Supplementary Information

Title: Global impact of COVID-19 lockdown on surface concentration and health risk of atmospheric benzene

Authors: Chaohao Ling[a,b,1], Lulu Cui[d,1], Rui Li[a, c] [*]

Affiliations: [a] *Key Laboratory of Geographic Information Science of the Ministry of Education, School of Geographic Sciences, East China Normal University, Shanghai, 200241, PR China*

[b] *School of History and Geography, Minnan Normal University, Zhangzhou, 363000, China*

[c] *Institute of Eco-Chongming (IEC), 20 Cuiniao Road, Chenjia Town, Chongming District, Shanghai, 202162, China*

[d] *State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing, 100084, China*

* Corresponding author

Prof. Li (rli@geo.ecnu.edu.cn)
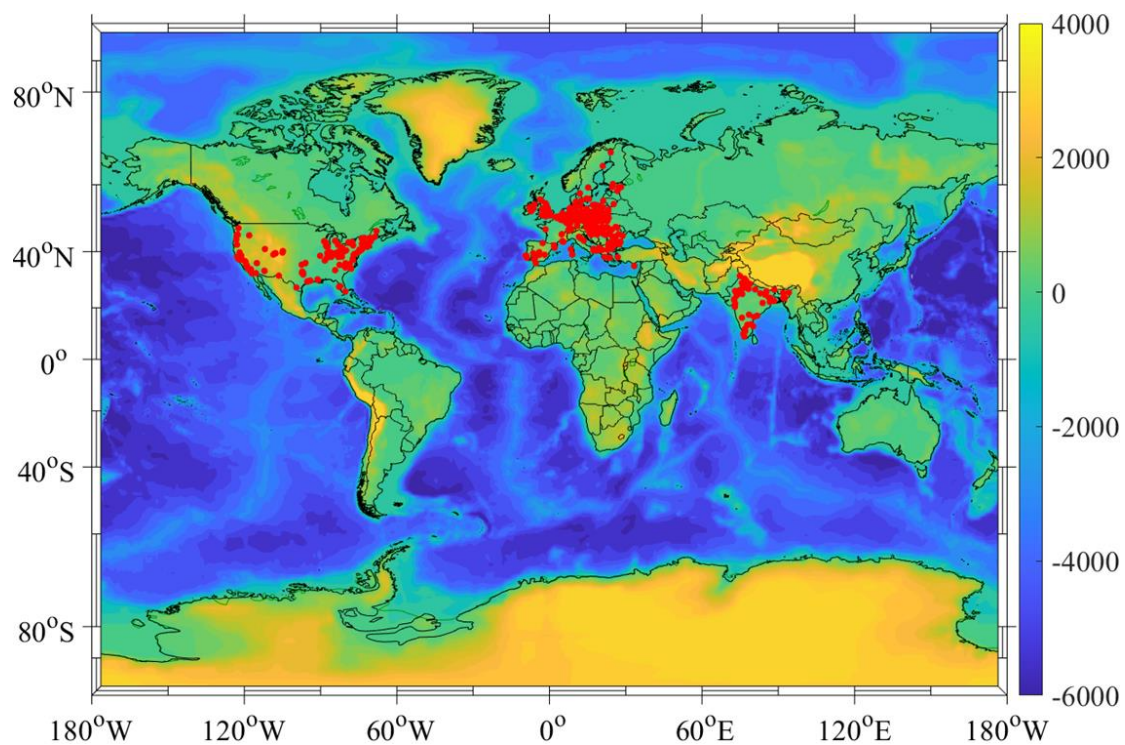
Number of pages: 13

Number of figures: 10

Number of tables: 1

Table S1 Exposure parameters for the calculation of residential inhalation cancer risk.

Parameter	Unit	Value
ET	$h\ d^{-1}$	8
EF	$d\ a^{-1}$	250
ED	a	20
ATnca	a	25
ATca	a	70
Rfc	mg/m^3	0.03
IUR	$1/(\mu g/m^3)$	0.0000078

Figure S1 Spatial distributions of ground-observed benzene monitoring sites used for model development. Red circles represent the ground-level sites during 2019-2020. The colormap denotes the elevation distribution around the world.

Figure S2 Correlation coefficient matrix between atmospheric benzene concentration and all of the explanatory variables around the world. The full names of Benz, Emis, CTM, PS, RH, T2M, TPREC, U10M, V10M, and ZPBL represent benzene concentration, benzene emission, simulated benzene level based on the process-based model from GEOS-CF reanalysis, surface pressure, relative humidity, 2-m air temperature, total precipitation, 10-m latitudinal wind component, 10-m longitudinal wind component, and boundary layer height, respectively. The colorbar reflects the correlation coefficient of these variables.

Figure S3 (a) represents the predictive performance of global atmospheric benzene concentration (μg/m³) based on ensemble decision tree model. (b) denotes the predictive accuracy of ambient benzene based on RF model. (c) denotes the predictive accuracy of ambient benzene using XGBoost algorithm. (d) represents the predictive accuracy of ambient benzene based on LightGBM model. The black solid line denotes the best-fitting curve for all of the points, while the black dashed line represents the diagonal, which means the same observed and simulated values.
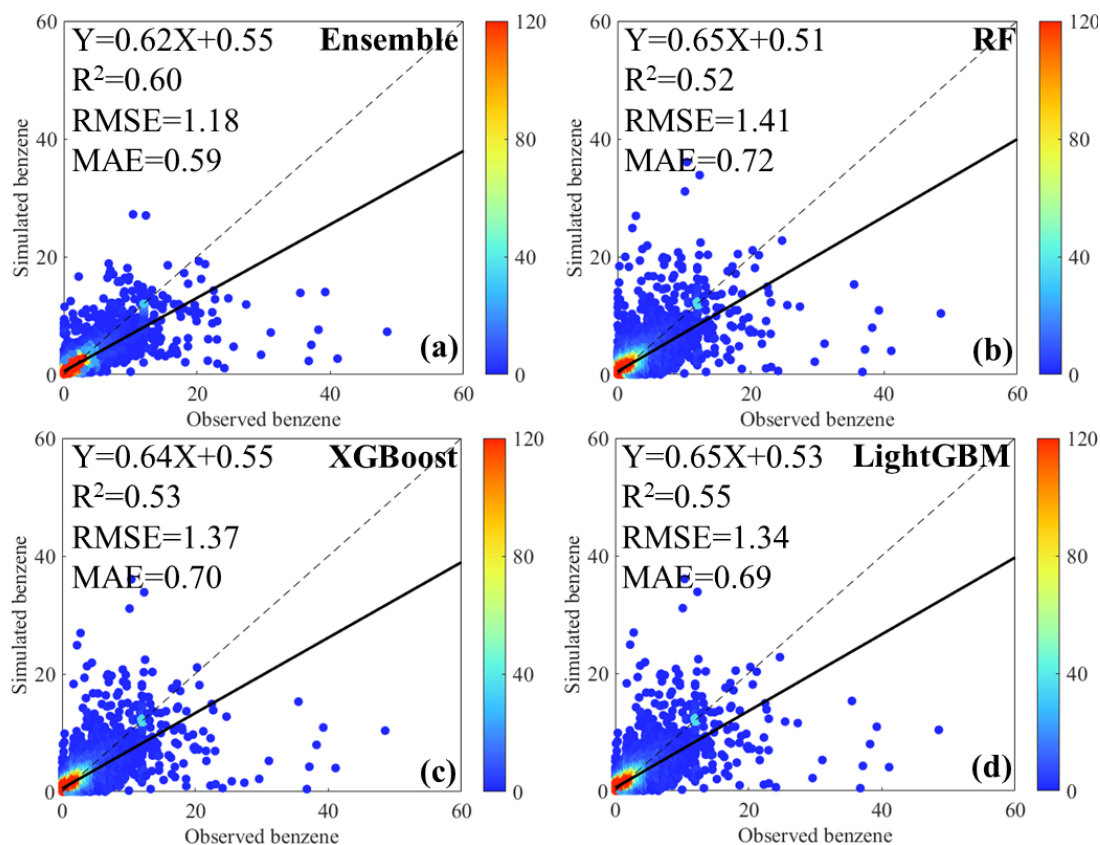
Figure S4 The transferability validation of the ensemble model in estimating global ambient benzene based on the site-based validation method. The linear regression curve is added in the figure. The blue circular represents the data points, and the black solid line denotes the best-fit line through the data points.
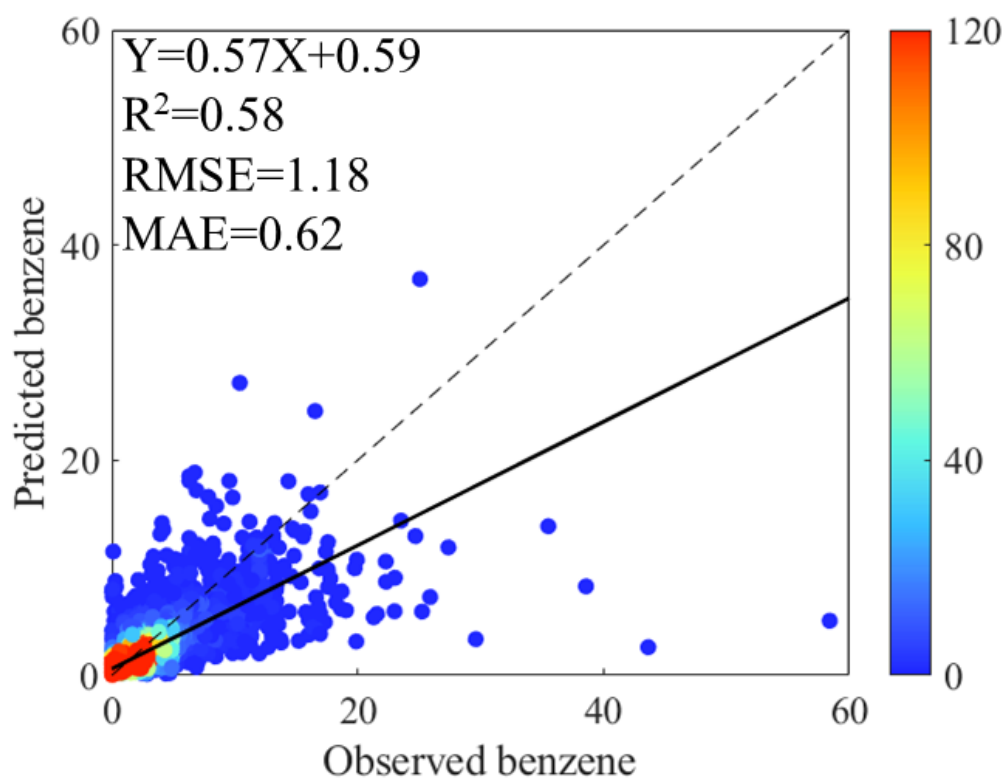
Figure S5 The performance of XGBoost method in predicting the daily meteorology-normalized benzene concentrations around the world. The model was developed with 90% original data and the remained one was applied to validate the model.
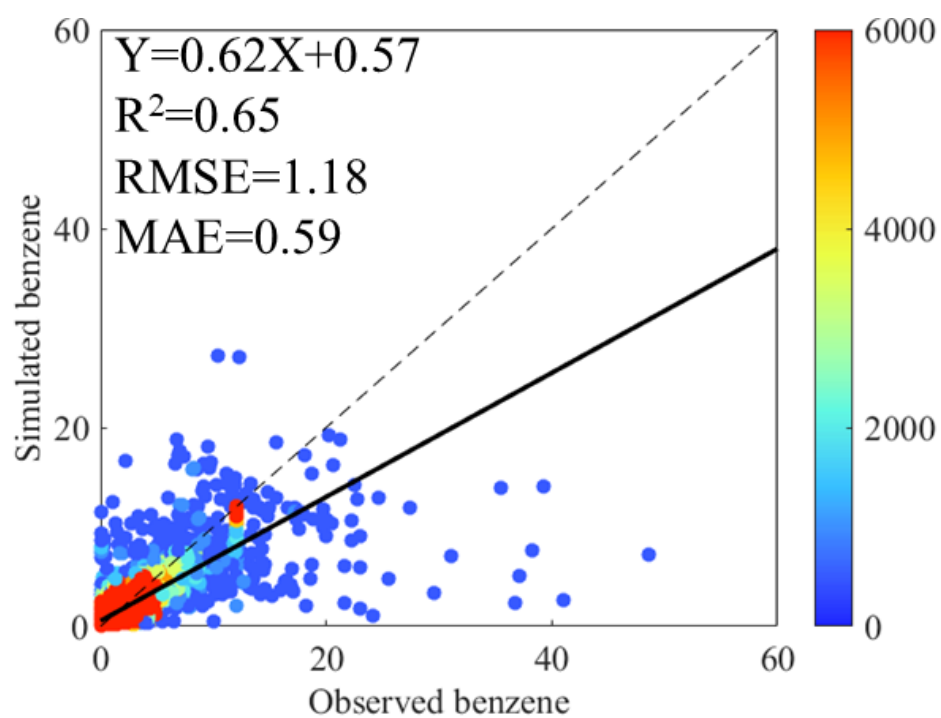
Figure S6 The global average concentrations of predicted (total) atmospheric benzene (μg/m³) during Jan. 23-Jun. 30 in 2019 (a) and predicted ambient benzene (μg/m³) (b) during the same period in 2020.
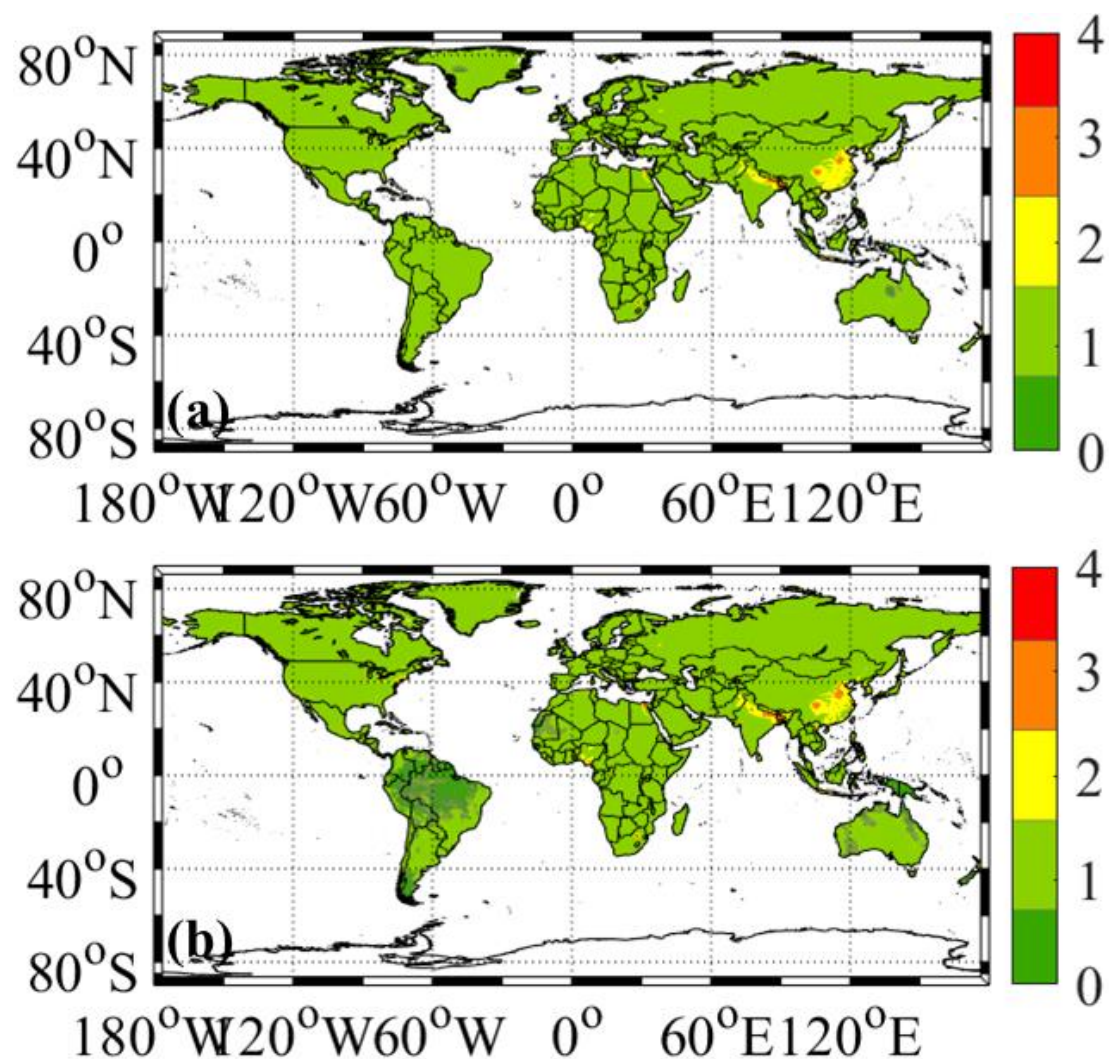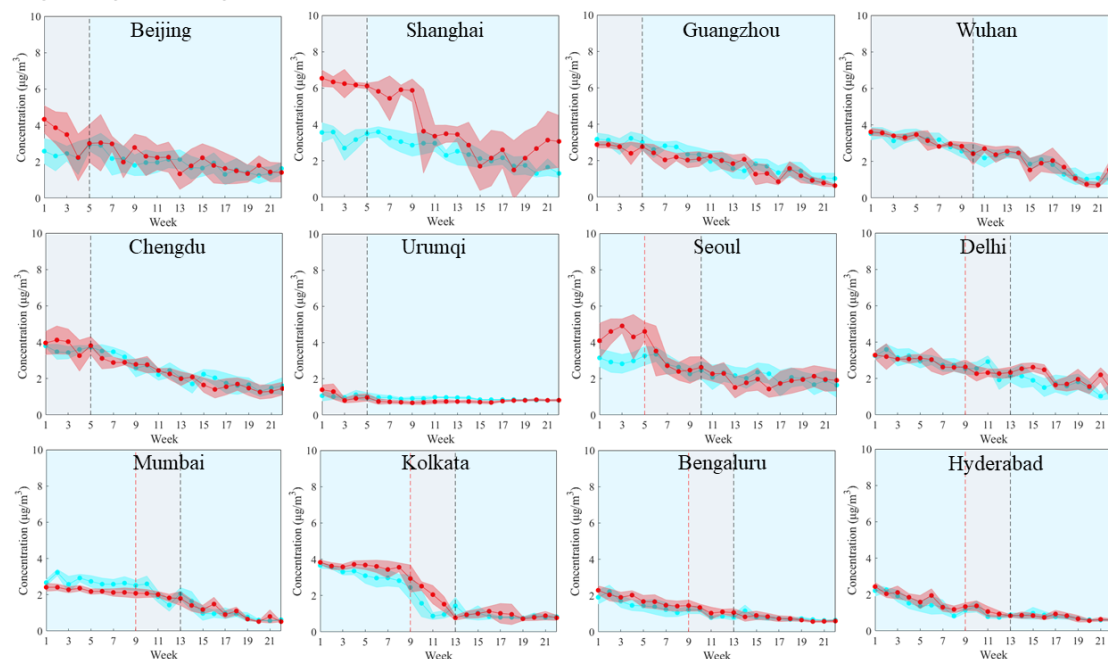
Figure S7 The weekly variations of atmospheric benzene concentrations (μg/m³) in some major regions around the world during Jan. 23-Jun. 30. The red line and background denote mean values and standard deviation of simulated weekly benzene concentrations in 2020. The cyan line and background denote mean values and standard deviation of simulated weekly benzene levels in 2019. The dashed vertical red line suggests COVID-19 restriction dates, and the black line indicates the beginning of easing measures.
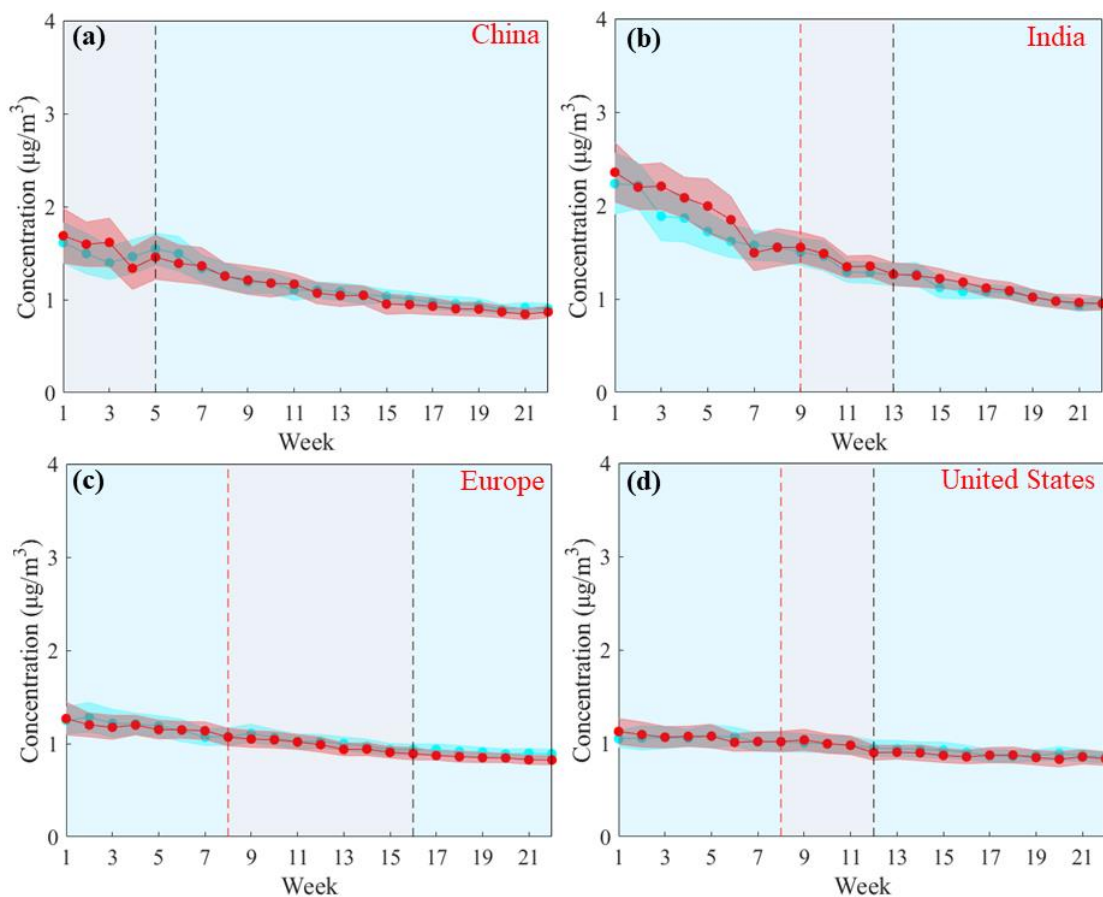
Figure S8 The weekly variations of atmospheric benzene concentrations (μg/m³) in some megacities around the world. The red line and background denote mean values and standard deviation of deweathered weekly benzene concentrations in 2020. The cyan line and background denote mean values and standard deviation of deweathered weekly benzene levels in 2019. The dashed vertical red line suggests COVID-19 restriction dates, and the black line indicates the beginning of easing measures.

Figure S9 The weekly variations of atmospheric benzene concentrations (μg/m³) in some megacities around the world. The red line and background denote mean values and standard deviation of deweathered weekly benzene concentrations in 2020. The cyan line and background denote mean values and standard deviation of estimated deweathered weekly benzene levels in 2019. The dashed vertical red line suggests COVID-19 restriction dates, and the black line indicates the beginning of easing measures.
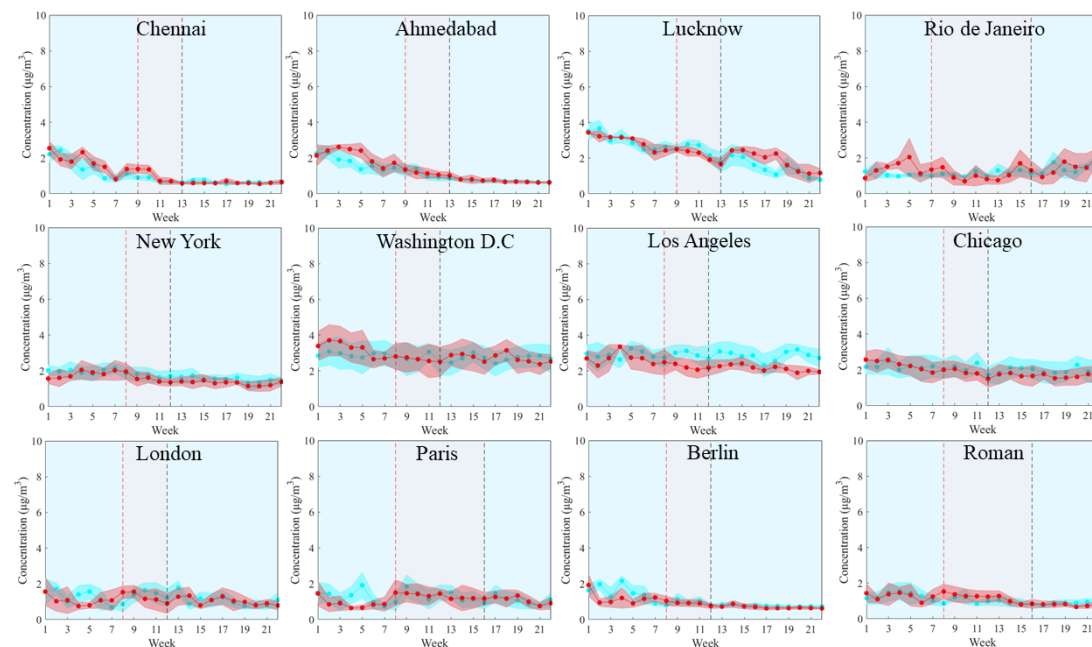
Figure S10 The carcinogenic (Unit: 10^{-7}) and non-carcinogenic risk (Unit: 10^{-3}) differences for global atmospheric benzene between COVID-19 period in 2020 and the same period in 2019 (Difference = benzene concentration in 2020-benzene concentration in 2019).