 (242.7 to 532.8)	5.15 (3.31 to 7.65)
Sc7	NA	NA	332.5 (212.9 to 460.6)	-8.58 (-9.40 to -6.94)
Sc8	211.1 (129.6 to 293.3)	190.5 (116.8 to 265.0)	401.6 (246.4 to 558.3)	10.4 (4.85 to 12.8)

SUPPLEMENTARY FIGURES

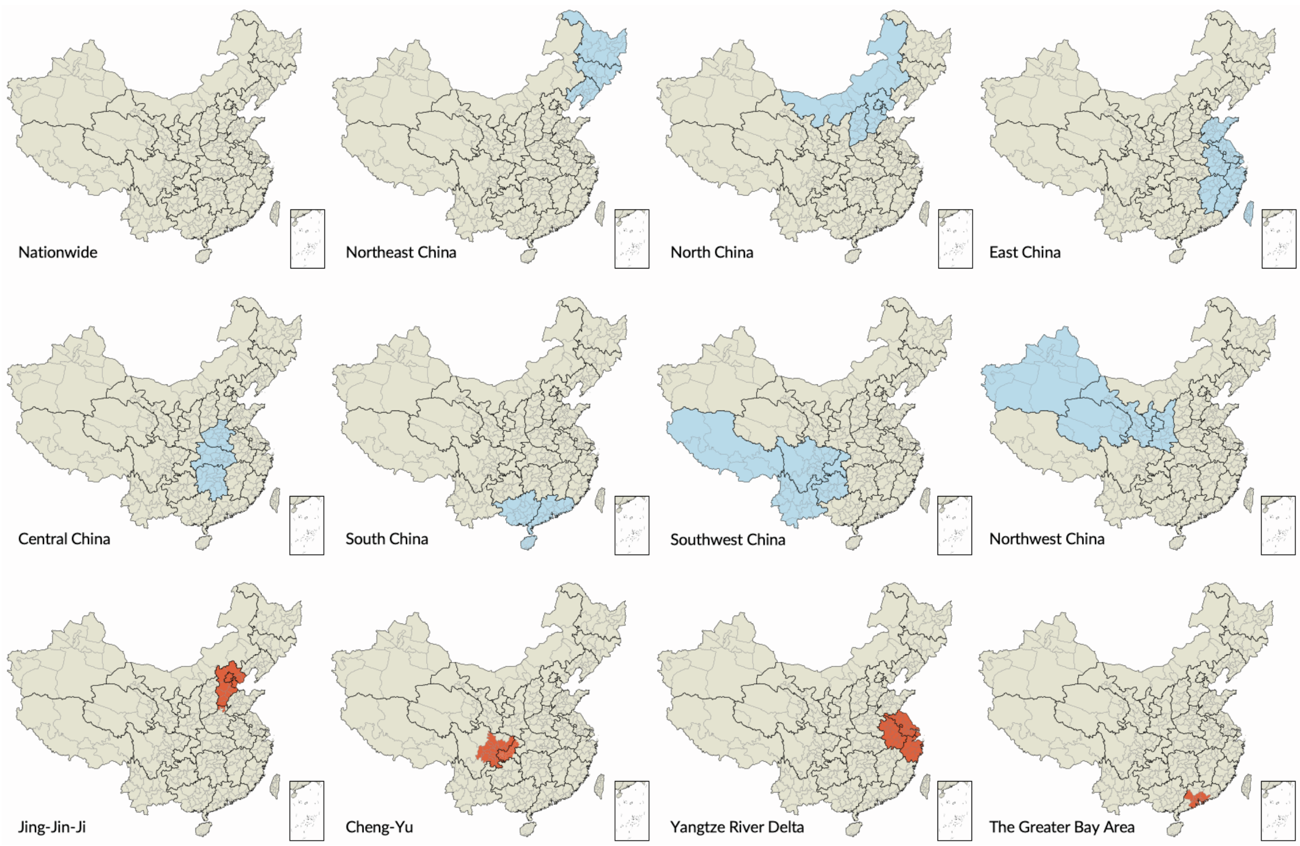


Figure S1 | Mapping of 7 Chinese administrative divisions and 4 megalopolises.

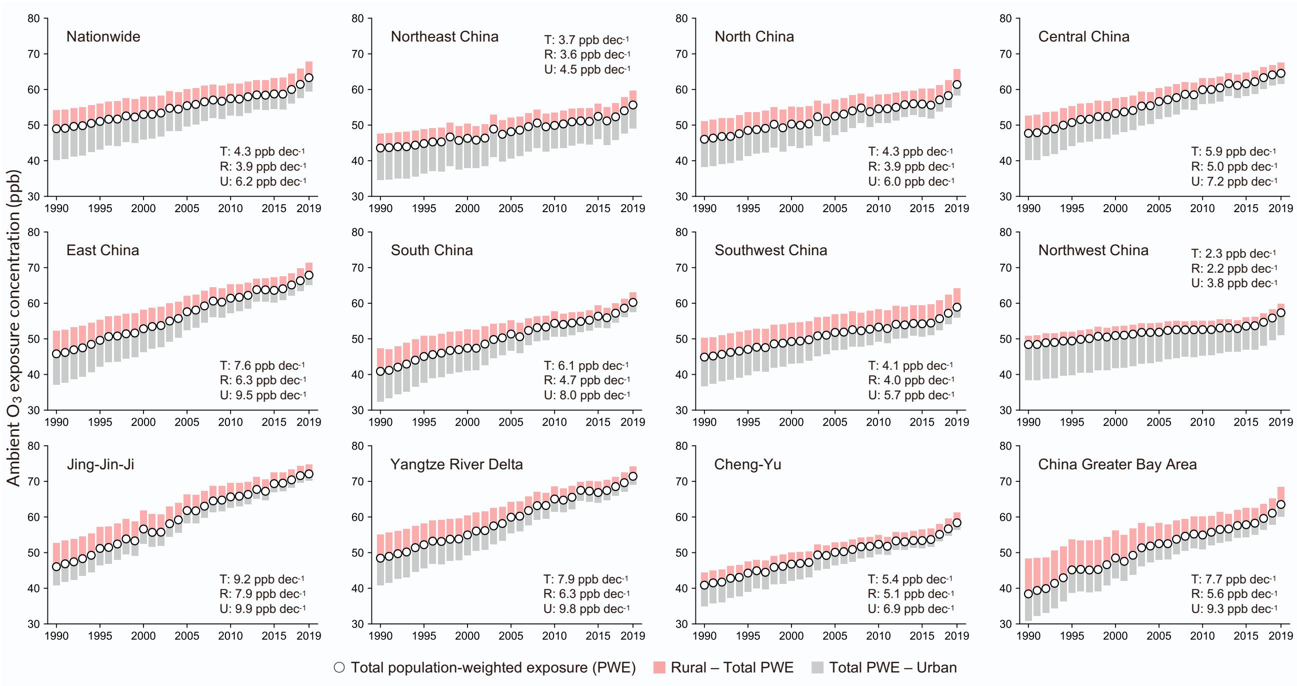


Figure S2 | Nationwide and regional 30-year longitudinal trends of ambient ozone exposure.

Population-weighted exposure (PWE) of total, rural- and urban-specified average exposure levels to ambient ozone are scaled in metric of OSDMA8. PWE levels are indicated by circles, based on which the rural–total (defined as rural-population average minus total PWE, similarly hereinafter) and total–urban differences are marked with directional bars. Upper apexes and lower vertexes represent nationwide or regional average ambient ozone exposure concentrations for rural and urban residents, respectively. Decadal average increasing rates (ppb per decade) are estimated by generalised linear model, as inserted in each subplot (T for total PWE; R for rural population exposure levels; U for urban population exposure levels). Longitudinal trends are summarised for nationwide, 7 geographical divisions, and 4 megalopolises (see Figure S1 for detailed definition).

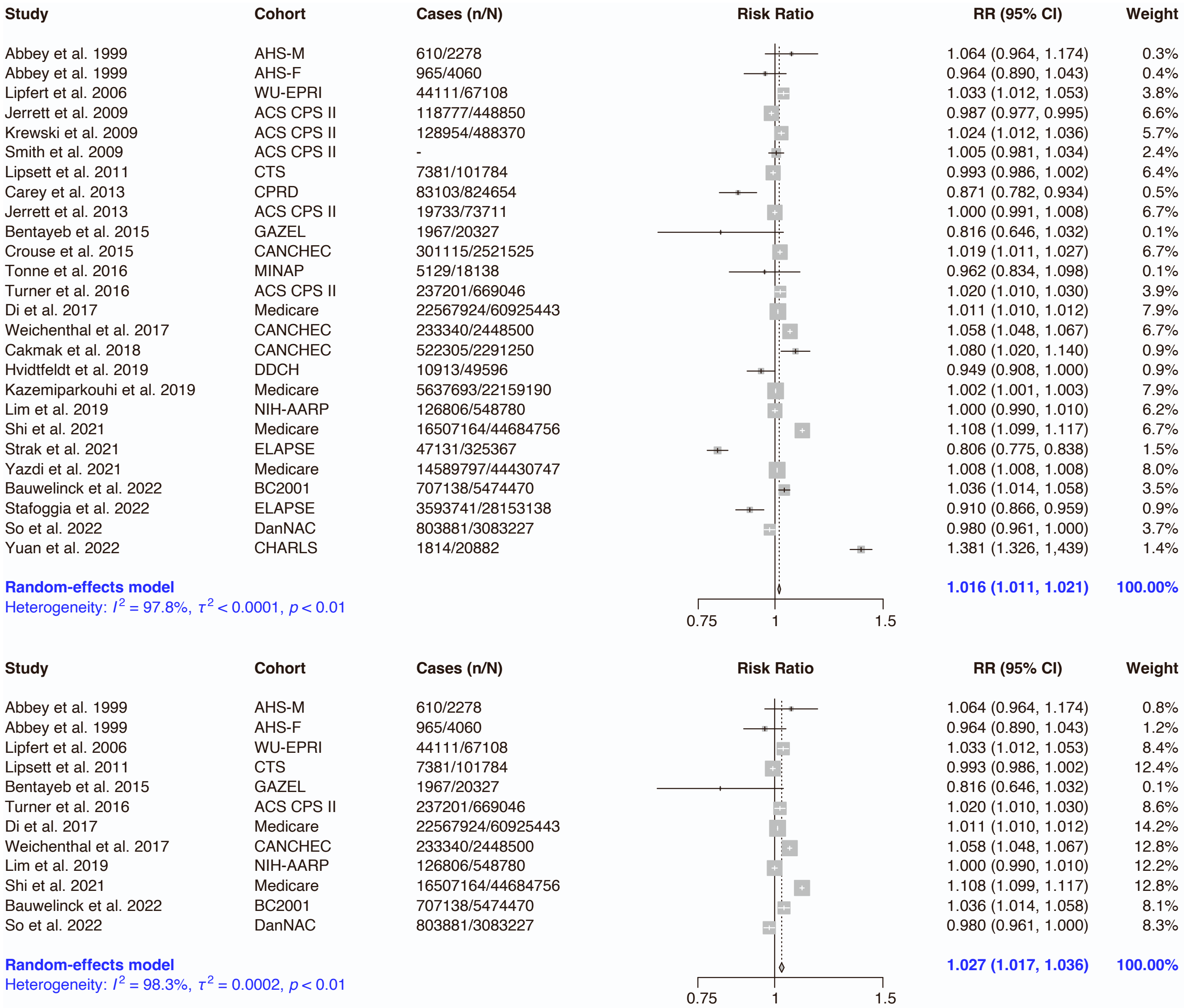
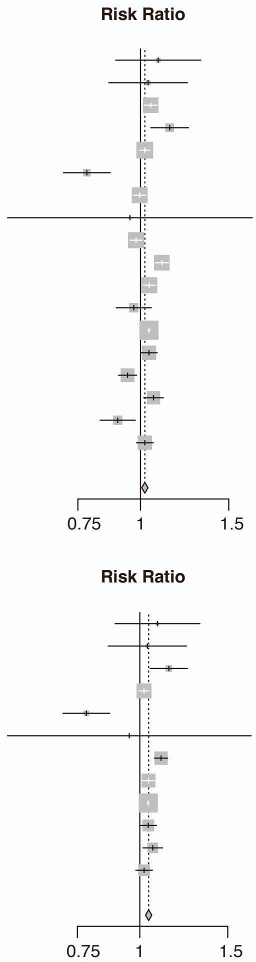


Figure S3 | Multi-study pooled mortality RR of NCDs associated with long-term ozone exposure.

Risk strengths are defined as RRs per 10-ppb incremental exposure by OSDMA8 metric. The upper panel displays the meta-analysis results for all relevant cohort studies identified from systematic review, and the lower panel, censored meta-analysis, excludes i) studies conducted from the same cohort; ii) studies using over-smoothed metrics (e.g. 24-hour average) to quantify the individual-level exposure; iii) studies showing significant publication bias by trim-and-fill test (Figure S8); and iv) studies in which ozone hazards are mistakenly confounded by correlated or anticorrelated air pollutant species (e.g. NO_2). For cohort duplication censoring, only one study covering the widest population is reserved in principle; unless different participant inclusion criteria are clearly stated (e.g. Di et al.[64] conducted study on the whole Medicare cohort participants while Shi et al.[71] focused on the low-exposure participants, thus both included for meta-analysis). Methodology of metric and unit unification has been illustrated in a previous review[14]. Supplementary Figs. 4-7 follow the same configuration.



Study	Cohort	Cases (n/N)	RR (95% CI)	Weight
Abbey et al. 1999	AHS-M	63/2278	1.085 (0.890, 1.319)	0.5%
Abbey et al. 1999	AHS-F	72/4060	1.036 (0.867, 1.241)	0.6%
Jerrett et al. 2009	ACS CPS II	9819/448850	1.048 (1.016, 1.081)	8.1%
Smith et al. 2009	ACS CPS II	-	1.144 (1.048, 1.247)	2.3%
Lipsett et al. 2011	CTS	702/101784	1.020 (0.993, 1.044)	9.5%
Carey et al. 2013	CPRD	10583/824654	0.782 (0.699, 0.871)	1.6%
Jerrett et al. 2013	ACS CPS II	1973/73711	1.004 (0.978, 1.030)	8.5%
Bentayeb et al. 2015	GAZEL	284/20327	0.953 (0.554, 1.671)	0.1%
Crouse et al. 2015	CANCHEC	24900/2521525	0.980 (0.953, 1.007)	8.7%
Turner et al. 2016	ACS CPS II	20484/669046	1.080 (1.060, 1.110)	8.2%
Weichenthal et al. 2017	CANCHEC	21100/2448500	1.041 (1.011, 1.070)	8.7%
Hvidtfeldt et al. 2019	DDCH	2093/49596	0.970 (0.888, 1.051)	2.6%
Kazemiparkouhi et al. 2019	Medicare	633216/22159190	1.033 (1.030, 1.037)	12.6%
Lim et al. 2019	NIH-AARP	12459/548780	1.040 (1.000, 1.080)	6.8%
Strak et al. 2021	ELAPSE	2865/325367	0.796 (0.679, 0.934)	6.2%
Bauwelinck et al. 2022	BC2001	82341/5474470	1.062 (1.014, 1.111)	5.7%
Stafoggia et al. 2022	ELAPSE	371990/28153138	0.901 (0.831, 0.977)	2.6%
So et al. 2022	DanNAC	223553/3083227	1.020 (0.982, 1.060)	6.7%
Random-effects model			**1.020 (1.006, 1.035)**	**100.0%**

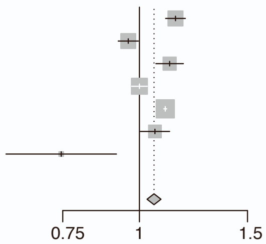
Heterogeneity: $I^2 = 84.9\%$, $\tau^2 = 0.0004$, $p < 0.01$

Study	Cohort	Cases (n/N)	RR (95% CI)	Weight
Abbey et al. 1999	AHS-M	63/2278	1.085 (0.890, 1.319)	0.4%
Abbey et al. 1999	AHS-F	72/4060	1.036 (0.867, 1.241)	0.5%
Smith et al. 2009	ACS CPS II	-	1.144 (1.048, 1.247)	1.9%
Lipsett et al. 2011	CTS	702/101784	1.020 (0.993, 1.044)	13.9%
Carey et al. 2013	CPRD	10583/824654	0.782 (0.699, 0.871)	1.2%
Bentayeb et al. 2015	GAZEL	284/20327	0.953 (0.554, 1.671)	0.1%
Turner et al. 2016	ACS CPS II	20484/669046	1.080 (1.060, 1.110)	11.0%
Weichenthal et al. 2017	CANCHEC	21100/2448500	1.041 (1.011, 1.070)	12.4%
Kazemiparkouhi et al. 2019	Medicare	633216/22159190	1.033 (1.030, 1.037)	37.2%
Lim et al. 2019	NIH-AARP	12459/548780	1.040 (1.000, 1.080)	7.9%
Bauwelinck et al. 2022	BC2001	82341/5474470	1.062 (1.014, 1.111)	6.0%
So et al. 2022	DanNAC	223553/3083227	1.020 (0.982, 1.060)	7.6%
Random-effects model			**1.042 (1.029, 1.055)**	**100.0%**

Heterogeneity: $I^2 = 76.8\%$, $\tau^2 = 0.0001$, $p < 0.01$

Figure S4 | Multi-study pooled mortality RR of CRDs associated with ozone exposure.

Study	Cohort	Cases (n/N)	Risk Ratio	RR (95% CI)	Weight
Zanobetti et al. 2011	Medicare	1445000/3210511		1.145 (1.082, 1.188)	16.2%
Crouse et al. 2015	CANCHEC	14170/2521525		0.959 (0.924, 0.996)	16.0%
Turner et al. 2016	ACS CPS II	9967/669046		1.090 (1.050, 1.130)	12.4%
Cakmak et al. 2018	CANCHEC	16470/2291250		1.000 (0.970, 1.030)	18.4%
Kazemiparkouhi et al. 2019	Medicare	328957/22159190		1.084 (1.079, 1.089)	23.8%
Lim et al. 2019	NIH-AARP	7748/548780		1.060 (1.010, 1.120)	11.6%
Strak et al. 2021	ELAPSE	1711/325367		0.746 (0.605, 0.917)	1.6%
Random-effects model				**1.056 (1.029, 1.084)**	**100.0%**
Heterogeneity: $I^2 = 94.5\%$, $\tau^2 = 0.0007$, $p < 0.01$					



Study	Cohort	Cases (n/N)	Risk Ratio	RR (95% CI)	Weight
Turner et al. 2016	ACS CPS II	9967/669046		1.090 (1.050, 1.130)	18.0%
Cakmak et al. 2018	CANCHEC	16470/2291250		1.000 (0.970, 1.030)	22.8%
Kazemiparkouhi et al. 2019	Medicare	328957/22159190		1.084 (1.079, 1.089)	42.7%
Lim et al. 2019	NIH-AARP	7748/548780		1.060 (1.010, 1.120)	9.7%
Strak et al. 2021	ELAPSE	1711/325367		0.746 (0.605, 0.917)	0.8%
Random-effects model				**1.060 (1.040, 1.080)**	**100.0%**
Heterogeneity: $I^2 = 90.2\%$, $\tau^2 = 0.0002$, $p < 0.01$					

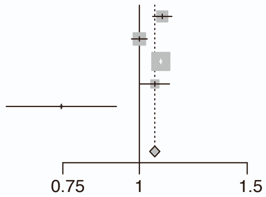
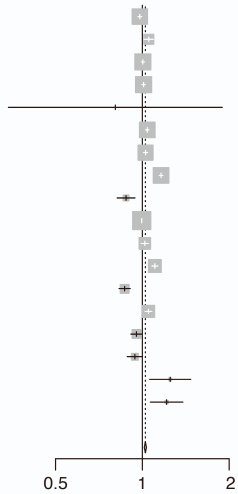


Figure S5 | Multi-study pooled mortality RR of COPD associated with ozone exposure.

Study	Cohort	Cases (n/N)	Risk Ratio	RR (95% CI)	Weight
Jerrett et al. 2009	ACS CPS II	48884/448850		0.980 (0.965, 0.993)	9.1%
Smith et al. 2009	ACS CPS II	-		1.053 (1.014, 1.114)	3.8%
Lipsett et al. 2011	CTS	2919/101784		1.004 (0.991, 1.015)	9.9%
Jerrett et al. 2013	ACS CPS II	8046/73711		1.010 (0.997, 1.022)	10.0%
Bentayeb et al. 2015	GAZEL	165/20327		0.831 (0.397, 1.729)	0.0%
Crouse et al. 2015	CANCHEC	98970/2521525		1.040 (1.025, 1.055)	9.4%
Turner et al. 2016	ACS CPS II	85132/669046		1.026 (1.009, 1.043)	8.7%
Weichenthal et al. 2017	CANCHEC	77000/2448500		1.161 (1.144, 1.178)	9.4%
Hvidtfeldt et al. 2019	DDCH	2319/49596		0.878 (0.817, 0.959)	1.3%
Kazemiparkouhi et al. 2019	Medicare	2333681/22159190		0.997 (0.995, 0.999)	12.7%
Lim et al. 2019	NIH-AARP	39529/548780		1.020 (0.990, 1.030)	5.1%
Paul et al. 2020	ONPHEC	64773/452590		1.105 (1.078, 1.133)	6.3%
Strak et al. 2021	ELAPSE	15542/325367		0.791 (0.734, 0.853)	3.0%
Bauwelinck et al. 2022	BC2001	234549/5474470		1.050 (1.022, 1.076)	5.8%
Stafoggia et al. 2022	ELAPSE	1186101/28153138		0.954 (0.912, 0.996)	2.9%
So et al. 2022	DanNAC	90028/3083227		0.942 (0.886, 0.980)	1.8%
Liu et al. 2022	CHERRY	7308/744882		1.249 (1.060, 1.500)	0.3%
Niu et al. 2022	CCDRFS	2064/96955		1.214 (1.066, 1.383)	0.4%
Random-effects model				**1.024 (1.015, 1.033)**	**100.0%**
Heterogeneity: $I^2 = 97.3\%$, $\tau^2 = 0.0002$, $p < 0.01$					



Study	Cohort	Cases (n/N)	Risk Ratio	RR (95% CI)	Weight
Lipsett et al. 2011	CTS	2919/101784		1.004 (0.991, 1.015)	7.6%
Smith et al. 2009	ACS CPS II	-		1.053 (1.014, 1.114)	3.8%
Bentayeb et al. 2015	GAZEL	165/20327		0.831 (0.397, 1.729)	0.0%
Crouse et al. 2015	CANCHEC	98970/2521525		1.040 (1.025, 1.055)	13.7%
Turner et al. 2016	ACS CPS II	85132/669046		1.026 (1.009, 1.043)	21.7%
Hvidtfeldt et al. 2019	DDCH	2319/49596		0.878 (0.817, 0.959)	1.2%
Kazemiparkouhi et al. 2019	Medicare	2333681/22159190		0.997 (0.995, 0.999)	24.5%
Lim et al. 2019	NIH-AARP	39529/548780		1.020 (0.990, 1.030)	5.6%
Paul et al. 2020	ONPHEC	64773/452590		1.105 (1.078, 1.133)	7.4%
Strak et al. 2021	ELAPSE	15542/325367		0.791 (0.734, 0.853)	2.9%
Bauwelinck et al. 2022	BC2001	234549/5474470		1.050 (1.022, 1.076)	6.5%
Stafoggia et al. 2022	ELAPSE	1186101/28153138		0.954 (0.912, 0.996)	2.8%
So et al. 2022	DanNAC	90028/3083227		0.942 (0.886, 0.980)	1.6%
Liu et al. 2022	CHERRY	7308/744882		1.249 (1.060, 1.500)	0.2%
Niu et al. 2022	CCDRFS	2064/96955		1.214 (1.066, 1.383)	0.4%
Random-effects model				**1.017 (1.009, 1.025)**	**100.0%**
Heterogeneity: $I^2 = 95.1\%$, $\tau^2 < 0.0001$, $p < 0.01$					

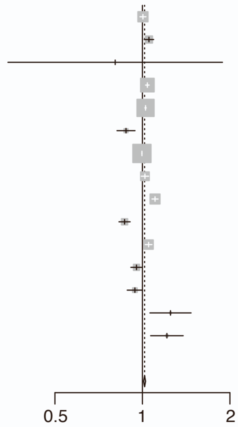


Figure S6 | Multi-study pooled mortality RR of CVDs associated with ozone exposure.

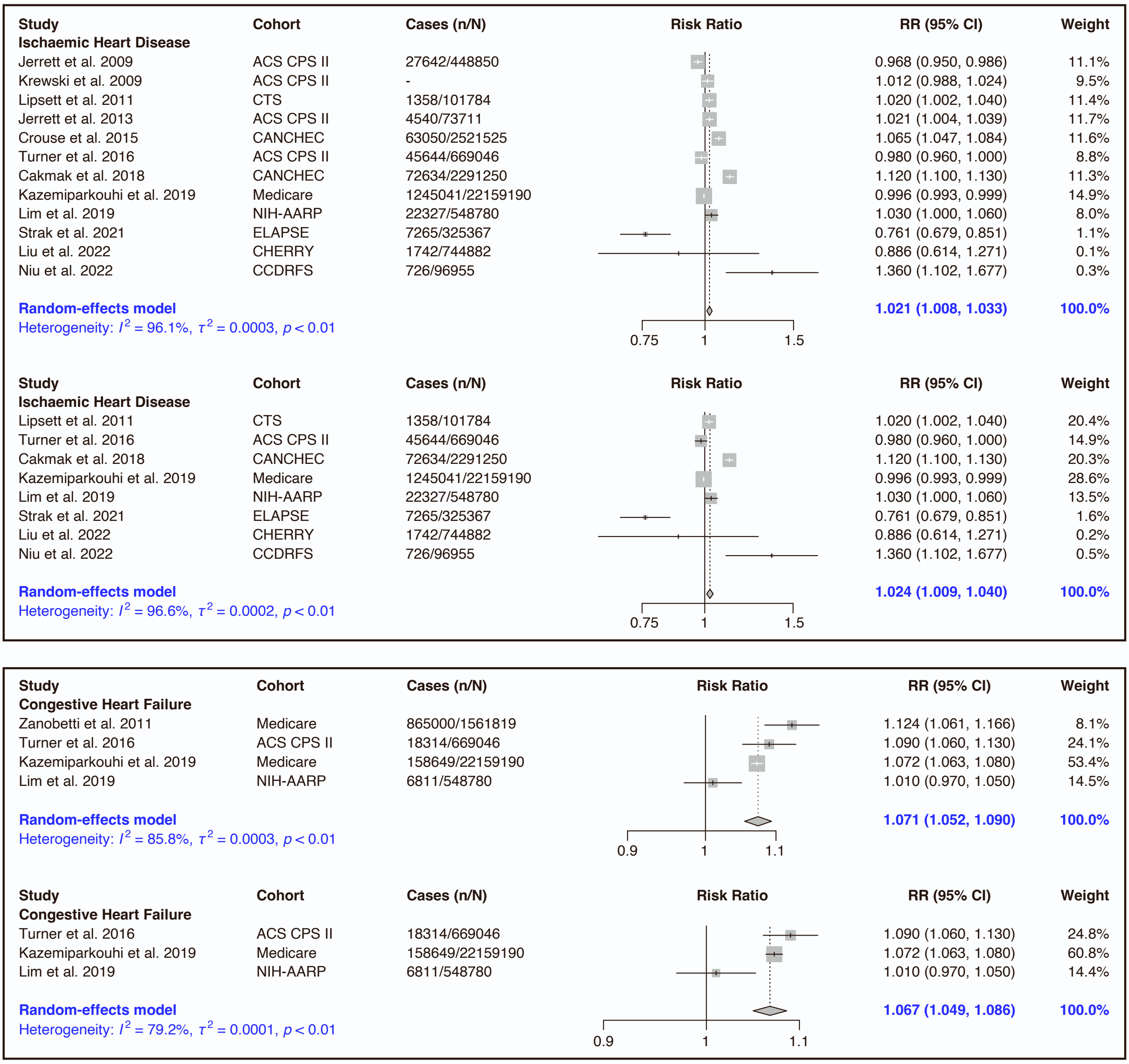


Figure S7 | Multi-study pooled mortality RR of IHD and CHF associated with ozone exposure.

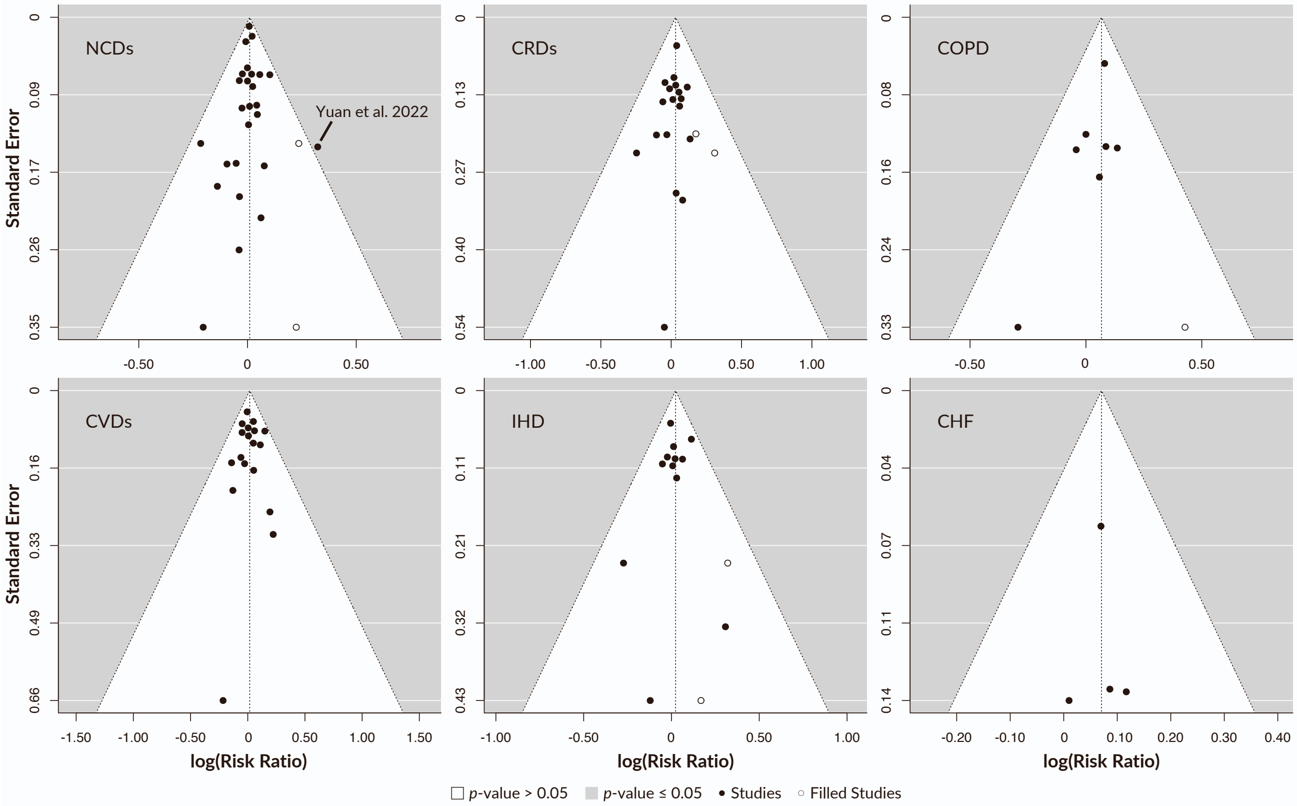


Figure S8 | Examination of publication biases by trim-and-fill method.

Scatter points are jittered appropriately to avoid excessive overlap.

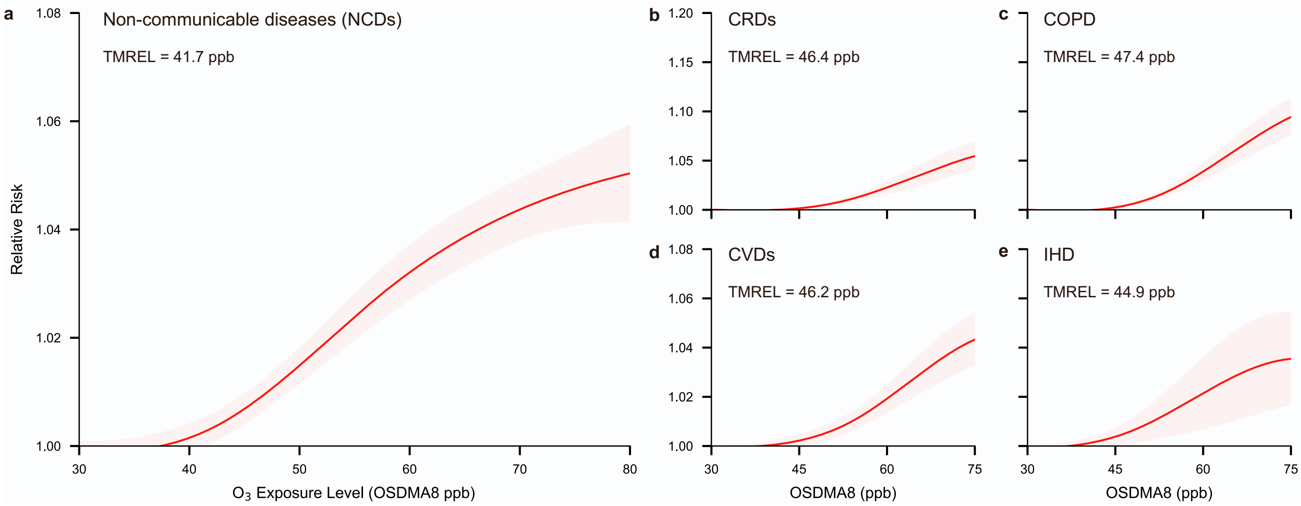


Figure S9 | Multi-study pooled ozone-associated RR curves of multi-cause mortality.

The exposure-response (ER) curves are estimated for (**a**) ozone-associated mortality risks of non-communicable diseases (NCDs), (**b**) chronic respiratory diseases (CRDs), (**c**) chronic obstructive pulmonary disease (COPD), (**d**) cardiovascular diseases (CVDs), and (**e**) ischaemic heart disease (IHD) by mean of exposure range resampled meta-regression, Bayesian, regularised, and trimmed (MR-BRT). Exposures are quantified by 6-month (April–September) ozone-season 8-hour daily maximum average (OSDMA8) metric in ppb. Meta-regressions are performed on censored epidemiological evidence removing studies on duplicated cohort, unless the ER curved are clearly reported in the original literatures. Threshold exposure levels, also known as theoretical minimum risk exposure levels (TMREL), are indicated in each panel. The curved relative risks are used for mortality estimations as main analyses.

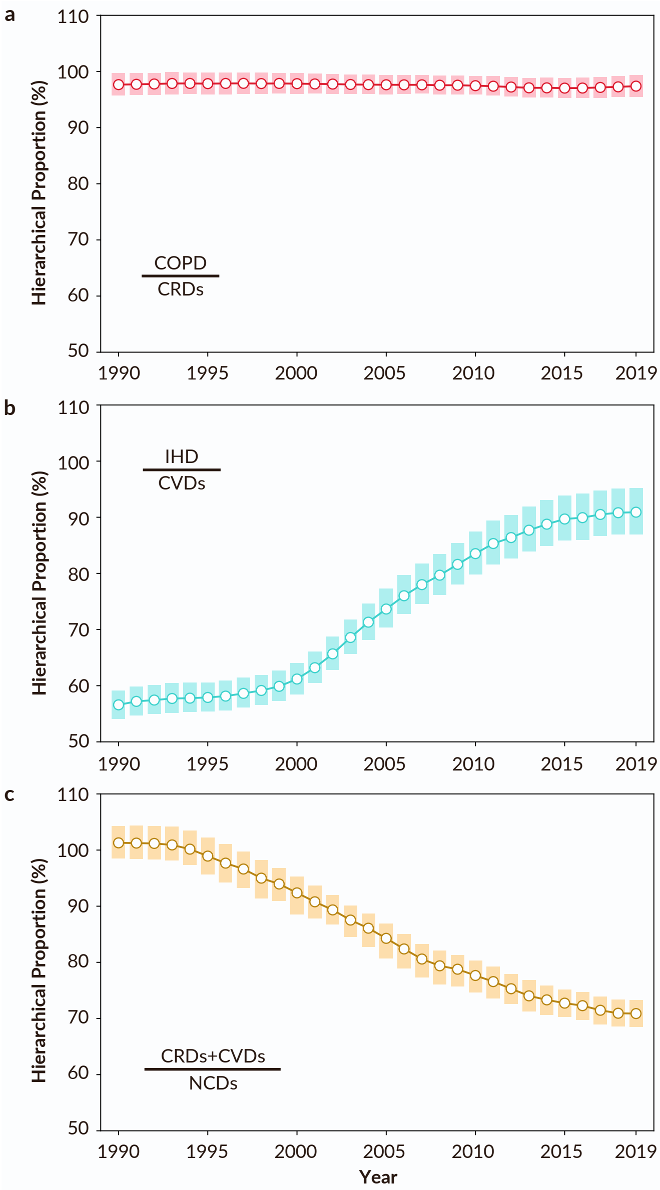


Figure S10 | 30-year trend of hierarchical multi-cause mortality fractions.

Three hierarchical fraction values are calculated, as **a**) chronic obstructive pulmonary disease (COPD) excess deaths out of all chronic respiratory deaths, COPD/CRDs; **b**) ischaemic heart disease (IHD) excess deaths out of all cardiovascular deaths, IHD/CVDs; and **c**) total chronic respiratory and cardiovascular excess deaths out of deaths due to all non-communicable diseases, (CRDs+CVDs)/NCDs. The median values for each year are indicated by dots, with 95% uncertainty intervals presented by shades.

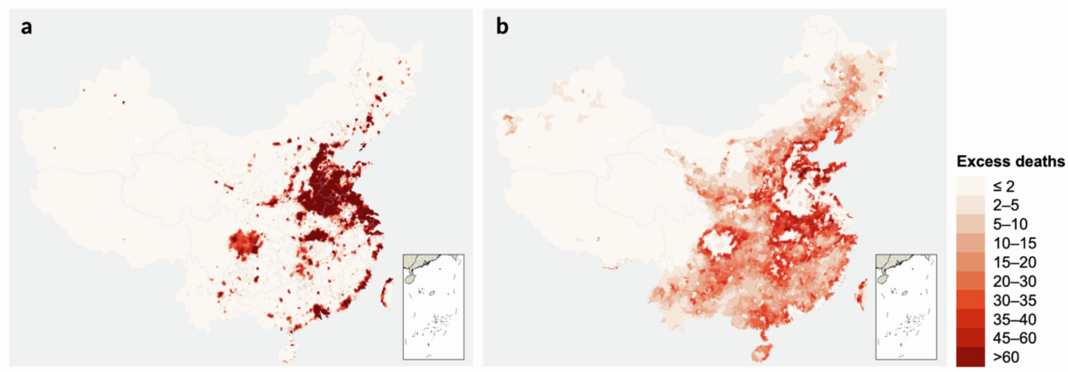


Figure S11 | Gridded mapping of urban and rural cardiopulmonary premature deaths in 2019.

The spatial resolution for grid-specific population ambient O_3 exposure assignment and associated mortality estimation with **(a)** urban and **(b)** rural differentiation is 1/8°×1/8° (approximately 10×10 km^2). Long-term ambient O_3 exposure-associated excess cardiopulmonary premature deaths are defined as the total mortality cases caused from chronic obstructive pulmonary diseases (COPD) and all-type cardiovascular diseases. Intervals of colourbar are defined by Jenks natural breaks.

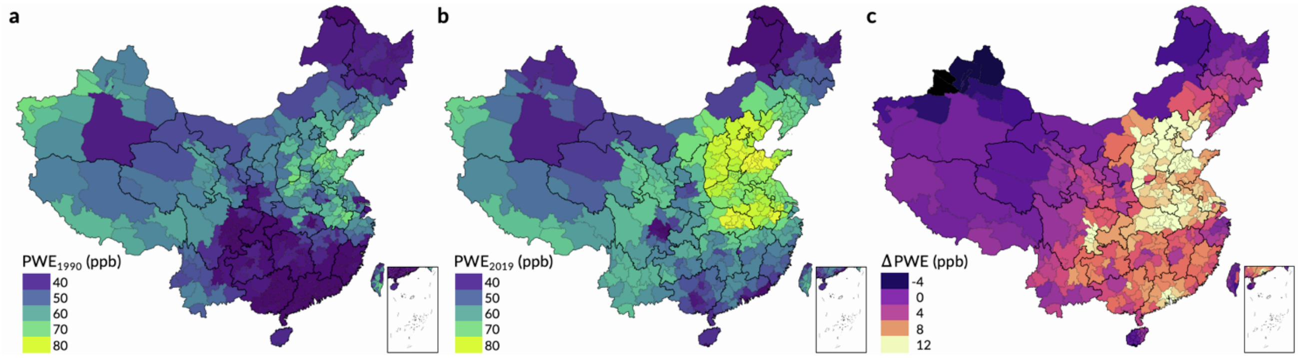


Figure S12 | Changes in population-weighted ozone exposure comparing 1990 with 2019.

Panel **a** and **b** map population-weighted exposure (PWE) concentrations to ambient ozone (ppb) by OSDMA8 metric in year 1990 and 2019, respectively. Panel **c** presents the change of PWE (∆PWE) from 1990 to 2019. Only 2 years of PWE are considered for comparison.

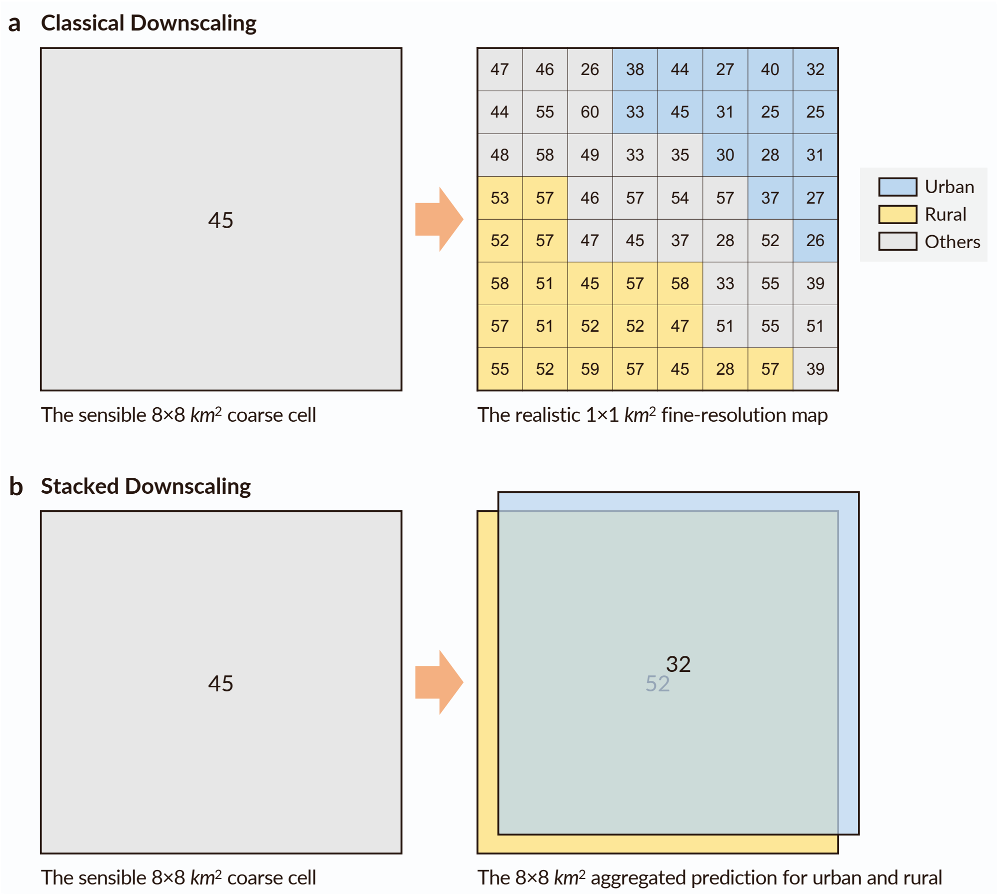


Figure S13 | Schematic diagram of (a) classical high-resolution downscaling and (b) urban-rural differentiated stacked downscaling.

a. Classical downscaling requires predictions precise to target finer resolution (from 45 ppb to 47, 46, 26, ... ppb for each finer cell, a total of 8×8=64 times of predictions), which however is frequently unfeasible in practice due to lack of high-resolution auxiliary datasets as predictors. Note in the diagram, spatial resolution and gridded values are manually faked, simply for illustration purpose. **b.** The left panel presents an 8×8 km^2 coarse cell of which the cell-level ambient O_3 concentrations (like 45 ppb) are sensible as an integrity (e.g. by remote-sensing measurement, model fusion calibrated by deep learning algorithms, etc.) to represent the average level of the whole cell. However, 8×8 km^2 is still a large domain with substantial intra-cell variability in term of ambient O_3, as shown in the right part of panel **a**. Under the circumstance when it is unfeasible to realise higher-resolution downscaling (e.g. 1×1 km^2) but there are multi-site urban- and rural-classified observations inside the studied cell, the urban and rural average ambient O_3 concentrations, 32 and 52 ppb, can be calculated and stacked to the cell, as shown in the right panel. The stacked downscaling only requires two times of predictions, from 45 to 32 ppb for urban concentration, and from 45 to 52 ppb for rural concentration. Note in the diagram, spatial resolution and gridded values are manually faked, aiming at illustrative presentations.

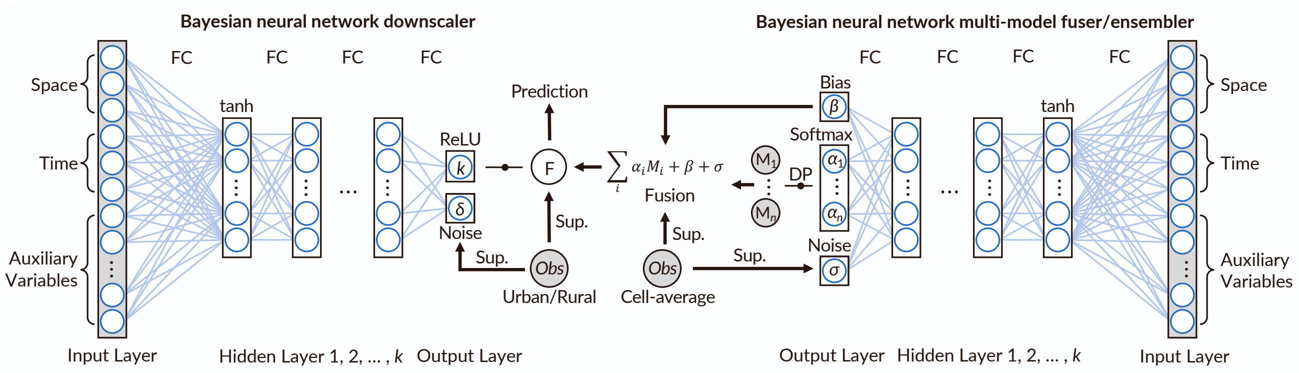


Figure S14 | Schematic diagram of Bayesian neural network multi-model fuser and downscaler.

Right part demonstrates deep-learning-based multi-model fuser, and left part depicts urban-rural downscaler. The shaded elements refer to the external datasets not affected by neural network; the rectangle circumscribed elements indicate the input, processing and output variates inside the neural network; and non-rectangle circumscribed elements represent the final products. The schematic diagram is appropriately modified from a publication[2] with full consents from American Chemical Society Publications and involved authors.

Abbreviations and denotations: FC, *fully connected*; Sup., *supervised training*; DP, *dot product*; F, *multi-model fused output*; Obs, *observations*; ReLU, *rectified linear unit*; M, *calibrated CMIP6 models*; Softmax, *normalised exponential function*; tanh, *hyperbolic tangent function*.

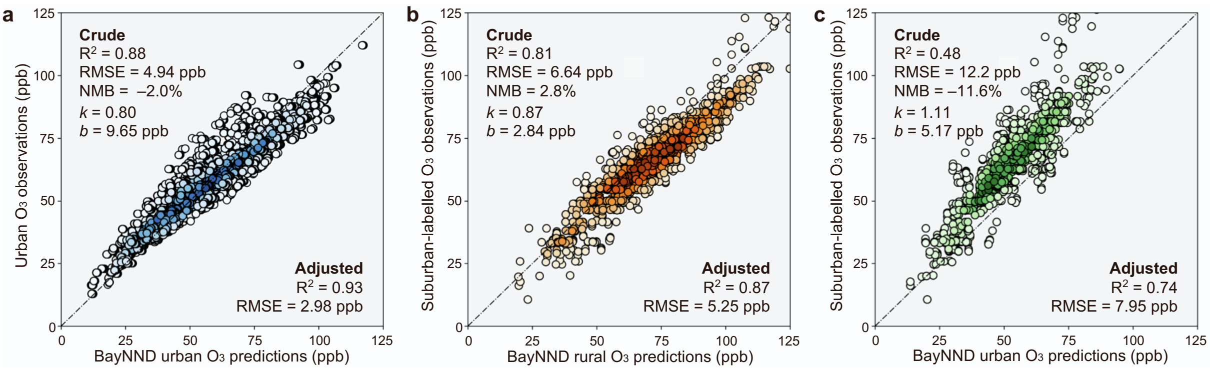


Figure S15 | Extrapolation validations on Chinese *in situ* observations with (a) urban, (b) rural, and (c) suburban differentiation by metric of monthly average of daily 8-hour maximum.

Prediction-observation extrapolation evaluations span from May 2014 to December 2019, including statistics of coefficient of determination (R^2), root-mean-square error (RMSE, ppb), normalised mean bias (NMB, %, defined as difference that prediction minus observation proportion to observation), linear regression slope (k) and intercept (b). No Chinese *in situ* observations are included for Bayesian neural network framework training; predictions for urban and rural ambient O_3 in China are results of spatial extrapolation. Crude evaluations are performed on the observations and raw predictions by BayNND, and adjusted evaluations on the observations and 1:1-linearly calibrated predictions by BayNND. Adjusted evaluations are all of fixed NMB =0%, slope (k=1), and intercept (b=0). Panel (**b**) evaluates the coherence between "suburban"-labelled observations and rural O_3 predictions, and (**c**) evaluates the consistency between "suburban"-labelled observations and urban O_3 predictions. Data-based evidence reveals the "suburban"-labelled ambient O_3 concentrations are closer to rural than urban pattern.

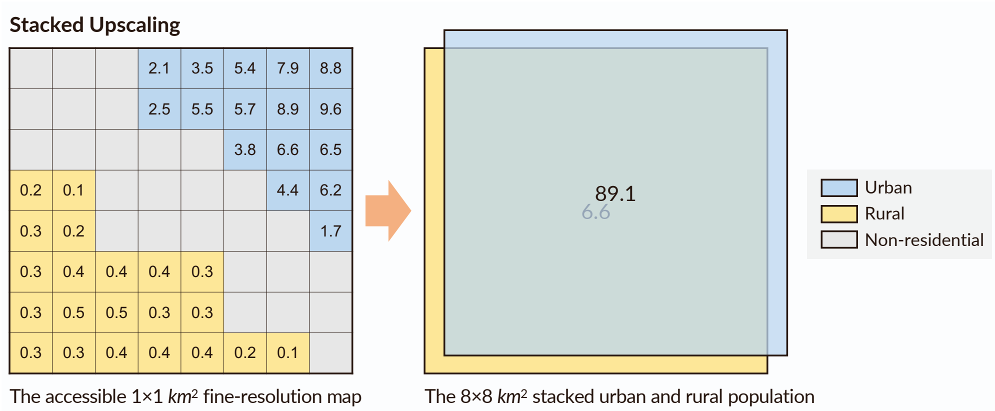


Figure S16 | Schematic diagram of urban-rural stacked gridded population upscaling.

The left panel presents 1×1 km^2 higher-resolution population (in thousand) distribution in a target coarser 8×8 km^2 cell, in which urban and rural populations are defined based on population density. The upscaling process sums up the total finely gridded populations separately for urban and rural regions, and stacked the total urban population count 89,100 and rural population count 6,600 into the upscale coarse cell, as shown in the right panel. In further analyses, it will only be considered the upscaled cell-level total urban and rural populations (i.e. 89,100 and 6,600), rather than how the residents are spatially distributed (i.e. 2,100, 3,500, etc.). The populations scaled in coarse cell will be linked with ambient O_3 in same spatial resolution. Note in the diagram, spatial resolution and gridded values are manually faked, aiming at illustrative presentations.

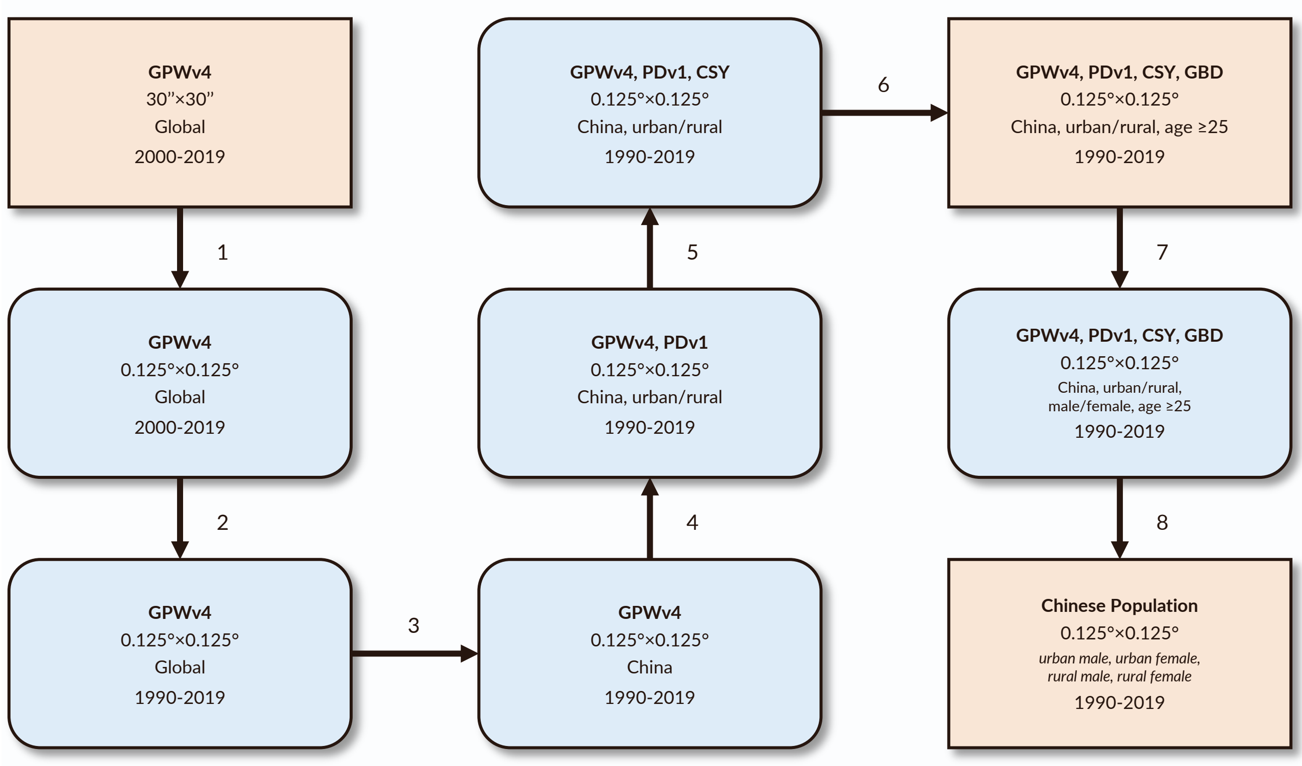


Figure S17 | Flowchart of gridded population dataset construction and calibration.

Rounded rectangles represent procedural data products; two rectangles refer to the initial input and final output datasets; number-marked arrows note manual operations for database development. Spatial resolution, space-time coverage, and population features are indicated in each dataset.

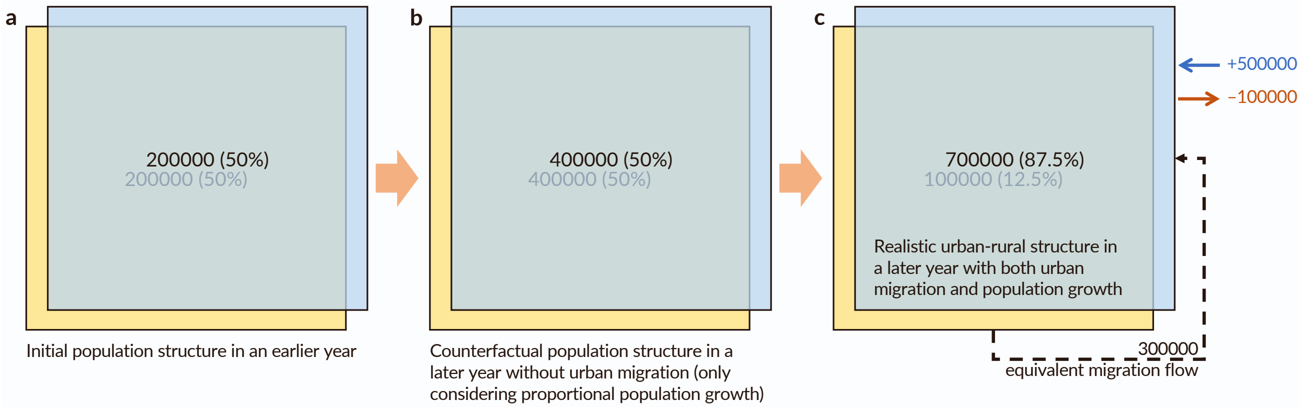


Figure S18 | Schematic diagram of cross-sectional population migration at cell-level definition.

Panel (**a**) represents the initial population structure in an earlier year, when urban and rural populations are both 200,000. Panel (**b**) indicates a counterfactual scenario in a later year, that only population growth occurs without any urban-rural population structure change. The cell-level total population doubles from 400,000 to 800,000, among which urban and rural populations increase proportionally to 400,000. Panel (**c**) reflects the realistic population structure in the later year, when urban population is 700,000 and rural population is 100,000. Directly comparing the realistic situation (**a** and **c**), urban population expands by 500,000 and rural population shrinks by 100,000, which is affected both by population growth and migration. Adjusting the effect from population growth assuming urban and rural populations are of the same growing rate, the population migration flow can be equivalently perceived as 300,000 rural population inside the studied cell migrate to the urban environments in the same cell (comparing **b** and **c**), so that rural population can be perceived as 400,000−300,000=100,000, and urban population as 400,000+300,000=700,000.

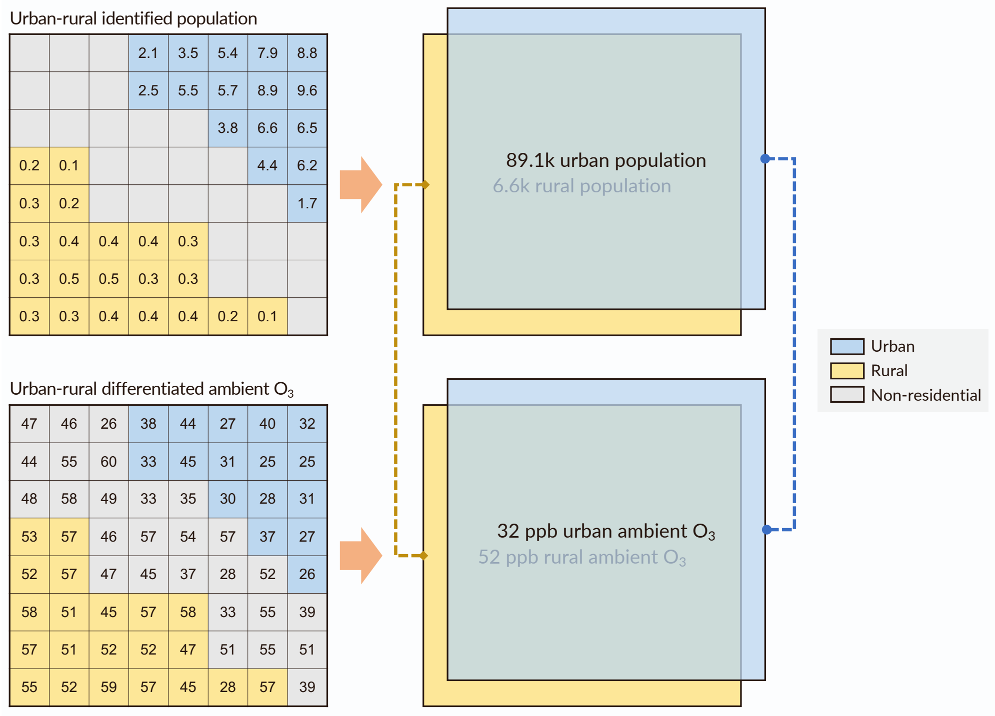


Figure S19 | Schematic diagram of cell-level population exposure assignment in stacked context.

The upper part presents upscaling of stacked urban-rural population, and the lower part shows the urban-rural differentiation of ambient O_3 concentrations. The right part demonstrates how urban (or rural) populations are linked to urban (or rural) ambient O_3 exposure in the stacked context, as 89,100 urban population are exposed to 32 ppb O_3 on average, and 6,600 rural population are exposed to 52 ppb O_3.

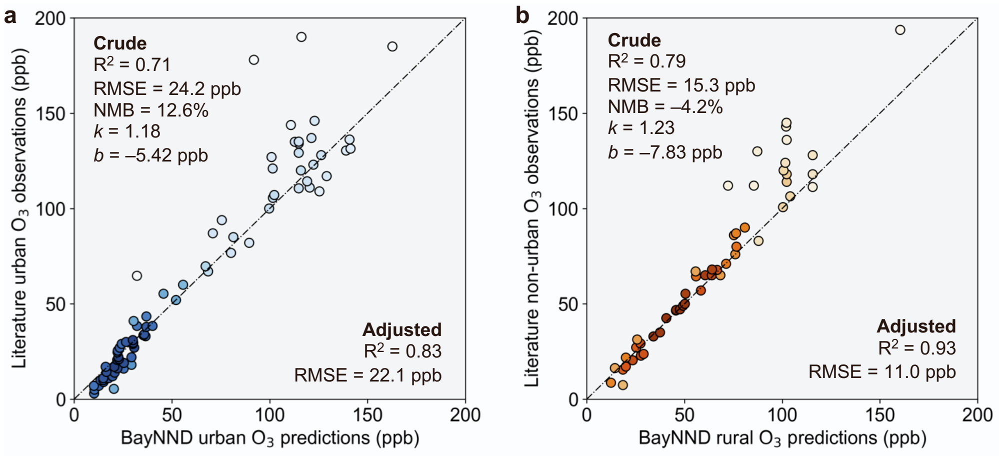


Figure S20 | External ozone prediction validations with literature reported observations.

Enhanced external evaluations beyond CNEMC span from October 1993 to December 2019, including statistics of coefficient of determination (R^2), root-mean-square error (RMSE), normalised mean bias (NMB), linear regression slope (k) and intercept (b). Only point-to-point evaluations are performed, excluding literatures only reporting concentration ranges. All available metrics in monthly smoothed values are included with necessary cross-metric conversion. When multiple metrics are provided in literature, the daily 24-h average and diurnal maximum 8-h average are preferred. Crude evaluations are performed on the observations and raw predictions by BayNND, and adjusted evaluations on the observations and 1:1-linearly calibrated predictions by BayNND. Adjusted evaluations are all of fixed NMB = 0%, slope (k = 1), and intercept (b = 0). Full information can be found at Content S2.

SUPPLEMENTARY CONTENTS

Content S1 | Population density of "suburban"-labelled CNEMC observation stations in 2019.

A total of 245 "suburban"-labelled CNEMC stations are projected to gridded population (see "*Population gridding and calibration*" section in *Methods*). Planar cell (approximated as rectangles) areas are calculated by planar meridional distance multiplied by planar parallel distance, where meridional (m) and parallel (p) distance follow the two formulae below, where R is the average Earth radius, 6378.137 km. Population densities are calculated by cell-specific total population divided by cell area. Urban locations (U) are categorised by population density >1,500 people per km^2 (C1), and more conservatively, an additional urban categorisation by population density threshold >1,000 people per km^2 (C2) is provided as a sensitivity analysis. By C1, 242 out of 245 sites are classified as rural (R); by C2, 232 out sites are classified as rural, indicating "suburban"-labelled sites are more of rural sociodemographic characteristics.

Station	Longitude (°E)	Latitude (°N)	Pop. Des.	C1	C2	Station	Longitude (°E)	Latitude (°N)	Pop. Des.	C1	C2
1002A	116.220	40.292	503	R	R	1855A	111.675	29.024	205	R	R
1013A	117.151	39.097	1543	U	U	1856A	111.679	29.038	205	R	R
1014A	117.193	39.173	2175	U	U	1857A	111.716	29.146	259	R	R
1016A	117.184	39.121	1543	U	U	1861A	110.442	29.315	141	R	R
1020A	117.269	39.134	1483	R	U	1862A	110.414	25.317	258	R	R
1025A	117.401	39.124	568	R	R	1866A	109.226	21.588	232	R	R
1027A	117.157	38.919	296	R	R	1882A	104.563	28.793	265	R	R
1028A	114.564	38.055	1312	R	U	1887A	104.679	28.799	277	R	R
1035A	114.352	37.891	286	R	R	1893A	105.432	28.963	262	R	R
1036A	118.166	39.631	539	R	R	1897A	104.755	29.363	590	R	R
1039A	118.219	39.668	540	R	R	1905A	106.056	30.806	563	R	R
1058A	114.892	40.795	328	R	R	1918A	108.720	34.396	776	R	R
1061A	114.892	40.866	328	R	R	1921A	108.737	34.316	559	R	R
1066A	117.927	41.003	183	R	R	1922A	109.066	35.099	196	R	R
1069A	116.715	39.557	358	R	R	1926A	109.413	36.628	85	R	R
1076A	115.691	37.739	263	R	R	1930A	107.185	34.306	239	R	R
1082A	112.573	37.910	1053	R	U	1938A	109.529	34.510	360	R	R
1083A	112.434	38.011	488	R	R	1942A	102.188	38.525	77	R	R
1092A	111.659	40.845	247	R	R	1947A	106.339	38.817	114	R	R
1093A	111.608	40.814	173	R	R	2073A	117.721	24.509	501	R	R
1098A	123.684	41.934	991	R	R	2074A	117.657	24.516	501	R	R
1101A	123.284	41.769	1232	R	U	2075A	117.634	24.467	260	R	R
1107A	123.361	41.781	1232	R	U	2165A	112.845	35.546	300	R	R
1125A	125.719	43.515	408	R	R	2177A	111.040	35.039	183	R	R
1129A	126.542	45.755	1430	R	U	2180A	112.736	38.419	144	R	R
1146A	120.978	31.094	585	R	R	2189A	122.260	43.627	100	R	R
1160A	120.561	31.247	398	R	R	2190A	122.304	43.616	70	R	R
1174A	119.141	34.590	599	R	R	2193A	119.728	49.201	24	R	R
1175A	119.368	34.751	26	R	R	2194A	107.594	40.916	89	R	R
1176A	119.348	34.698	300	R	R	2199A	122.062	46.087	106	R	R
1179A	117.192	34.308	933	R	R	2204A	105.647	38.836	4	R	R
1183A	117.166	34.181	801	R	R	2224A	124.342	43.175	277	R	R
1184A	119.460	32.388	751	R	R	2228A	125.157	42.895	248	R	R
1185A	119.404	32.410	751	R	R	2241A	130.982	45.305	195	R	R
1187A	119.439	32.403	751	R	R	2242A	131.010	45.295	132	R	R
1197A	119.933	31.779	885	R	R	2247A	130.110	47.338	74	R	R
1198A	119.962	31.809	885	R	R	2251A	131.120	46.566	132	R	R
1199A	120.039	31.764	1076	R	U	2254A	129.503	48.471	45	R	R
1208A	119.882	32.303	593	R	R	2259A	130.379	46.759	159	R	R
1217A	120.129	33.372	324	R	R	2260A	131.003	45.768	152	R	R
1219A	118.266	33.960	325	R	R	2261A	130.863	45.819	78	R	R
1222A	118.321	33.951	325	R	R	2262A	131.052	45.875	148	R	R
1225A	119.026	29.635	75	R	R	2265A	127.529	50.247	79	R	R
1237A	121.554	29.891	349	R	R	2268A	124.119	50.427	6	R	R
1238A	121.615	29.902	349	R	R	2272A	117.309	32.935	436	R	R
1248A	120.576	30.007	442	R	R	2276A	117.042	32.661	336	R	R
1249A	120.100	30.887	233	R	R	2281A	116.633	32.620	253	R	R
1251A	120.093	30.862	352	R	R	2293A	117.049	30.549	220	R	R
1273A	117.160	31.905	722	R	R	2300A	118.316	32.306	199	R	R
1280A	119.390	26.054	689	R	R	2304A	116.977	33.648	449	R	R
1286A	118.161	24.817	672	R	R	2305A	116.968	33.628	449	R	R
1293A	115.973	28.697	660	R	R	2313A	117.481	30.641	173	R	R
1296A	115.742	28.800	425	R	R	2315A	117.497	30.654	173	R	R
1297A	115.912	28.613	413	R	R	2323A	118.981	25.479	441	R	R
1302A	116.989	36.687	757	R	R	2327A	117.728	26.311	143	R	R
1307A	120.666	36.240	44	R	R	2331A	118.097	26.676	120	R	R
1324A	113.515	34.911	1071	R	U	2335A	117.019	25.118	103	R	R
1334A	113.845	30.292	253	R	R	2339A	119.500	26.695	131	R	R
1344A	112.958	28.361	878	R	R	2340A	119.520	26.661	131	R	R
1355A	113.443	23.304	1192	R	U	2342A	117.310	29.387	120	R	R

Station	Longitude (°E)	Latitude (°N)	Pop. Des.	C1	C2	Station	Longitude (°E)	Latitude (°N)	Pop. Des.	C1	C2
1382A	113.441	22.485	792	R	R	2347A	114.100	27.500	152	R	R
1400A	112.475	23.100	178	R	R	2352A	115.086	27.932	237	R	R
1405A	108.439	22.790	177	R	R	2356A	114.912	27.804	209	R	R
1409A	110.576	19.951	133	R	R	2357A	116.982	28.114	174	R	R
1414A	106.379	29.828	703	R	R	2362A	114.902	25.915	234	R	R
1415A	106.460	29.574	923	R	R	2371A	114.341	27.806	271	R	R
1423A	106.571	29.564	993	R	R	2376A	116.213	28.081	198	R	R
1424A	106.512	29.516	993	R	R	2381A	118.005	28.457	267	R	R
1430A	106.591	29.427	591	R	R	2384A	117.903	28.430	239	R	R
1438A	103.620	31.020	159	R	R	2393A	114.991	35.767	596	R	R
1449A	102.743	25.012	462	R	R	2411A	115.658	34.402	629	R	R
1451A	102.625	24.961	264	R	R	2428A	111.042	32.395	89	R	R
1472A	108.869	34.378	1121	R	U	2442A	112.193	31.038	164	R	R
1477A	104.137	35.945	124	R	R	2445A	114.886	30.452	552	R	R
1481A	101.749	36.692	316	R	R	2447A	114.318	29.814	240	R	R
1483A	101.524	36.687	214	R	R	2456A	112.500	26.917	446	R	R
1484A	105.951	38.602	150	R	R	2462A	111.524	27.303	402	R	R
1485A	106.268	38.474	150	R	R	2467A	112.407	28.643	310	R	R
1486A	106.217	38.454	162	R	R	2472A	113.007	25.906	213	R	R
1487A	106.072	38.486	161	R	R	2477A	111.622	26.208	193	R	R
1492A	87.475	43.947	750	R	R	2482A	109.598	27.572	163	R	R
1552A	106.805	26.300	157	R	R	2487A	111.959	27.890	292	R	R
1559A	113.251	27.834	303	R	R	2492A	109.641	28.256	215	R	R
1564A	112.488	27.916	349	R	R	2505A	109.568	23.148	234	R	R
1586A	109.810	40.658	331	R	R	2509A	110.111	22.702	275	R	R
1614A	118.612	24.960	915	R	R	2516A	108.201	24.715	88	R	R
1615A	117.685	36.205	493	R	R	2523A	105.895	32.454	88	R	R
1617A	117.715	36.208	493	R	R	2527A	105.545	30.568	552	R	R
1626A	115.997	36.457	570	R	R	2535A	103.772	29.546	330	R	R
1627A	115.984	36.480	570	R	R	2539A	104.031	30.048	327	R	R
1647A	121.595	37.387	245	R	R	2543A	106.631	30.528	579	R	R
1649A	119.092	36.731	556	R	R	2546A	106.641	30.484	409	R	R
1651A	119.161	36.657	565	R	R	2548A	107.528	31.283	250	R	R
1654A	116.586	35.414	566	R	R	2553A	103.009	30.013	211	R	R
1667A	118.586	37.444	196	R	R	2557A	106.758	31.848	389	R	R
1668A	118.819	37.378	174	R	R	2561A	104.662	30.137	513	R	R
1696A	114.678	23.757	111	R	R	2566A	102.188	31.914	6	R	R
1699A	111.980	21.859	161	R	R	2571A	102.343	27.810	131	R	R
1702A	113.043	23.691	444	R	R	2576A	104.800	26.589	161	R	R
1705A	116.637	23.672	830	R	R	2598A	100.214	26.858	49	R	R
1712A	112.039	22.917	142	R	R	2617A	98.578	24.441	93	R	R
1721A	113.382	40.110	160	R	R	2623A	97.181	31.125	6	R	R
1722A	113.286	40.096	261	R	R	2627A	88.893	29.237	19	R	R
1729A	113.147	36.196	314	R	R	2631A	80.090	32.504	1	R	R
1732A	111.513	36.098	238	R	R	2634A	106.989	33.184	169	R	R
1737A	111.492	36.042	337	R	R	2638A	109.741	38.334	55	R	R
1754A	123.129	41.023	461	R	R	2642A	109.032	32.654	127	R	R
1778A	126.706	43.713	137	R	R	2649A	106.006	34.343	126	R	R
1782A	123.626	47.203	132	R	R	2653A	102.647	37.936	105	R	R
1797A	118.402	31.384	525	R	R	2660A	107.683	35.729	226	R	R
1802A	118.625	31.724	527	R	R	2665A	105.082	33.326	59	R	R
1810A	115.977	29.570	276	R	R	2680A	105.003	37.464	51	R	R
1818A	114.484	36.062	744	R	R	2683A	106.232	36.142	110	R	R
1821A	114.393	36.088	745	R	R	2690A	88.124	43.889	18	R	R
1822A	114.286	36.110	839	R	R	2701A	79.949	37.115	52	R	R
1823A	114.341	34.802	523	R	R	2702A	79.912	37.101	52	R	R
1825A	114.373	34.798	523	R	R	2707A	88.121	47.905	15	R	R
1830A	113.199	35.270	683	R	R	2874A	117.041	32.646	336	R	R
1831A	113.306	33.721	663	R	R	2914A	106.768	31.879	340	R	R
1838A	111.143	34.796	188	R	R	2916A	117.490	30.660	173	R	R
1846A	112.289	30.306	225	R	R	2923A	113.280	40.111	261	R	R
1852A	113.212	29.355	186	R	R	3122A	105.961	26.261	319	R	R
1853A	111.704	29.024	205	R	R						

Note: $m = 2\,sin^{-1}\sqrt{cos\left(\left(lat - \frac{1}{16}\right)\cdot\frac{\pi}{180}\right)\cdot cos\left(\left(lat + \frac{1}{16}\right)\cdot\frac{\pi}{180}\right)\cdot sin^2\left(\frac{1}{16}\cdot\frac{\pi}{180}\right)} \times R$, $p = 2\,sin^{-1}\sqrt{sin^2\left(\frac{1}{16}\cdot\frac{\pi}{180}\right)} \times R$.

Content S2 | Literature-based external validations of urban-rural ambient ozone predictions.

Accuracy evaluations on CNEMC observations are limited to the latest six years (2014–2019). To check the reliability of 30-yr deep-learning-based prediction, totally 68 peer-reviewed studies reporting *in situ* observations of ambient O_3 are collected for enhanced model-observation comparison. The developed ambient O_3 database covers two metrics as i) monthly average of daily 24-h average, and ii) monthly average of daily maximum 8-h average. The metric, daily diurnal 7-h average, adopted in earlier literatures, are compared to daily maximum 8-h average as an alternative proxy. For prediction-observation comparisons on daily 1-h maximum metric, null-intercept linear conversion is applied to approximately project daily 8-h maximum average (DMA8h) concentrations onto daily 1-h maximum average (DMA1h) concentrations. The idea of null-intercept linear conversion was put forward by US EPA (Volume I, section 7.1.3.2)[78], and the conversion coefficients have been updated by 30-yr historical observations archived in TOAR and CNEMC[14]. At multi-season or multi-year scale, the conversion follows: DMA1h = DMA8h × 1.213; in warm seasons (i.e. April to September), the conversion follows: DMA1h = DMA8h×1.202, where O_3 concentrations in DMA8h metric are obtained from Bayesian neural network downscaler. Observed and deep-learning-modelled ambient O_3 concentrations are both unified into ppb. IGAC (International Global Atmospheric Chemistry project) TOAR-II Working Group has double-checked the external validation in August 2022, and recognised the credibility of the database for long-term population exposure tracking and risk assessment studies (https://igacproject.org/human-health-impacts-ozone-focus-working-group, accessed February 2023).

Site location	Longitude (°E)	Latitude (°N)	Period start	Period end	Metric	Type	Observed	Modelled	Refs
Chongqing	106.5	29.6	Oct-93		Period 24-h average	Urban	7	12.6	79
Chongqing	106.5	29.6	Oct-93		Daily 7-h average	Urban	12	19.6	79
Chongqing	106.5	29.6	Nov-93		Period 24-h average	Urban	10	13.5	79
Chongqing	106.5	29.6	Nov-93		Daily 7-h average	Urban	16	25.3	79
Chongqing	106.5	29.6	Dec-93		Period 24-h average	Urban	3	10.2	79
Chongqing	106.5	29.6	Dec-93		Daily 7-h average	Urban	7	10.1	79
Chongqing	106.5	29.6	Jan-94		Period 24-h average	Urban	5	10.4	79
Chongqing	106.5	29.6	Jan-94		Daily 7-h average	Urban	11	17.0	79
Chongqing	106.5	29.6	Feb-94		Period 24-h average	Urban	9	15.0	79
Chongqing	106.5	29.6	Feb-94		Daily 7-h average	Urban	17	20.5	79
Chongqing	106.5	29.6	Mar-94		Period 24-h average	Urban	11	17.6	79
Chongqing	106.5	29.6	Mar-94		Daily 7-h average	Urban	19	25.4	79
Hong Kong SAR	114.0	22.0	May-94		Period 24-h average	Urban	33	36.5	80
Hong Kong SAR	114.0	22.0	Jul-94		Period 24-h average	Urban	21	22.9	80
Lin'an, Zhejiang	119.7	30.4	Aug-94	Jul-95	Period maximum 1-h	Rural	120	100.7	81
Waliguan, Qinghai	100.9	36.3	Aug-94	Jul-95	Period maximum 1-h	Rural	130	87.3	81
Shazikou, Shandong	120.5	36.1	Aug-94	Jul-95	Period maximum 1-h	Rural	90	80.9	81
Longfengshan, Heilongjiang	127.6	44.7	Aug-94	Jul-95	Period maximum 1-h	Rural	80	76.6	81
Waliguan, Qinghai	100.9	36.3	Aug-94	Dec-13	Period 24-h average	Rural	65	60.6	82
Hong Kong SAR	114.0	22.0	Sep-94		Period 24-h average	Urban	52	52.1	80
Hong Kong SAR	114.0	22.0	Oct-94		Period 24-h average	Urban	60	55.7	80
Hong Kong SAR	114.2	22.3	Oct-94	Nov-94	Period 24-h average	Urban	53±13	53.3	83
Hong Kong SAR	114.0	22.3	Oct-94	Nov-94	Period 24-h average	Urban	69±23	52.4	83
Longfengshan, Heilongjiang	127.6	44.7	Oct-94	Jan-95	Period maximum 1-h	Rural	86	75.3	84
Lin'an, Zhejiang	119.7	30.4	Oct-94	Jan-95	Period maximum 1-h	Rural	112	72.3	84
Hong Kong SAR	114.3	22.2	Oct-94	Jan-95	Period maximum 1-h	Urban	87	70.9	84
Qingdao, Shandong	120.5	36.1	Oct-94	Jan-95	Period maximum 1-h	Urban	67	68.5	84
Mt Waliguan, Qinghai	100.9	36.3	Jan-95	Dec-18	Annual average	Rural	47–56	53.7–57.9	85
Beijing	117.1	40.7	Jan-95	Dec-18	Annual average	Rural	33–46	35.0–53.3	85
Lin'an, Zhejiang	119.7	30.4	Jan-95	Dec-18	Annual average	Rural	30–35	28.6–36.3	85
Chongqing	106.5	29.6	Jun-95		Period 24-h average	Urban	11	14.9	79
Chongqing	106.5	29.6	Jun-95		Daily 7-h average	Urban	22	29.3	79
Chongqing	106.5	29.6	Jul-95		Period 24-h average	Urban	10	14.9	79
Chongqing	106.5	29.6	Jul-95		Daily 7-h average	Urban	18	29.3	79
Chongqing	106.5	29.6	Aug-95		Period 24-h average	Urban	17	16.1	79
Chongqing	106.5	29.6	Aug-95		Daily 7-h average	Urban	27	30.7	79
Chongqing	106.5	29.6	Jun-96		Period 24-h average	Urban	29	24.3	79
Chongqing	106.5	29.6	Jun-96		Daily 7-h average	Urban	41	30.4	79
Chongqing	106.5	29.6	Jul-96		Period 24-h average	Urban	27	23.3	79
Chongqing	106.5	29.6	Jul-96		Daily 7-h average	Urban	30	26.5	79
Chongqing	106.5	29.6	Aug-96		Period 24-h average	Urban	31	30.0	79
Chongqing	106.5	29.6	Aug-96		Daily 7-h average	Urban	34	35.8	79
Lin'an, Zhejiang	119.7	30.4	Sep-99		Period maximum 1-h	Rural	136	102.3	86
Lin'an, Zhejiang	119.7	30.4	Oct-99		Period maximum 1-h	Rural	112	85.5	86
Lin'an, Zhejiang	119.7	30.4	Nov-99		Period maximum 1-h	Rural	87	76.4	86
Lin'an, Zhejiang	119.7	30.4	Dec-99		Period maximum 1-h	Rural	68	64.1	86
Nanjing, Jiangsu	118.7	32.1	Jan-00	Feb-03	Period 24-h average	Urban	20.4±18.3	20.9	87
Lin'an, Zhejiang	119.7	30.4	Jan-00		Period maximum 1-h	Rural	65	64.0	86
Lin'an, Zhejiang	119.7	30.4	Feb-00		Period maximum 1-h	Rural	71	71.3	86
Lin'an, Zhejiang	119.7	30.4	Mar-00		Period maximum 1-h	Rural	76	76.0	86
Lin'an, Zhejiang	119.7	30.4	Apr-00		Period maximum 1-h	Rural	83	87.9	86
Lin'an, Zhejiang	119.7	30.4	May-00		Monthly average	Rural	57	58.4	88
Lin'an, Zhejiang	119.7	30.4	May-00		Period maximum 1-h	Rural	124	101.6	86
Lin'an, Zhejiang	119.7	30.4	Jun-00		Period maximum 1-h	Rural	118	102.3	86
Lin'an, Zhejiang	119.7	30.4	Jul-00		Period maximum 1-h	Rural	145	102.2	86
Nanjing, Jiangsu	118.7	32.1	2000–2003	Spring	Monthly average	Urban	27±20.6	27.3	87
Nanjing, Jiangsu	118.7	32.1	2000–2003	Summer	Monthly average	Urban	22.8±19.4	21.8	87
Nanjing, Jiangsu	118.7	32.1	2000–2003	Autumn	Monthly average	Urban	18.4±16.7	19.8	87
Nanjing, Jiangsu	118.7	32.1	2000–2003	Winter	Monthly average	Urban	14.1±12.9	17.6	87
Shanghai	121.5	31.2	Jan-01	Jan-04	DMA8h	Urban	32.3±18.7	35.5	89

Site location	Longitude (°E)	Latitude (°N)	Period start	Period end	Metric	Type	Observed	Modelled	Refs
Lin'an, Zhejiang	119.7	30.4	Feb-01	Apr-01	Period 24-h average	Rural	34±18	32.8	90
Mt Tai, Shandong	117.1	36.3	Jul-03	Nov-03	Period 24-h average	NA	58±16	58.6	91
Beijing	117.1	40.7	Sep-03	Dec-03	Period 24-h average	Rural	26.8±27.7	25.9	92
Jinan, Shandong	117.1	36.7	2003	Spring	Period 24-h average	Urban	38.4	40.1	93
Jinan, Shandong	117.1	36.7	2003	Summer	Period 24-h average	Urban	43.4	36.9	93
Jinan, Shandong	117.1	36.7	2003	Autumn	Period 24-h average	Urban	22.1	22.6	93
Jinan, Shandong	117.1	36.7	2003	Winter	Period 24-h average	Urban	14.3	16.5	93
Jinan, Shandong	117.1	36.7	2003	Summer	Period 24-h median	Urban	37.9	36.9	93
Mt Waliguan, Qinghai	100.9	36.3	2003	Spring	Period 24-h average	Rural	58±9	59.4	94
Mt Waliguan, Qinghai	100.9	36.3	2003	Summer	Period 24-h average	Rural	54±11	52.0	94
Beijing	117.1	40.7	Jan-04	Dec-04	Period 24-h average	Rural	30.1±26.7	28.5	92
Shanghai	121.5	31.2	Mar-04	Dec-05	DMA8h	Urban	39.3±1.5	38.5	95
Jinan, Shandong	117.1	36.7	Apr-04		Period maximum 1-h	Urban	105.6	101.6	96
Guangzhou, Guangdong	113.6	22.7	Apr-04	May-04	Period maximum 1-h	Urban	178.0	91.9	97
Jinan, Shandong	117.1	36.7	May-04		Period maximum 1-h	Urban	131.4	141.3	96
Mt Huang, Anhui	118.2	30.1	May-04		Period 24-h average	Rural	67.8	66.7	98
Mt Tai, Shandong	117.2	36.4	May-04		Period 24-h average	Rural	64.4	55.9	98
Beijing	117.1	40.7	May-04		Period 24-h average	Rural	42.5	40.7	98
Wan-Li, Taiwan	121.7	25.2	May-04		Period 24-h average	Rural	32.9	34.1	98
Hong Kong SAR	114.1	22.4	May-04		Period 24-h average	Urban	25.5	22.3	98
Mt Tai, Shandong	117.2	36.4	May-04		Period maximum 1-h	Urban	111.0	120.4	98
Mt Huang, Anhui	118.2	30.1	May-04		Period maximum 1-h	Rural	114.0	102.3	98
Jinan, Shandong	117.1	36.7	Jun-04		Period maximum 1-h	Urban	143.8	110.8	96
Jinan, Shandong	117.1	36.7	Jul-04		Period maximum 1-h	Urban	136.2	140.8	96
Jinan, Shandong	117.1	36.7	Aug-04		Period maximum 1-h	Urban	109.0	125.4	96
Jinan, Shandong	117.1	36.7	Sep-04		Period maximum 1-h	Urban	114.3	119.1	96
Guangzhou, Guangdong	113.6	22.6	Oct-04	Nov-04	Period 24-h average	Rural	49	49.3	99
Guangzhou, Guangdong	113.3	23.1	Oct-04	Nov-04	Period 24-h average	Urban	29	29.9	99
Jinan, Shandong	117.1	36.7	Oct-04		Period maximum 1-h	Urban	107.1	102.3	96
Beijing	117.1	40.7	Jan-05	Dec-05	Period 24-h average	Rural	32.8±30.4	30.3	92
Shanghai	121.1	31.5	May-05		Period maximum 1-h	Urban	127	100.9	97
Beijing	116.3	40.4	Jun-05	Jul-05	Period maximum 1-h	Urban	286	129.7	100
Beijing	117.1	40.7	Jan-06	Dec-06	Period 24-h average	Rural	30.9±29.3	30.2	92
Mt Tai, Shandong	117.1	36.3	May-06	Jun-06	Period 24-h average	Urban	82	89.5	101
Shanghai	121.4	31.2	Jun-06	Jun-07	Period maximum 1-h	Urban	128	126.3	102
Shanghai	121.4	31.2	Jun-06	Jun-07	Monthly avg daily max	Urban	17-70	15.8-66.7	102
Shanghai	121.4	31.2	Jun-06	Jun-07	Period 24-h average	Urban	6-28	12.0-29.9	102
Lanzhou, Gansu	103.7	36.1	Jun-06	Jul-06	Period maximum 1-h	Rural	143	102.0	97
Lanzhou, Gansu	103.7	36.1	Jun-06	Jul-06	Period 24-h average	Rural	53±24	48.1	103
Beijing	115.7	39.1	Jul-06	Sep-07	Period maximum 1-h	Rural	100.7	100.4	104
Qingyuan, Guangdong	113.0	23.5	Jul-06		Diurnal average	Rural	54±18	58.8	105
Guangzhou, Guangdong	113.3	23.1	Jul-06		Diurnal average	Urban	51±29	53.5	105
Mt Waliguan, Qinghai	100.9	36.3	Jul-06	Aug-06	Period 24-h average	Rural	59±8	61.0	103
Beijing	116.8	40.5	Aug-06		Diurnal average	Rural	65	68.4	106
Peking Uni, Beijing	116.3	40.0	Aug-06		Period maximum 1-h	Urban	123	122.3	107
Tianjin	117.2	39.1	Sep-06	Oct-06	Diurnal maximum	Urban	117	129.1	108
Beijing	116.4	39.9	Jan-07	Jan-10	Period maximum 1-h	Urban	60–120	47.3–108.2	109
Beijing	115.7	39.1	Jun-07		Monthly average	Rural	54.8±18.1	57.3	104
Beijing	115.7	39.1	Jun-07		DMA8h	Rural	108.6±23.6	113.9	104
Beijing	115.7	39.1	Jun-07		Daily mean values	Rural	70.0±13.1	78.4	104
Beijing	117.1	40.7	Jun-07	Sep-07	Period 24-h average	Rural	58.2±32.1	54.2	110
Beijing	116.3	39.8	Jun-07	Sep-07	Period 24-h average	Urban	36.2±34.1	37.4	110
Beijing	116.6	40.1	Jun-07	Sep-07	Period 24-h average	Urban	39.6±36.6	37.1	110
Beijing	116.4	39.9	Jun-07	Sep-07	Period 24-h average	Urban	47.0±41.6	43.7	110
Songyuan, Jilin	125.0	45.0	Jun-07		Period 24-h average	Urban	100	99.7	111
Beijing	116.8	40.5	Aug-07		Diurnal average	Rural	50	50.2	106
Shanghai	121.5	31.2	Sep-07		Period 24-h average	Urban	20–60	32.7	112
Guangzhou, Guangdong	113.6	22.7	Oct-07	Dec-07	Period 24-h average	Rural	40±3	42.4	113
Hong Kong SAR	113.9	22.3	Oct-07	Dec-07	Period 24-h average	Urban	32±1	31.0	113
Beijing	116.3	40.0	Nov-07	Mar-08	Period 24-h average	Urban	11.9±0.8	15.9	114
Beijing	116.3	40.0	Nov-07	Mar-08	Period maximum 1-h	Urban	69.7	67.2	114
Guangzhou, Guangdong	113.6	22.7	Nov-07		Period 24-h average	Rural	59±5	55.9	115
Shangri-La, Yunnan	99.7	28.0	2007–2009	January	Monthly average	Rural	45.4±5.6	47.4	116
Shangri-La, Yunnan	99.7	28.0	2007–2009	February	Monthly average	Rural	50.6±5.8	51.8	116
Shangri-La, Yunnan	99.7	28.0	2007–2009	March	Monthly average	Rural	57.1±6.9	59.4	116
Shangri-La, Yunnan	99.7	28.0	2007–2009	April	Monthly average	Rural	58.3±8.8	60.9	116
Shangri-La, Yunnan	99.7	28.0	2007–2009	May	Monthly average	Rural	50.2±9.8	49.1	116
Shangri-La, Yunnan	99.7	28.0	2007–2009	June	Monthly average	Rural	37.4±11.6	33.6	116
Shangri-La, Yunnan	99.7	28.0	2007–2009	July	Monthly average	Rural	26.8±12.5	24.6	116
Shangri-La, Yunnan	99.7	28.0	2007–2009	August	Monthly average	Rural	24.2±8.8	26.2	116
Shangri-La, Yunnan	99.7	28.0	2007–2009	September	Monthly average	Rural	29.6±9.2	29.0	116
Shangri-La, Yunnan	99.7	28.0	2007–2009	October	Monthly average	Rural	31.4±10.1	29.9	116
Shangri-La, Yunnan	99.7	28.0	2007–2009	November	Monthly average	Rural	38.1±7.8	35.2	116
Shangri-La, Yunnan	99.7	28.0	2007–2009	December	Monthly average	Rural	39.7±5.0	36.3	116
Xi'an, Shaanxi	108.9	34.3	Jun-08		Monthly average	Urban	33.5	35.4	117
Beijing	117.5	40.4	Jul-08	Aug-08	Period 24-h average	Rural	67.0	55.7	118
Baoding, Hebei	115.5	38.9	Jul-08	Aug-08	Period 24-h average	Rural	55.3	50.5	118
Beijing	116.0	39.5	Jul-08	Aug-08	Period 24-h average	Rural	47.1	47.7	118
Olympic Vill, Beijing	116.6	40.0	Jul-08	Aug-08	Period 24-h average	Urban	55.3	45.7	118
Langfang, Hebei	116.8	39.6	Jul-08	Aug-08	Period 24-h average	Rural	46.8	45.7	118

Site location	Longitude (°E)	Latitude (°N)	Period start	Period end	Metric	Type	Observed	Modelled	Refs
Beijing	116.1	40.1	Jul-08	Aug-08	Period 24-h average	Rural	46.5	45.6	118
Beijing	116.4	40.0	Jul-08		Period maximum 1-h	Urban	190.0	116.2	119
Peking Uni, Beijing	116.3	40.0	Aug-08	Sep-08	Period maximum 1-h	Urban	135.0	112.8	120
Beijing	116.4	40.0	Aug-08		Period maximum 1-h	Rural	128.0	115.6	121
Beijing	116.4	40.0	Aug-08		Period maximum 1-h	Rural	111.3	115.6	121
Beijing	116.4	40.0	Aug-08		Period maximum 1-h	Rural	118.0	115.6	121
Xi'an, Shaanxi	108.9	34.3	2008	Spring	Period 24-h average	Urban	21.8±10.1	22.3	117
Xi'an, Shaanxi	108.9	34.3	2008	Summer	Period 24-h average	Urban	32.5±11.6	31.1	117
Xi'an, Shaanxi	108.9	34.3	2008	Autumn	Period 24-h average	Urban	8.8±8.1	21.0	117
Xi'an, Shaanxi	108.9	34.3	2009	Winter	Period 24-h average	Urban	3.0±2.5	15.3	117
Tianjin	117.2	39.1	Jan-09	Dec-15	DMA8h	Urban	120.0	115.9	122
Tianjin	117.0	39.4	Jul-09	Sep-09	Period maximum 1-h	Rural	193.7	160.4	123
Tianjin	117.2	39.1	Jul-09	Sep-09	Period maximum 1-h	Urban	130.4	139.0	123
Beijing	116.4	40.0	Jul-10		Period 24-h average	Urban	3.1-66.3	36.4	124
Beijing	116.0	39.7	Jul-10		Period 24-h average	Urban	8.2-105.1	35.8	124
Beijing	116.0	39.7	Aug-10		Period 24-h average	Urban	22.3-89.1	33.7	124
Beijing	116.4	40.0	Aug-10		Period 24-h average	Urban	11.7-53.1	33.2	124
Hong Kong SAR	114.1	22.4	Oct-10		Period 24-h average	Urban	31.9-47.5	65.7	125
Nam Co, Tibet	91.0	30.8	Jan-11	Dec-11	Period 24-h average	Rural	23.5±6.2	42.6	126
Mt Huang, Anhui	118.2	30.1	Jun-11		Period 24-h average	Rural	12.8-51.0	39.8	127
Beijing	116.0	39.7	Jul-11		Period 24-h average	Urban	36.3-80.8	52.2	124
Beijing	116.4	40.0	Jul-11		Period 24-h average	Urban	4.6-54.1	37.3	124
Beijing	116.0	39.7	Aug-11		Period 24-h average	Urban	30.9-74.6	35.4	124
Beijing	116.4	40.0	Aug-11		Period 24-h average	Urban	6.6-56.1	33.9	124
Nanjing, Jiangsu	119.0	32.1	Aug-11		Monthly average	Rural	23.7	29.1	128
Nanjing, Jiangsu	119.0	32.1	Sep-11		Monthly average	Rural	29.2	27.5	128
Nanjing, Jiangsu	119.0	32.1	Oct-11		Monthly average	Rural	22.8	27.7	128
Nanjing, Jiangsu	119.0	32.1	Nov-11		Monthly average	Rural	7.4	18.5	128
Nanjing, Jiangsu	119.0	32.1	Dec-11		Monthly average	Rural	8.6	12.4	128
Nanjing, Jiangsu	119.0	32.1	Jan-12		Monthly average	Rural	16.3	14.4	128
Nam Co, Tibetan	91.0	30.8	Jan-12	Dec-12	Period 24-h average	Rural	48.1±11.4	44.0	126
Nanjing, Jiangsu	119.0	32.1	Feb-12		Monthly average	Rural	15.5	18.5	128
Nanjing, Jiangsu	119.0	32.1	Mar-12		Monthly average	Rural	17.1	20.0	128
Nanjing, Jiangsu	119.0	32.1	Apr-12		Monthly average	Rural	21.8	20.0	128
Nanjing, Jiangsu	119.0	32.1	May-12		Monthly average	Rural	20.4	23.5	128
Nanjing, Jiangsu	119.0	32.1	Jun-12		Monthly average	Rural	27.1	25.3	128
Nanjing, Jiangsu	119.0	32.1	Jul-12		Monthly average	Rural	31.3	25.8	128
Nam Co, Tibetan	91.0	30.8	Jan-13	Dec-13	Period 24-h average	Rural	47.5±12.3	42.2	126
Wuhan, Hubei	114.4	30.5	Feb-13	Oct-14	Daily Maximum average	Urban	85.0	81.4	129
Nanjing, Jiangsu	118.7	32.2	Jun-13	Aug-13	Period maximum 1-h	Urban	110.6	114.8	130
Nanjing, Jiangsu	118.7	32.2	Jun-13	Aug-13	Period maximum 1-h	Urban	129.2	114.8	130
Nanjing, Jiangsu	118.7	32.2	Jun-13	Aug-13	Period maximum 1-h	Urban	135.1	114.8	130
Nanjing, Jiangsu	118.7	32.1	Jun-13	Aug-13	Period maximum 1-h	Urban	134.1	114.8	130
Lanzhou, Gansu	103.8	36.1	Jun-13	Jul-13	Diurnal maximum	Urban	48–98	83.3	131
Lanzhou, Gansu	103.7	36.1	Jun-13	Jul-13	Diurnal maximum	Rural	66–138	88.3	131
Hangzhou, Zhejiang	120.2	30.3	Jul-13	Aug-13	Period 24-h average	Urban	45.5±15.1	52.5	132
Hangzhou, Zhejiang	120.3	30.3	Jul-13	Aug-13	Period 24-h average	Rural	42.0±10.8	36.8	132
Hangzhou, Zhejiang	119.0	29.6	Jul-13	Aug-13	Period 24-h average	Rural	42.0±10.8	36.0	132
Fudan Uni, Shanghai	121.5	31.3	Aug-13		DMA8h	Urban	15.8–117.0	81.1	133
Hong Kong SAR	114.1	22.4	Nov-13	Dec-13	Period 24-h average	Rural	30.6–32.7	58.5	134
Nam Co, Tibetan	91.0	30.8	Jan-14	Dec-14	Period 24-h average	Rural	24.2±5.4	45.3	126
North China	114.5–119.5	36.5–40.5	May-14	Jul-17	DMA8h	NA	98.5	104.3	135
North China	114.5–119.5	36.5–40.5	May-14	Jul-17	Maximum of DMA8h	NA	124.4	116.2	135
Ningbo, Zhejiang	121.5	29.9	Sep-14	Aug-15	Period hourly average	Urban	11–39	21.9	136
Ningbo, Zhejiang	121.6	29.8	Sep-14	Aug-15	Period hourly average	Rural	22–53	30.2	136
Ningbo, Zhejiang	121.9	29.8	Sep-14	Aug-15	Period hourly average	Rural	22–53	30.3	136
Mt Tai, Shandong	117.0	36.3	Jan-15	Dec-15	Daily maximum average	NA	~100	58.3	137
Nam Co, Tibetan	91.0	30.8	Jan-15	Dec-15	Period 24-h average	Rural	48.9±12.0	46.1	126
Kashgar, Xinjiang	76.0	39.5	2015	Autumn	Period 24-h average	Urban	13.9	20.7	138
Nanjing, Jiangsu	118.8	32.1	Jan-16	Dec-16	DM8h 90th percentile	Urban	93.9	75.5	139
Shanghai	121.5	30.8	May-16		DMA8h	Rural	106.4	104.1	140
Hangzhou, Zhejiang	120.2	30.2	Aug-16	Sep-16	Period 24-h average	Urban	64.7	32.1	141
Hangzhou, Zhejiang	120.2	30.2	Sep-16		Period 24-h average	Urban	38.4	32.1	141
Kashgar, Xinjiang	76.0	39.5	2016	Spring	Period 24-h average	Urban	16.2	22.1	138
Shanghai	121.5	30.8	Dec-17		Period 24-h average	Rural	35.0	37.5	142
Shanghai	121.4	31.2	2017	Autumn	Period maximum 1-h	Urban	146.0	122.9	143
Kashgar, Xinjiang	76.0	39.5	2017	Summer	Period 24-h average	Urban	29.6	30.2	138
Fuzhou, Fujian	119.3	26.1	May-18		Period 24-h average	Urban	24.6	22.2	144
Fuzhou, Fujian	119.4	26.0	May-18		Period 24-h average	Urban	20.6	22.0	144
Shenzhen, Guangdong	114.0	22.6	Sep-18	Oct-18	Period maximum 1-h	Urban	121.0	101.5	145
Shanghai	121.4	31.2	2018	Spring	Period maximum 1-h	Urban	137.0	121.3	143
Kashgar, Xinjiang	76.0	39.5	2018	Winter	Period 24-h average	Urban	5.4	20.3	138
Shanghai	121.5	31.3	May-19	Sep-19	Period 24-h average	Urban	35.14±18.72	32.5	146
Shanghai	121.4	31.2	2019	Summer	Period maximum 1-h	Urban	185.0	162.7	143
Shanghai	121.4	31.2	2019	Winter	Period maximum 1-h	Urban	76.7	80.2	143

REFERENCES

1. Sun Z, Archibald AT. Multi-stage ensemble-learning-based model fusion for surface ozone simulations: A focus on CMIP6 models. *Environ Sci Ecotechnol* 2021; **8**: 100124.
2. Sun H, Shin YM, Xia M, et al. Spatial Resolved Surface Ozone with Urban and Rural Differentiation during 1990-2019: A Space-Time Bayesian Neural Network Downscaler. *Environ Sci Technol* 2022; **56**(11): 7337-49.
3. Wu TW, Zhang F, Zhang J, et al. Beijing Climate Center Earth System Model version 1 (BCC-ESM1): model description and evaluation of aerosol simulations. *Geosci Model Dev* 2020; **13**(3): 977-1005.
4. Danabasoglu G, Lamarque JF, Bacmeister J, et al. The Community Earth System Model Version 2 (CESM2). *J Adv Model Earth Sy* 2020; **12**(2): e2019MS001916.
5. van Noije T, Bergman T, Le Sager P, et al. EC-Earth3-AerChem: a global climate model with interactive aerosols and atmospheric chemistry participating in CMIP6. *Geosci Model Dev* 2021; **14**(9): 5637-68.
6. Dunne JP, Horowitz LW, Adcroft AJ, et al. The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics. *J Adv Model Earth Sy* 2020; **12**(11): e2019MS002015.
7. Miller RL, Schmidt GA, Nazarenko LS, et al. CMIP6 Historical Simulations (1850-2014) With GISS-E2.1. *J Adv Model Earth Sy* 2021; **13**(1): e2019MS002034.
8. Yukimoto S, Kawai H, Koshiro T, et al. The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component. *J Meteorol Soc Jpn* 2019; **97**(5): 931-65.
9. Sellar AA, Jones CG, Mulcahy JP, et al. UKESM1: Description and Evaluation of the U.K. Earth System Model. *J Adv Model Earth Sy* 2019; **11**(12): 4513-58.
10. Hegglin M, Kinnison D, Lamarque J-F, Plummer D. CCMI ozone in support of CMIP6 - version 1.0. Earth System Grid Federation; 2016.
11. Schultz MG, Schroder S, Lyapina O, et al. Tropospheric Ozone Assessment Report: Database and metrics data of global surface ozone observations. *Elementa-Sci Anthrop* 2017; **5**(58): 1-26.
12. Sengupta U, Amos M, Hosking S, Rasmussen CE, Juniper M, Young P. Ensembling geophysical models with Bayesian neural networks. *Adv Neural Inf Process Syst* 2020; **33**: 1205-17.
13. Byun G, Choi Y, Kim S, Lee JT. Long-term exposure to ambient ozone and mortality in a population-based cohort of South Korea: Considering for an alternative exposure time metric. *Environ Pollut* 2022; **314**: 120300.
14. Sun HZ, Yu P, Lan C, et al. Cohort-based long-term ozone exposure-associated mortality risks with adjusted metrics: A systematic review and meta-analysis. *The Innovation* 2022; **3**(3): 100246.
15. Global Burden of Disease Mortality and Causes of Death Collaborators. GBD 2019 Results. 2022. http://ghdx.healthdata.org/gbd-results-tool.
16. Zheng P, Barber R, Sorensen RJD, Murray CJL, Aravkin AY. Trimmed Constrained Mixed Effects Models: Formulations and Algorithms. *J Comput Graph Stat* 2021; **30**(3): 544-56.
17. Murray CJL, Aravkin AY, Zheng P, et al. 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* 2020; **396**(10258): 1223-49.
18. Burnett RT, Pope III CA, Ezzati M, et al. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environ Health Persp* 2014; **122**(4): 397-403.
19. Global Burden of Disease Collaborative Network. Ambient ozone pollution — Level 3 risk. 2020. https://www.healthdata.org/results/gbd_summaries/2019/ambient-ozone-pollution-level-3-risk.
20. Yin P, Brauer M, Cohen AJ, et al. The effect of air pollution on deaths, disease burden, and life expectancy across China and its provinces, 1990-2017: an analysis for the Global Burden of Disease Study 2017. *Lancet Planet Health* 2020; **4**(9): e386-e98.
21. Jerrett M, Burnett RT, Pope III CA, et al. Long-term ozone exposure and mortality. *N Engl J Med* 2009; **360**(11): 1085-95.
22. Malley CS, Henze DK, Kuylenstierna JCI, et al. Updated Global Estimates of Respiratory Mortality in Adults ≥30 Years of Age Attributable to Long-Term Ozone Exposure. *Environ Health Persp* 2017; **125**(8): 087021.
23. Center for International Earth Science Information Network (Columbia University). Gridded Population of the World, Version 4 (GPWv4): Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals, Revision 11. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC); 2018.
24. Yang J, Zhou M, Ren Z, et al. Projecting heat-related excess mortality under climate change scenarios in China. *Nat Commun* 2021; **12**(1): 1039.
25. Jones B, O'Neill BC. Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways. *Environ Res Lett* 2016; **11**(8): 084003.
26. Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 (GBD 2019) Population Estimates 1950–2019. Seattle, WA: Institute for Health Metrics and Evaluation; 2020.
27. National Bureau of Statistics of China. China Statistical Yearbook 1981–2021. www.yearbookchina.com. 2022.
28. Liu X, Zhu YJ, Xue L, Desai AR, Wang HK. Cluster-Enhanced Ensemble Learning for Mapping Global Monthly Surface Ozone From 2003 to 2019. *Geophys Res Lett* 2022; **49**(11): e2022GL097947.

29. Zhou M, Wang H, Zeng X, et al. Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2019; **394**(10204): 1145-58.
30. Niu Y, Zhou Y, Chen R, et al. Long-term exposure to ozone and cardiovascular mortality in China: a nationwide cohort study. *Lancet Planet Health* 2022; **6**(6): e496-e503.
31. Liu S, Zhang Y, Ma R, et al. Long-term exposure to ozone and cardiovascular mortality in a large Chinese cohort. *Environ Int* 2022; **165**: 107280.
32. Wang R, Tao S, Ciais P, et al. High-resolution mapping of combustion processes and implications for CO_2 emissions. *Atmos Chem Phys* 2013; **13**(10): 5189-203.
33. Liu Z, Guan D, Wei W, et al. Reduced carbon emission estimates from fossil fuel combustion and cement production in China. *Nature* 2015; **524**(7565): 335-8.
34. Zhong Q, Huang Y, Shen H, et al. Global estimates of carbon monoxide emissions from 1960 to 2013. *Environ Sci Pollut Res Int* 2017; **24**(1): 864-73.
35. Huang Y, Shen H, Chen H, et al. Quantification of global primary emissions of $PM_{2.5}$, PM_{10}, and TSP from combustion and industrial process sources. *Environ Sci Technol* 2014; **48**(23): 13834-43.
36. Wang R, Tao S, Balkanski Y, et al. Exposure to ambient black carbon derived from a unique inventory and high-resolution model. *Proc Natl Acad Sci USA* 2014; **111**(7): 2459-63.
37. Huang Y, Shen HZ, Chen YL, et al. Global organic carbon emissions from primary sources from 1960 to 2009. *Atmos Environ* 2015; **122**: 505-12.
38. Zhong Q, Shen H, Yun X, et al. Global Sulfur Dioxide Emissions and the Driving Forces. *Environ Sci Technol* 2020; **54**(11): 6508-17.
39. Zhong QR, Shen HZ, Yun X, et al. Effects of International Fuel Trade on Global Sulfur Dioxide Emissions. *Environ Sci Tech Let* 2019; **6**(12): 727-31.
40. Huang T, Zhu X, Zhong Q, et al. Spatial and Temporal Trends in Global Emissions of Nitrogen Oxides from 1960 to 2014. *Environ Sci Technol* 2017; **51**(14): 7992-8000.
41. Meng W, Zhong Q, Yun X, et al. Improvement of a Global High-Resolution Ammonia Emission Inventory for Combustion and Industrial Sources with New Data from the Residential and Transportation Sectors. *Environ Sci Technol* 2017; **51**(5): 2821-9.
42. Shen H, Huang Y, Wang R, et al. Global atmospheric emissions of polycyclic aromatic hydrocarbons from 1960 to 2008 and future predictions. *Environ Sci Technol* 2013; **47**(12): 6415-24.
43. Huang X, Li MM, Li JF, Song Y. A high-resolution emission inventory of crop burning in fields in China based on MODIS Thermal Anomalies/Fire products. *Atmos Environ* 2012; **50**: 9-15.
44. Huang Z, Zhong Z, Sha Q, et al. An updated model-ready emission inventory for Guangdong Province by incorporating big data and mapping onto multiple chemical mechanisms. *Sci Total Environ* 2021; **769**: 144535.
45. Kang YN, Liu MX, Song Y, et al. High-resolution ammonia emissions inventories in China from 1980 to 2012. *Atmos Chem Phys* 2016; **16**(4): 2043-58.
46. Liu H, Fu ML, Jin XX, et al. Health and climate impacts of ocean-going vessels in East Asia. *Nat Clim Change* 2016; **6**(11): 1037-41.
47. Zheng B, Cheng J, Geng GN, et al. Mapping anthropogenic emissions in China at 1 km spatial resolution and its application in air quality modeling. *Chinese Sci Bull* 2021; **66**(6): 612-20.
48. Zhou YD, Zhao Y, Mao P, et al. Development of a high-resolution emission inventory and its evaluation and application through air quality modeling for Jiangsu Province, China. *Atmos Chem Phys* 2017; **17**(1): 211-33.
49. Zhou Y, Zhang Y, Zhao B, et al. Estimating air pollutant emissions from crop residue open burning through a calculation of open burning proportion based on satellite-derived fire radiative energy. *Environ Pollut* 2021; **286**: 117477.
50. Hurtt GC, Chini L, Sahajpal R, et al. Harmonisation of global land use change and management for the period 850-2100 (LUH2) for CMIP6. *Geosci Model Dev* 2020; **13**(11): 5425-64.
51. Popp A, Calvin K, Fujimori S, et al. Land-use futures in the shared socio-economic pathways. *Global Environ Chang* 2017; **42**: 331-45.
52. Abbey DE, Nishino N, McDonnell WF, et al. Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. *Am J Respir Crit Care Med* 1999; **159**(2): 373-82.
53. Lipfert FW, Wyzga RE, Baty JD, Miller JP. Traffic density as a surrogate measure of environmental exposures in studies of air pollution health effects: Long-term mortality in a cohort of US veterans. *Atmos Environ* 2006; **40**(1): 154-69.
54. Krewski D, Jerrett M, Burnett RT, et al. Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality: Health Effects Institute Boston, MA; 2009.
55. Smith KR, Jerrett M, Anderson HR, et al. Public health benefits of strategies to reduce greenhouse-gas emissions: health implications of short-lived greenhouse pollutants. *Lancet* 2009; **374**(9707): 2091-103.
56. Lipsett MJ, Ostro BD, Reynolds P, et al. Long-term exposure to air pollution and cardiorespiratory disease in the California teachers study cohort. *Am J Respir Crit Care Med* 2011; **184**(7): 828-35.
57. Zanobetti A, Schwartz J. Ozone and survival in four cohorts with potentially predisposing diseases. *Am J Respir Crit Care Med* 2011; **184**(7): 836-41.
58. Carey IM, Atkinson RW, Kent AJ, van Staa T, Cook DG, Anderson HR. Mortality associations with long-term exposure to outdoor air pollution in a national English cohort. *Am J Respir Crit Care Med* 2013; **187**(11): 1226-33.
59. Jerrett M, Burnett RT, Beckerman BS, et al. Spatial analysis of air pollution and mortality in California. *Am J Respir Crit Care Med* 2013; **188**(5): 593-9.

60. Bentayeb M, Wagner V, Stempfelet M, et al. Association between long-term exposure to air pollution and mortality in France: A 25-year follow-up study. *Environ Int* 2015; **85**: 5-14.

61. Crouse DL, Peters PA, Hystad P, et al. Ambient $PM_{2.5}$, O_3, and NO_2 Exposures and Associations with Mortality over 16 Years of Follow-Up in the Canadian Census Health and Environment Cohort (CanCHEC). *Environ Health Persp* 2015; **123**(11): 1180-6.

62. Tonne C, Halonen JI, Beevers SD, et al. Long-term traffic air and noise pollution in relation to mortality and hospital readmission among myocardial infarction survivors. *Int J Hyg Envir Heal* 2016; **219**(1): 72-8.

63. Turner MC, Jerrett M, Pope III CA, et al. Long-Term Ozone Exposure and Mortality in a Large Prospective Study. *Am J Respir Crit Care Med* 2016; **193**(10): 1134-42.

64. Di Q, Wang Y, Zanobetti A, et al. Air Pollution and Mortality in the Medicare Population. *N Engl J Med* 2017; **376**(26): 2513-22.

65. Weichenthal S, Pinault LL, Burnett RT. Impact of Oxidant Gases on the Relationship between Outdoor Fine Particulate Air Pollution and Nonaccidental, Cardiovascular, and Respiratory Mortality. *Sci Rep* 2017; **7**(1): 16401.

66. Cakmak S, Hebbern C, Pinault L, et al. Associations between long-term $PM_{2.5}$ and ozone exposure and mortality in the Canadian Census Health and Environment Cohort (CANCHEC), by spatial synoptic classification zone. *Environ Int* 2018; **111**: 200-11.

67. Hvidtfeldt UA, Sorensen M, Geels C, et al. Long-term residential exposure to $PM_{2.5}$, PM_{10}, black carbon, NO_2, and ozone and mortality in a Danish cohort. *Environ Int* 2019; **123**: 265-72.

68. Kazemiparkouhi F, Eum KD, Wang B, Manjourides J, Suh HH. Long-term ozone exposures and cause-specific mortality in a US Medicare cohort. *J Expo Sci Environ Epidemiol* 2020; **30**(4): 650-8.

69. Lim CC, Hayes RB, Ahn J, et al. Long-Term Exposure to Ozone and Cause-Specific Mortality Risk in the United States. *Am J Respir Crit Care Med* 2019; **200**(8): 1022-31.

70. Paul LA, Burnett RT, Kwong JC, et al. The impact of air pollution on the incidence of diabetes and survival among prevalent diabetes cases. *Environ Int* 2020; **134**: 105333.

71. Shi L, Rosenberg A, Wang Y, et al. Low-Concentration Air Pollution and Mortality in American Older Adults: A National Cohort Analysis (2001-2017). *Environ Sci Technol* 2022; **56**(11): 7194-202.

72. Strak M, Weinmayr G, Rodopoulou S, et al. Long term exposure to low level air pollution and mortality in eight European cohorts within the ELAPSE project: pooled analysis. *BMJ* 2021; **374**: n1904.

73. Yazdi MD, Wang Y, Di Q, et al. Long-term effect of exposure to lower concentrations of air pollution on mortality among US Medicare participants and vulnerable subgroups: a doubly-robust approach. *Lancet Planet Health* 2021; **5**(10): E689-E97.

74. Bauwelinck M, Chen J, de Hoogh K, et al. Variability in the association between long-term exposure to ambient air pollution and mortality by exposure assessment method and covariate adjustment: A census-based country-wide cohort study. *Sci Total Environ* 2022; **804**: 150091.

75. Stafoggia M, Oftedal B, Chen J, et al. Long-term exposure to low ambient air pollution concentrations and mortality among 28 million people: results from seven large European cohorts within the ELAPSE project. *Lancet Planet Health* 2022; **6**(1): E9-E18.

76. So R, Andersen ZJ, Chen J, et al. Long-term exposure to air pollution and mortality in a Danish nationwide administrative cohort study: Beyond mortality from cardiopulmonary disease and lung cancer. *Environ Int* 2022; **164**: 107241.

77. Yuan Y, Wang K, Sun HZ, et al. Excess mortality associated with high ozone exposure: A national cohort study in China. *Environ Sci Ecotechnol* 2023: 100241.

78. United States Environmental Protection Agency. Air Quality Criteria for Ozone and Related Photochemical Oxidants, 2006.

79. Zheng Y, Stevenson KJ, Barrowcliffe R, Chen S, Wang H, Barnes JD. Ozone levels in Chongqing: a potential threat to crop plants commonly grown in the region? *Environ Pollut* 1998; **99**(3): 299-308.

80. Wang T, Lam KS, Lee ASY, Pang SW, Tsui WS. Meteorological and chemical characteristics of the photochemical ozone episodes observed at Cape D'Aguilar in Hong Kong. *J Appl Meteorol* 1998; **37**(10): 1167-78.

81. Li X, He Z, Fang X, Zhou X. Distribution of surface ozone concentration in the clean areas of china and its possible impact on crop yields. *Adv Atmos Sci* 1999; **16**(1): 154-8.

82. Xu WY, Lin WL, Xu XB, et al. Long-term trends of surface ozone and its influencing factors at the Mt Waliguan GAW station, China - Part 1: Overall trends and characteristics. *Atmos Chem Phys* 2016; **16**(10): 6191-205.

83. Kok GL, Lind JA, Fang M. An airborne study of air quality around the Hong Kong Territory. *J Geophys Res-Atmos* 1997; **102**(D15): 19043-57.

84. Luo C, John JCS, Zhou XJ, Lam KS, Wang T, Chameides WL. A nonurban ozone air pollution episode over eastern China: Observations and model simulations. *J Geophys Res-Atmos* 2000; **105**(D2): 1889-908.

85. Zhang Y, Jin JL, Yan P, et al. Long-term variations of major atmospheric compositions observed at the background stations in three key areas of China. *Adv Clim Chang Res* 2020; **11**(4): 370-80.

86. Wang T, Cheung TF, Li YS, Yu XM, Blake DR. Emission characteristics of CO, NO_X, SO_2 and indications of biomass burning observed at a rural site in eastern China. *J Geophys Res-Atmos* 2002; **107**(D12): ACH-9.

87. Tu J, Xia ZG, Wang HS, Li WQ. Temporal variations in surface ozone and its precursors and meteorological effects at an urban site in China. *Atmos Res* 2007; **85**(3-4): 310-37.

88. Wang T, Cheung VTF, Anson M, Li YS. Ozone and related gaseous pollutants in the boundary layer of eastern China: Overview of the recent measurements at a rural site. *Geophys Res Lett* 2001; **28**(12): 2373-6.

89. Zhang Y, Huang W, London SJ, et al. Ozone and daily mortality in Shanghai, China. *Environ Health Persp* 2006; **114**(8): 1227-32.

90. Wang T, Wong CH, Cheung TF, et al. Relationships of trace gases and aerosols and the emission characteristics at Lin'an, a rural site in eastern China, during spring 2001. *J Geophys Res-Atmos* 2004; **109**(D19): S05.

91. Gao J, Wang T, Ding AJ, Liu CB. Observational study of ozone and carbon monoxide at the summit of mount Tai (1534m a.s.l.) in central-eastern China. *Atmos Environ* 2005; **39**(26): 4779-91.

92. Meng ZY, Xu XB, Yan P, et al. Characteristics of trace gaseous pollutants at a regional background station in Northern China. *Atmos Chem Phys* 2009; **9**(3): 927-36.

93. Shan W, Yin Y, Zhang J, Ji X, Deng X. Surface ozone and meteorological condition in a single year at an urban site in central-eastern China. *Environ Monit Assess* 2009; **151**(1-4): 127-41.

94. Wang T, Wong HLA, Tang J, Ding A, Wu WS, Zhang XC. On the origin of surface ozone and reactive nitrogen observed at a remote mountain site in the northeastern Qinghai-Tibetan Plateau, western China. *J Geophys Res-Atmos* 2006; **111**(D8).

95. Huang W, Tan J, Kan H, et al. Visibility, air quality and daily mortality in Shanghai, China. *Sci Total Environ* 2009; **407**(10): 3295-300.

96. Shan WP, Yin YQ, Lu HX, Liang SX. A meteorological analysis of ozone episodes using HYSPLIT model and surface data. *Atmos Res* 2009; **93**(4): 767-76.

97. Xue LK, Wang T, Gao J, et al. Ground-level ozone in four Chinese cities: precursors, regional transport and heterogeneous processes. *Atmos Chem Phys* 2014; **14**(23): 13175-88.

98. Wang ZF, Li J, Wang XQ, Pochanart P, Akimoto H. Modeling of regional high ozone episode observed at two mountain sites (Mt. Tai and Huang) in East China. *J Atmos Chem* 2006; **55**(3): 253-72.

99. Zhang YH, Hu M, Zhong LJ, et al. Regional Integrated Experiments on Air Quality over Pearl River Delta 2004 (PRIDE-PRD2004): Overview. *Atmos Environ* 2008; **42**(25): 6157-73.

100. Wang T, Ding AJ, Gao J, Wu WS. Strong ozone production in urban plumes from Beijing, China. *Geophys Res Lett* 2006; **33**(21): 806-10.

101. Kanaya Y, Akimoto H, Wang ZF, et al. Overview of the Mount Tai Experiment (MTX2006) in central East China in June 2006: studies of significant regional air pollution. *Atmos Chem Phys* 2013; **13**(16): 8265-83.

102. Ran L, Zhao CS, Geng FH, et al. Ozone photochemical production in urban Shanghai, China: Analysis based on ground level observations. *J Geophys Res-Atmos* 2009; **114**(D15): 301-14.

103. Zhang JM, Wang T, Ding AJ, et al. Continuous measurement of peroxyacetyl nitrate (PAN) in suburban and remote areas of western China. *Atmos Environ* 2009; **43**(2): 228-37.

104. Lin WL, Xu XB, Ge BZ, Zhang XC. Characteristics of gaseous pollutants at Gucheng, a rural site southwest of Beijing. *J Geophys Res-Atmos* 2009; **114**(D2): G14.

105. Lu KD, Zhang YH, Su H, et al. Regional ozone pollution and key controlling factors of photochemical ozone production in Pearl River Delta during summer time. *Sci China Chem* 2010; **53**(3): 651-63.

106. Wang Y, Hao J, McElroy MB, et al. Ozone air quality during the 2008 Beijing Olympics: effectiveness of emission restrictions. *Atmos Chem Phys* 2009; **9**(14): 5237-51.

107. Chou CCK, Tsai CY, Shiu CJ, Liu SC, Zhu T. Measurement of NO_y during Campaign of Air Quality Research in Beijing 2006 (CAREBeijing-2006): Implications for the ozone production efficiency of NO_X. *J Geophys Res-Atmos* 2009; **114**(D2): G01.

108. Han S, Bian H, Feng Y, et al. Analysis of the Relationship between O_3, NO and NO_2 in Tianjin, China. *Aerosol Air Qual Res* 2011; **11**(2): 128-39.

109. Chen PF, Quan JN, Zhang Q, et al. Measurements of vertical and horizontal distributions of ozone over Beijing from 2007 to 2010. *Atmos Environ* 2013; **74**: 37-44.

110. Xu J, Ma JZ, Zhang XL, et al. Measurements of ozone and its precursors in Beijing during summertime: impact of urban plumes on ozone pollution in downwind rural areas. *Atmos Chem Phys* 2011; **11**(23): 12241-52.

111. Ding A, Wang T, Xue L, et al. Transport of north China air pollution by midlatitude cyclones: Case study of aircraft measurements in summer 2007. *J Geophys Res* 2009; **114**(D8): 304-19.

112. Geng F, Zhang Q, Tie X, et al. Aircraft measurements of O_3, NO_X, CO, VOCs, and SO_2 in the Yangtze River Delta region. *Atmos Environ* 2009; **43**(3): 584-93.

113. Guo H, Jiang F, Cheng HR, et al. Concurrent observations of air pollutants at two sites in the Pearl River Delta and the implication of regional transport. *Atmos Chem Phys* 2009; **9**(19): 7343-60.

114. Lin W, Xu X, Ge B, Liu X. Gaseous pollutants in Beijing urban area during the heating period 2007-2008: variability, sources, meteorological, and chemical impacts. *Atmos Chem Phys* 2011; **11**(15): 8157-70.

115. Zhang YL, Wang XM, Blake DR, et al. Aromatic hydrocarbons as ozone precursors before and after outbreak of the 2008 financial crisis in the Pearl River Delta region, south China. *J Geophys Res-Atmos* 2012; **117**(D15): 306-321.

116. Ma J, Lin WL, Zheng XD, Xu XB, Li Z, Yang LL. Influence of air mass downward transport on the variability of surface ozone at Xianggelila Regional Atmosphere Background Station, southwest China. *Atmos Chem Phys* 2014; **14**(11): 5311-25.

117. Wang X, Shen Z, Cao J, et al. Characteristics of surface ozone at an urban site of Xi'an in Northwest China. *J Environ Monit* 2012; **14**(1): 116-26.

118. Gao Y, Zhang M. Sensitivity analysis of surface ozone to emission controls in Beijing and its neighboring area during the 2008 Olympic Games. *J Environ Sci* 2012; **24**(1): 50-61.

119. Wang T, Nie W, Gao J, et al. Air quality during the 2008 Beijing Olympics: secondary pollutants and regional impact. *Atmos Chem Phys* 2010; **10**(16): 7603-15.

120. Chou CCK, Tsai CY, Chang CC, Lin PH, Liu SC, Zhu T. Photochemical production of ozone in Beijing during the 2008 Olympic Games. *Atmos Chem Phys* 2011; **11**(18): 9825-37.

121. Sun Y, Wang L, Wang Y, Quan L, Zirui L. In situ measurements of SO_2, NO_X, NO_y, and O_3 in Beijing, China during August 2008. *Sci Total Environ* 2011; **409**(5): 933-40.

122. Yang JB, Liu JL, Han SQ, Yao Q, Cai ZY. Study of the meteorological influence on ozone in urban areas and their use in assessing ozone trends in all seasons from 2009 to 2015 in Tianjin, China. *Meteorol Atmos Phys* 2019; **131**(6): 1661-75.

123. Han SQ, Zhang M, Zhao CS, et al. Differences in ozone photochemical characteristics between the megacity Tianjin and its rural surroundings. *Atmos Environ* 2013; **79**: 209-16.

124. Wei W, Lv Z, Cheng S, et al. Characterizing ozone pollution in a petrochemical industrial area in Beijing, China: a case study using a chemical reaction model. *Environ Monit Assess* 2015; **187**(6): 377.

125. Guo H, Wang DW, Cheung K, Ling ZH, Chan CK, Yao XH. Observation of aerosol size distribution and new particle formation at a mountain site in subtropical Hong Kong. *Atmos Chem Phys* 2012; **12**(20): 9923-39.

126. Yin XF, Kang SC, de Foy B, et al. Surface ozone at Nam Co in the inland Tibetan Plateau: variation, synthesis comparison and regional representativeness. *Atmos Chem Phys* 2017; **17**(18): 11293-311.

127. Gao J, Zhu B, Xiao H, et al. Diurnal variations and source apportionment of ozone at the summit of Mount Huang, a rural site in Eastern China. *Environ Pollut* 2017; **222**: 513-22.

128. Ding AJ, Fu CB, Yang XQ, et al. Ozone and fine particle in the western Yangtze River Delta: an overview of 1 yr data at the SORPES station. *Atmos Chem Phys* 2013; **13**(11): 5813-30.

129. Lyu XP, Chen N, Guo H, et al. Ambient volatile organic compounds and their effect on ozone production in Wuhan, central China. *Sci Total Environ* 2016; **541**: 200-9.

130. An J, Zou J, Wang J, Lin X, Zhu B. Differences in ozone photochemical characteristics between the megacity Nanjing and its suburban surroundings, Yangtze River Delta, China. *Environ Sci Pollut Res Int* 2015; **22**(24): 19607-17.

131. Jia CH, Mao XX, Huang T, et al. Non-methane hydrocarbons (NMHCs) and their contribution to ozone formation potential in a petrochemical industrialized city, Northwest China. *Atmos Res* 2016; **169**: 225-36.

132. Li K, Chen LH, Ying F, et al. Meteorological and chemical impacts on ozone formation: A case study in Hangzhou, China. *Atmos Res* 2017; **196**: 40-52.

133. Shi CZ, Wang SS, Liu R, et al. A study of aerosol optical properties during ozone pollution episodes in 2013 over Shanghai, China. *Atmos Res* 2015; **153**: 235-49.

134. Wang T, Tham YJ, Xue LK, et al. Observations of nitryl chloride and modeling its source and effect on ozone in the planetary boundary layer of southern China. *J Geophys Res-Atmos* 2016; **121**(5): 2476-89.

135. Gong C, Liao H. A typical weather pattern for ozone pollution events in North China. *Atmos Chem Phys* 2019; **19**(22): 13725-40.

136. Tong L, Zhang HL, Yu J, et al. Characteristics of surface ozone and nitrogen oxides at urban, suburban and rural sites in Ningbo, China. *Atmos Res* 2017; **187**: 57-68.

137. Hayashida S, Kajino M, Deushi M, Sekiyama TT, Liu X. Seasonality of the lower tropospheric ozone over China observed by the Ozone Monitoring Instrument. *Atmos Environ* 2018; **184**: 244-53.

138. Wang HQ, Mamtimin M, Chen F. Some characteristics of ozone concentrations and their linkages with meteorological factors in Kashgar, northwestern China. *Iop C Ser Earth Env* 2017; **74**(1): 012001.

139. Wang M, Chen W, Zhang L, et al. Ozone pollution characteristics and sensitivity analysis using an observation-based model in Nanjing, Yangtze River Delta Region of China. *J Environ Sci* 2020; **93**: 13-22.

140. Zhang K, Zhou L, Fu QY, et al. Vertical distribution of ozone over Shanghai during late spring: A balloon-borne observation. *Atmos Environ* 2019; **208**: 48-60.

141. Li P, Wang L, Guo P, et al. High reduction of ozone and particulate matter during the 2016 G-20 summit in Hangzhou by forced emission controls of industry and traffic. *Environ Chem Lett* 2017; **15**(4): 709-15.

142. Chen Q, Wang D, Li X, et al. Vertical Characteristics of Winter Ozone Distribution within the Boundary Layer in Shanghai Based on Hexacopter Unmanned Aerial Vehicle Platform. *Sustainability* 2019; **11**(24): 7026.

143. Zhang G, Jing S, Xu W, et al. Simultaneous observation of atmospheric peroxyacetyl nitrate and ozone in the megacity of Shanghai, China: Regional transport and thermal decomposition. *Environ Pollut* 2021; **274**: 116570.

144. He Z, Zhang X, Li YF, et al. Characterizing carbonyl compounds and their sources in Fuzhou ambient air, southeast of China. *PeerJ* 2020; **8**: e10227.

145. Yu D, Tan ZF, Lu KD, et al. An explicit study of local ozone budget and NO_X-VOCs sensitivity in Shenzhen China. *Atmos Environ* 2020; **224**: 117304.

146. Zhu J, Wang SS, Wang HL, et al. Observationally constrained modeling of atmospheric oxidation capacity and photochemical reactivity in Shanghai, China. *Atmos Chem Phys* 2020; **20**(3): 1217-32.