Supplementary Materials for

Machine-learning-based corrections of CMIP6 historical surface ozone in China during 1950-2014

Yuanxi Tong [1], Yingying Yan [1*], Jintai Lin [2], Shaofei Kong [1, 3], Zhixuan Tong [1], Yifei Zhu [1], Yukun Yan [1], Zhan Sun [1]

[1] Department of Atmospheric Science, School of Environmental Sciences, China University of Geosciences, Wuhan, 430074, China

[2] Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China

[3] Research Centre for Complex Air Pollution of Hubei Province, Wuhan, 430074, China

Correspondence to: Yingying Yan (yanyingying@cug.edu.cn)

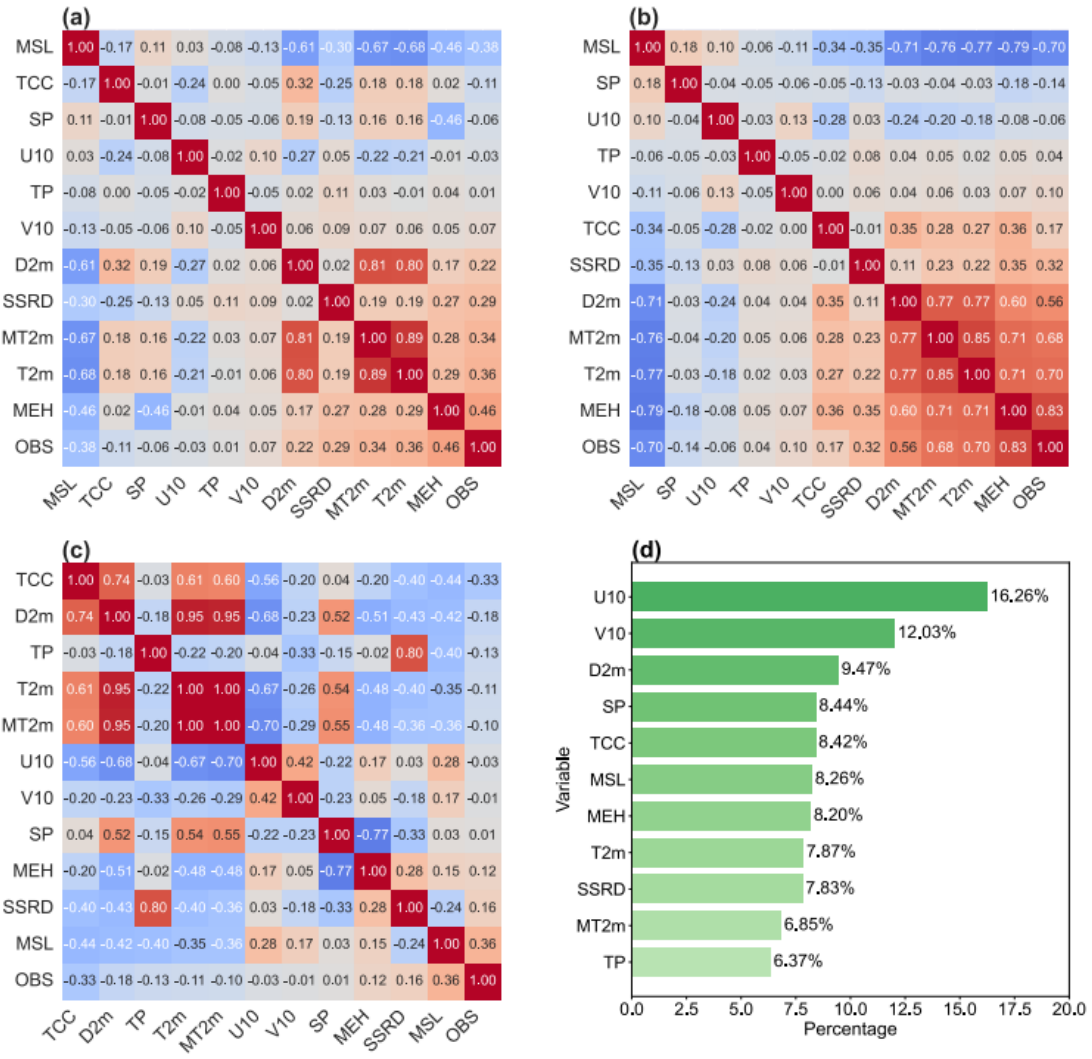


Figure S1 Thermodynamic matrix of spatiotemporal correlation coefficients (a), spatial correlation coefficients (b), temporal correlation coefficients (c), based on meteorological elements, MEH ozone concentration and observed ozone concentration. (d) XGBoost ranks the feature importance of the input data (2014–2022).

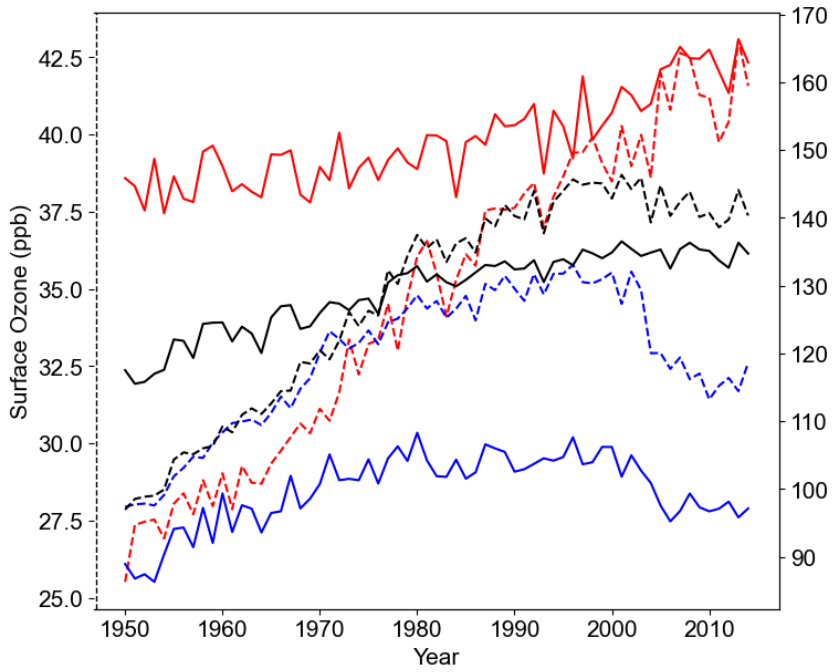


Fig. S2 Annual variation for MEH (dashed lines) and XGB (solid lines) ozone (red represents summer (JJA), blue represents winter (DJF) and black represents the whole year) from 1950 to 2014.

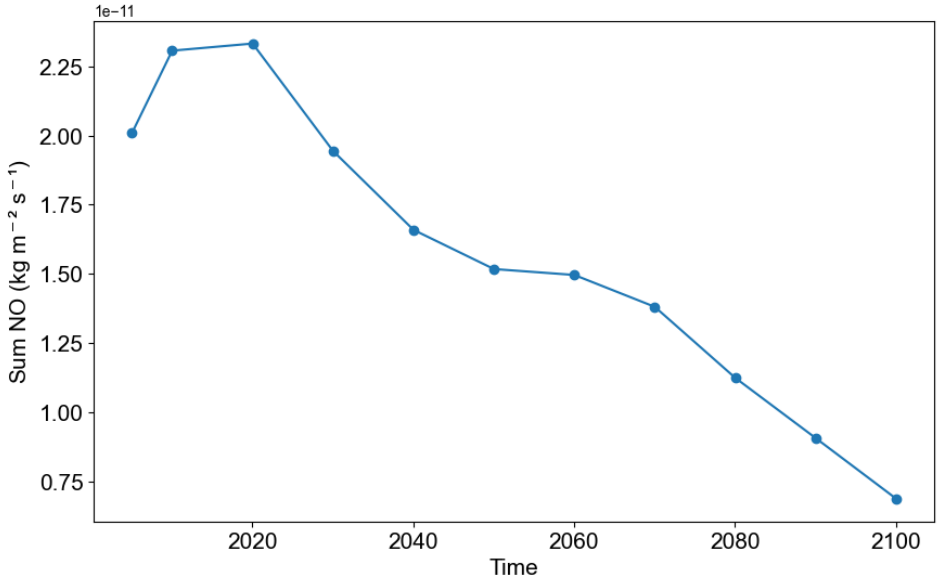


Figure S3 NO emissions of China from 2005 to 2100. Emission data is downloaded from ECCAD (https://eccad.sedoo.fr/#/metadata/393).

Table S1 The ERA5 meteorological data and modeled ozone data used as input to the XGBoost model

Abbreviation	Full name	Unit
U10	10m u-component of wind	m s-1
V10	10m v-component of wind	m s-1
D2m	2m dewpoint temperature	K
T2m	2m temperature	K
MT2m	Maximum 2m temperature since previous post-processing	K
SLP	Mean sea level pressure	Pa
SP	Surface pressure	Pa
SSRD	Surface solar radiation downwards	J m-2
TCC	Total cloud cover	Dimensionless
TP	Total precipitation	m
MEH	Surface ozone simulated by MPI-ESM1.2-HAM climate model	ppb