Environ Health Perspect

DOI: 10.1289/EHP9406

Supplemental Material

Satellite-Based Long-Term Spatiotemporal Patterns of Surface Ozone Concentrations in China: 2005–2019

Qingyang Zhu, Jianzhao Bi, Xiong Liu, Shenshen Li, Wenhao Wang, Yu Zhao, and Yang Liu

Table of Contents

Figure S4. An example of the spatial artifacts in the predicted monthly mean MDA8 ozone concentrations in the YRD (April 2019). Panel A was generated from the model with a priori ozone profile L22-L24; panel B was generated from the original model with OMI L24 retrieval. The green box represents YRD, and the yellow dashed line encircles an example of the spatial artifacts (the hard horizontal gap in predicted ozone concentrations). Abbreviations: MDA8, daily maximum 8-hour average; OMI, Ozone Monitoring Instrument; YRD, Yangtze River Delta.

Figure S5. Season-specific cross-validation results for the original model with OMI L24 retrieval. Left column: random CV; middle column: spatial CV; right column: temporal CV. Top row: spring (March-April-May); second row (from top to bottom): summer (June-July-August); third row: fall (September, October, November); bottom row: winter (December-January-February). The functions on the bottom-right corners are the regression functions between the predicted and observed MDA8 ozone concentrations. Red dashed lines: the regression line between the predictions and observations; black solid lines: the x = y line; The color scale represents the density of the points. Abbreviations: MDA8, daily maximum 8-hour average; OMI, Ozone Monitoring Instrument.

Figure S6. Comparison of model performance in rural/urban areas using OMI L24 retrieval. Left column: random CV; middle column: spatial CV; right column: temporal CV. Upper row: model performance at rural CNEMC sites; Lower row: model performance in urban CNEMC sites. The functions on the bottom-right corners are the regression functions between the predicted and observed MDA8 ozone concentrations. Red dashed lines: the regression line between the predictions and observations; black solid lines: the x = y line; The color scale represents the density of the points. Abbreviations: CNEMC, China National Environmental Monitoring Centre; MDA8, daily maximum 8-hour average; OMI, Ozone Monitoring Instrument.

Figure S7. Trend comparison between model predictions and the TOAR historical monitoring data (original model using OMI L24 retrieval). Each row represents a TOAR monitoring site. Blue lines: the observed monthly mean MDA8 ozone concentrations. Orange lines: the model predicted monthly MDA8 ozone concentrations. Abbreviations: MDA8, daily maximum 8-hour average; OMI, Ozone Monitoring Instrument. TOAR, Tropospheric Ozone Assessment Report.

Figure S8. Trend comparison between model predictions and the TOAR historical monitoring data (model trained with OMI L24 retrieval + MODIS FRP). Each row represents a TOAR monitoring site. Blue lines: the observed monthly mean MDA8 ozone concentrations. Orange lines: the model predicted monthly MDA8 ozone concentrations. Abbreviations: FRP, fire radiative power; MDA8, daily maximum 8-hour average; MODIS, Moderate Resolution Imaging Spectrometer; OMI, Ozone Monitoring Instrument. TOAR, Tropospheric Ozone Assessment Report.

Figure S9. Raw correlation between OMPROFOZ L24 partial column ozone and ground-level monitoring data. The function on the upper-left corner is the regression function between the OMPROFOZ L24 partial column ozone and the CNEMC surface MDA8 ozone concentrations. Red dashed line: the regression line as the regression function represents; black solid line: the x = y line; The color scale represents the density of the points. Abbreviations: CNEMC, China National Environmental Monitoring Centre; MDA8, daily maximum 8-hour average.

Figure S10. model performance without the OMI L24 retrieval. Left panel: random CV; middle panel: spatial CV; right panel: temporal CV. The functions on the bottom-right corners are the regression functions between the predicted and observed monthly mean MDA8 ozone concentrations. Red dashed lines: the regression line between the predictions and observations; black solid lines: the x = y line; The color scale represents the density of the points. Abbreviations: CV, cross-validation; MDA8, daily maximum 8-hour average.

Figure S11. Trend comparison between model predictions and the TOAR historical monitoring data (OMI-excluded model). Each row represents a TOAR monitoring site. Blue lines: the observed monthly mean MDA8 ozone concentrations. Orange lines: the model predicted monthly MDA8 ozone concentrations. Abbreviations: MDA8, daily maximum 8-hour average; OMI, Ozone Monitoring Instrument. TOAR, Tropospheric Ozone Assessment Report.

Figure S12. The importance ranking of the input predictors. Orange bars in the left panel: the spring-excluded (all months except for March-April-May) model; blue bars in the right panel: the spring (March-April-May) model. The red downward arrows denoted that the relative ranking of the variable dropped in the spring model compared to other seasons; the blue upward arrows denoted that the relative ranking of a variable increased in spring. Variables' importance was generated with a permutation method.

Figure S13. The OMI tropospheric column NO2 concentrations in the Yangtze River Delta region (August 2019, the marked regions were consistent to Fig. 5) The boxes represent some YRD cities and their surrounding area; A: Bengbu and Huainan City; B: Nanjing Metropolitan Area; C: Anqing City; D: Quzhou and Jinhua City. Abbreviations: MDA8, daily maximum 8-hour average.

Figure S14. The predicted long-term trend of summertime (June-July-August) mean MDA8 ozone concentrations over 2005-2019. The dots represented the predicted summertime mean MDA8 ozone concentrations; red lines are the linear smooth functions; P indicates the p-value of the linear trends. Abbreviations: MDA8, daily maximum 8-hour average; NCP: the North China Plain; PRD: the Pearl River Delta; YRD: the Yangtze River Delta.

Figure S15. The predicted long-term trend of summertime (June-July-August) mean MDA8 ozone concentrations over 2005-2013. The dots represented the predicted summertime mean MDA8 ozone concentrations; red lines are the linear smooth functions; P indicates the p-value of the linear trends. Abbreviations: MDA8, daily maximum 8-hour average; NCP: the North China Plain; PRD: the Pearl River Delta; YRD: the Yangtze River Delta.

Figure S16. The predicted long-term trend of springtime (March-April-May) mean MDA8 ozone concentrations over 2005-2019. The dots represented the predicted springtime mean MDA8 ozone concentrations; red lines are the linear smooth functions; P indicates the p-value of the linear trends. Abbreviations: MDA8, daily maximum 8-hour average; NCP: the North China Plain; PRD: the Pearl River Delta; YRD: the Yangtze River Delta.

Figure S17. The predicted long-term trend of springtime (March-April-May) mean MDA8 ozone concentrations over 2005-2013. The dots represented the predicted springtime mean MDA8 ozone concentrations; red lines are the linear smooth functions; P indicates the p-value of the linear trends. Abbreviations: MDA8, daily maximum 8-hour average; NCP: the North China Plain; PRD: the Pearl River Delta; YRD: the Yangtze River Delta.

Figure S18. The comparison of the mean ozone season (March-November) MDA8 ozone concentrations between our model predictions and the CNEMC monitoring data over 2014-2019. The blue columns on the left represent the CNEMC observations; the orange column on the right represents our model predictions. The height of the columns and the error bars represent the mean MDA8 ozone concentrations and the standard error. Abbreviations: CNEMC, China National Environmental Monitoring Centre; MDA8, daily maximum 8-hour average; NCP: the North China Plain; PRD: the Pearl River Delta; YRD: the Yangtze River Delta.

Figure S19. The comparison of the summertime (June-July-August) mean MDA8 ozone concentrations between our model predictions and the CNEMC monitoring data over 2014-2019. The blue columns on the left represent the CNEMC observations; the orange column on the right represents our model predictions. The height of the columns and the error bars represent the mean MDA8 ozone concentrations and the standard error. Abbreviations: CNEMC, China National Environmental Monitoring Centre; MDA8, daily maximum 8-hour average; NCP: the North China Plain; PRD: the Pearl River Delta; YRD: the Yangtze River Delta.

Figure S20. Locations of the Shandong Peninsula and Central East China. The solid line encircles the Shandong Peninsula, while the dashed line encircles the Central East China area.

Table S1. List of the input variables

Abbreviation	name	Source
DISPH	Zero plane displacement height	MERRA-2
PBLH	Planetary boundary layer height	MERRA-2
ALBEDO	Surface albedo	MERRA-2
EFLUX	Total latent energy flux	MERRA-2
T10M	10 meter air temperature	MERRA-2
USTAR	Surface velocity scale	MERRA-2
PRECTOT	Total precipitation	MERRA-2
SWGDN	Surface incoming shortwave flux	MERRA-2
RH	Relative humidity	MERRA-2
U	Eastward wind	MERRA-2
V	Northward wind	MERRA-2
PS	Surface pressure	MERRA-2
QV	Specific humidity	MERRA-2
T	Air temperature	MERRA-2
OMEGA	Vertical pressure velocity	MERRA-2
TROPPT	Tropopause pressure based on thermal estimate	MERRA-2
RH_C	Relative humidity	MERRA-2
U_C	Eastward wind	MERRA-2
V_C	Northward wind	MERRA-2
OMEGA_C	Vertical pressure velocity	MERRA-2
PS_C	Surface pressure	MERRA-2
QV_C	Specific humidity	MERRA-2
T_C	Air temperature	MERRA-2
PV_C	Ertels potential vorticity	MERRA-2
Boundary-layer		
O3	Gap-filled boundary-layer column ozone	OMPROFOZ
Elevation	Elevation	ASTER GDEM
LANDUSE_10	Cropland, rainfed	ESA CCI
LANDUSE_11	Herbaceous cover	ESA CCI
LANDUSE_12	Tree or shrub cover	ESA CCI
LANDUSE_20	Cropland, irrigated or post flooding	ESA CCI
LANDUSE_30	Mosaic cropland (>50%) / natural vegetation (tree, shrub, herbaceous cover) (<50%)	ESA CCI
LANDUSE_40	Mosaic natural vegetation (tree, shrub, herbaceous cover) (>50%) / cropland (<50%)	ESA CCI
LANDUSE_50	Tree cover, broadleaved, evergreen, closed to open (>15%)	ESA CCI
LANDUSE_60	Tree cover, broadleaved, deciduous, closed to open (>15%)	ESA CCI
LANDUSE_61	Tree cover, broadleaved, deciduous, closed (>40%)	ESA CCI
LANDUSE_62	Tree cover, broadleaved, deciduous, open (15□]40%)	ESA CCI

LANDUSE_70	Tree cover, needleleaved, evergreen, closed to open (>15%)	ESA CCI
LANDUSE_71	Tree cover, needleleaved, evergreen, closed (>40%)	ESA CCI
LANDUSE_72	Tree cover, needleleaved, evergreen, open (15□]40%)	ESA CCI
LANDUSE_80	Tree cover, needleleaved, deciduous, closed to open (>15%)	ESA CCI
LANDUSE_81	Tree cover, needleleaved, deciduous, closed (>40%)	ESA CCI
LANDUSE_90	Tree cover, mixed leaf type (broadleaved and needleleaved)	ESA CCI
LANDUSE_100	Mosaic tree and shrub (>50%) / herbaceous cover (<50%)	ESA CCI
LANDUSE_110	Mosaic herbaceous cover (>50%) / tree and shrub (<50%)	ESA CCI
LANDUSE_120	Shrubland	ESA CCI
LANDUSE_121	Evergreen shrubland	ESA CCI
LANDUSE_122	Deciduous shrubland	ESA CCI
LANDUSE_130	Grassland	ESA CCI
LANDUSE_140	Lichens and mosses	ESA CCI
LANDUSE_150	Sparse vegetation (tree, shrub, herbaceous cover) (<15%)	ESA CCI
LANDUSE_152	Sparse shrub (<15%)	ESA CCI
LANDUSE_153	Sparse herbaceous cover (<15%)	ESA CCI
LANDUSE_160	Tree cover, flooded, fresh or brakish water	ESA CCI
LANDUSE_170	Tree cover, flooded, saline water	ESA CCI
LANDUSE_180	Shrub or herbaceous cover, flooded, fresh/saline/brakish water	ESA CCI
LANDUSE_190	Urban areas	ESA CCI
LANDUSE_200	Bare areas	ESA CCI
LANDUSE_201	Consolidated bare areas	ESA CCI
LANDUSE_202	Unconsolidated bare areas	ESA CCI
LANDUSE_210	Water bodies	ESA CCI
LANDUSE_220	Permanent snow and ice	ESA CCI
ROAD_LENGTH	Total road length	gROADs
NO2	OMI tropospheric column Nitrogen dioxide	OMI
Population	Population	Landscan
FLASH	Lightning activity	GEOS-Chem
FRP*	Fire radiative power	MODIS
A priori L22_24 *	Gap-filled a priori L22-L24 partial column ozone	OMPROFOZ
A priori L24 *	Gap-filled a priori L24 partial column ozone	OMPROFOZ
OMI L22_L24*	Gap-filled retrieved L22-L24 partial column ozone	OMPROFOZ

* Variables not included in the original model but used for parameter comparison.

Table S2. Temporal trends of the TOAR monitoring data during 2005-2013

Stations	Spring mean MDA8 ozone		Summer mean MDA8 ozone	
	Linear trend $\mu g/m^3 * yr^{-1}$	p-value	Linear trend $\mu g/m^3 * yr^{-1}$	p-value
CMA	-0.43	0.839	-0.76	0.651
GCH	-5.68	0.031*	0.03	0.985
LFS	-3.34	0.016*	1.82	0.376
LAN	-1.74	0.247	-1.62	0.028*
SDZ	2.38	0.158	5.79	0.033*
WLG	0.94	0.057	1.17	0.098
XGLL	0.53	0.666	0.95	0.949

*: $p < 0.05$. The station AKDL was removed due to massive missingness. Spring: March-April-May; summer: June-July-August; Abbreviations: MDA8, daily maximum 8-hour average; WLG, Mt. Waliguan; SDZ, Shangdianzi; LAN, Lin'an; LFS, Longfengshan; XGLL, Xianggelila; AKDL, Akedala; GCH, Gucheng, CMA. China Meteorological Administration; TOAR: Tropospheric Ozone Assessment Report.

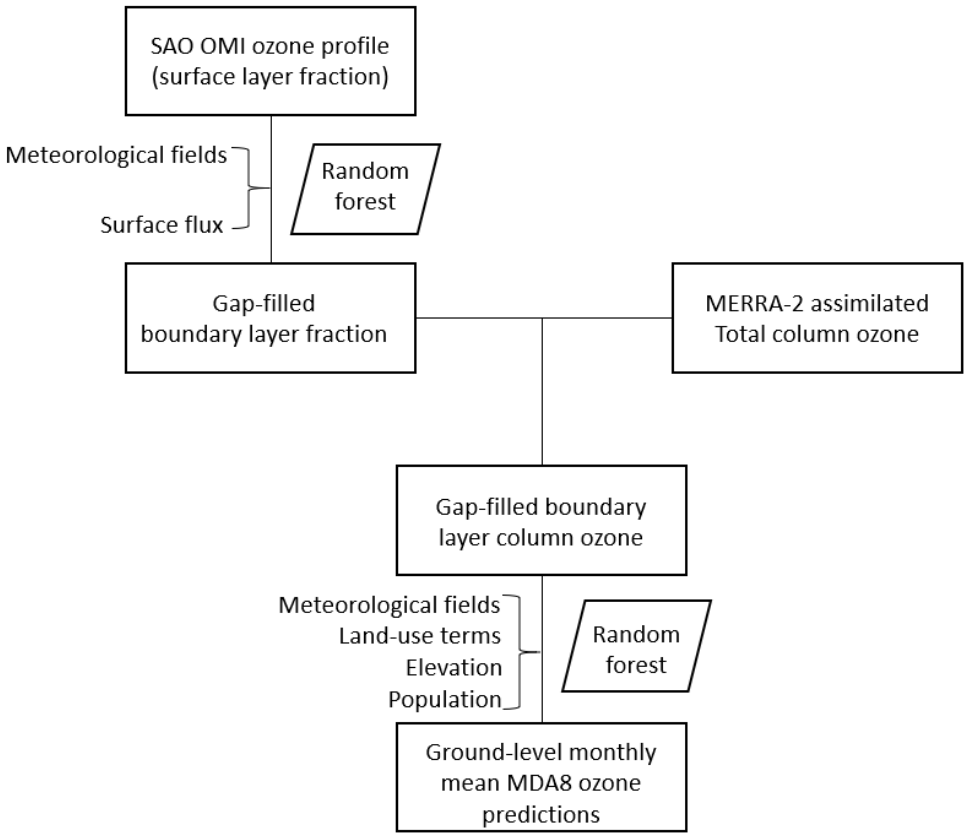

Fig.S1. Study workflow Abbreviations: SAO, Smithsonian Astrophysical Observatory; OMI, Ozone Monitoring Instrument; MERRA-2, the Modern-Era Retrospective analysis for Research and Applications, Version 2; MDA8, daily maximum 8-hour average.

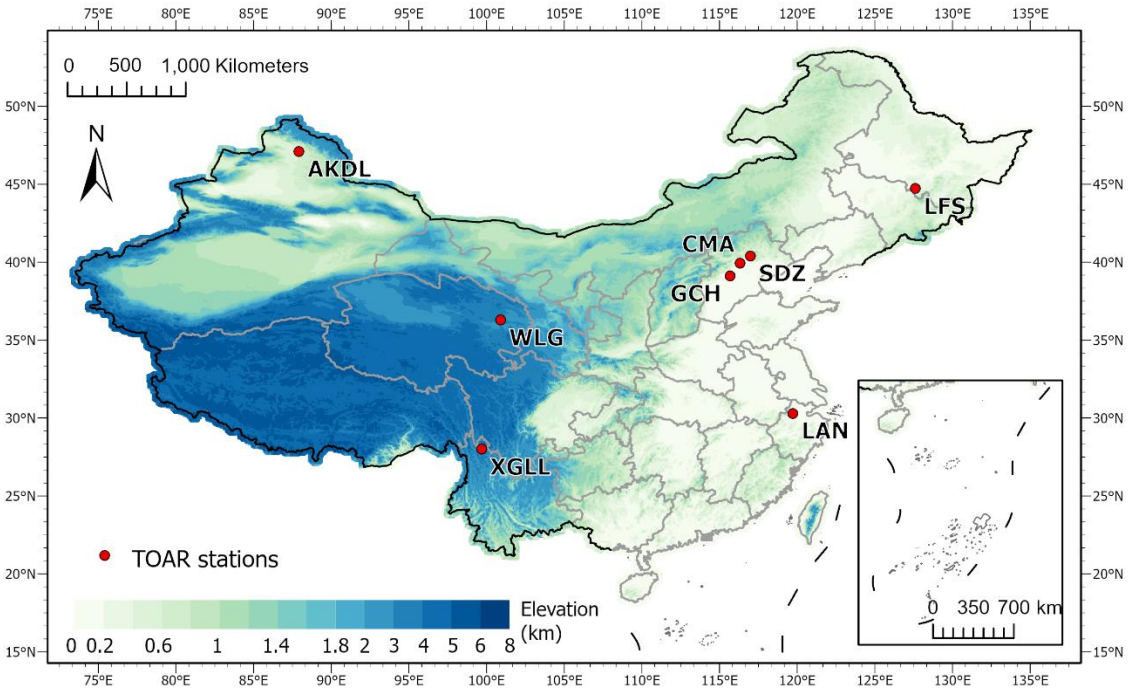

Fig. S2. Distribution of the Tropospheric Ozone Assessment Report (TOAR) monitoring stations. Red circle dots with black outlines represent the location of TOAR monitoring stations. Abbreviations: WLG, Mt. Waliguan; SDZ, Shangdianzi; LAN, Lin'an; LFS, Longfengshan; XGLL, Xianggelila; AKDL, Akedala; GCH, Gucheng, CMA. China Meteorological Administration; TOAR: Tropospheric Ozone Assessment Report.

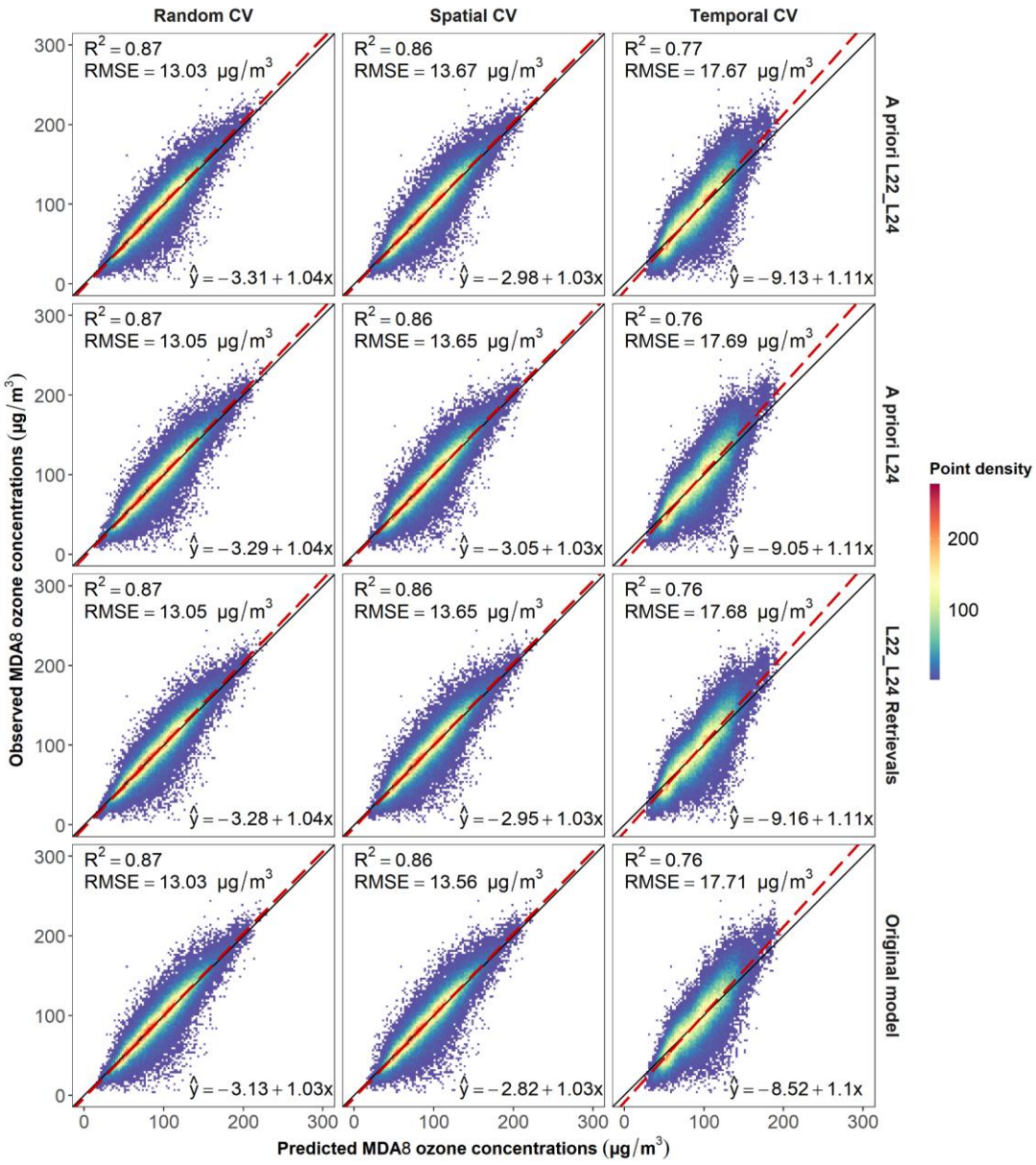

Fig. S3. Comparison of model performance on the training dataset (2014-2019) with different OMI-derived partial column ozone (original model). Left column: random CV; middle column: spatial CV; right column: temporal CV. Top row: model with a priori ozone profile L22-L24; second row (from top to bottom): model with a priori ozone profile L24; third row: model with L22-L24 retrievals; bottom row: the original model with retrieved ozone profile L24. The functions on the bottom-right corners are the regression functions between the predicted and observed monthly mean MDA8 ozone concentrations. Red dashed lines: the regression line between the predictions and observations; black solid lines: the x = y line; The color scale represents the density of the points. Abbreviations: MDA8, daily maximum 8-hour average; OMI, Ozone Monitoring Instrument.

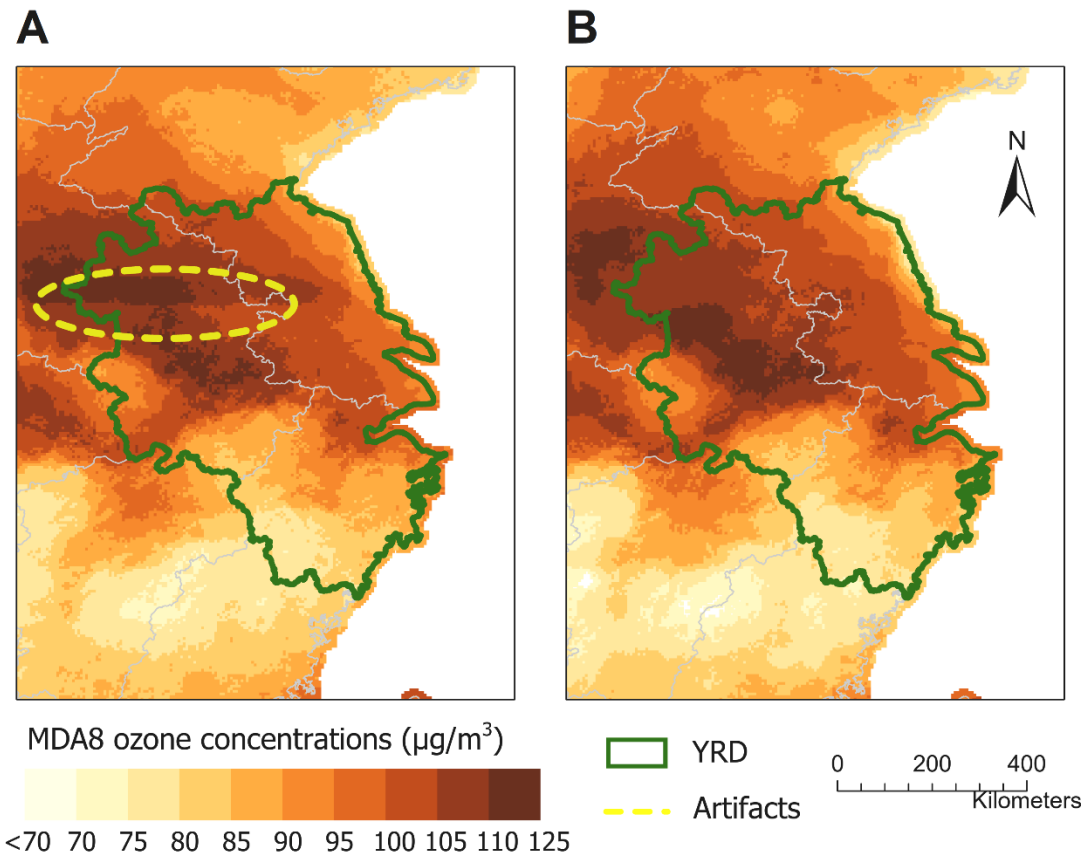

Fig. S4. An example of the spatial artifacts in the predicted monthly mean MDA8 ozone concentrations in the YRD (April 2019). Panel A was generated from the model with a priori ozone profile L22-L24; panel B was generated from the original model with OMI L24 retrieval. The green box represents YRD, and the yellow dashed line encircles an example of the spatial artifacts (the hard horizontal gap in predicted ozone concentrations). Abbreviations: MDA8, daily maximum 8-hour average; OMI, Ozone Monitoring Instrument; YRD, Yangtze River Delta.

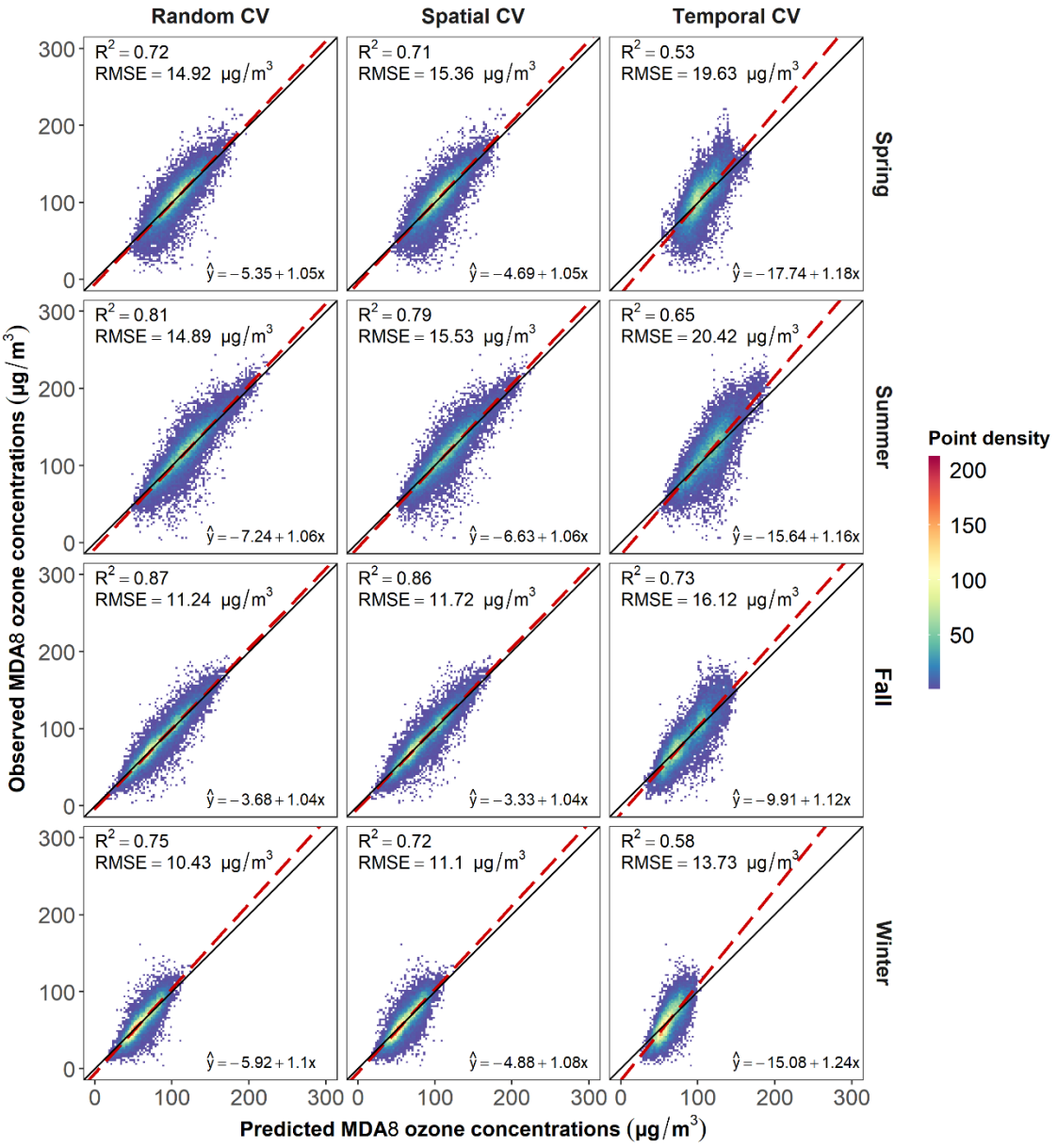

Fig.S5. Season-specific cross-validation results for the original model with OMI L24 retrieval. Left column: random CV; middle column: spatial CV; right column: temporal CV. Top row: spring (March-April-May); second row (from top to bottom): summer (June-July-August); third row: fall (September, October, November); bottom row: winter (December-January-February). The functions on the bottom-right corners are the regression functions between the predicted and observed MDA8 ozone concentrations. Red dashed lines: the regression line between the predictions and observations; black solid lines: the x = y line; The color scale represents the density of the points. Abbreviations: MDA8, daily maximum 8-hour average; OMI, Ozone Monitoring Instrument.

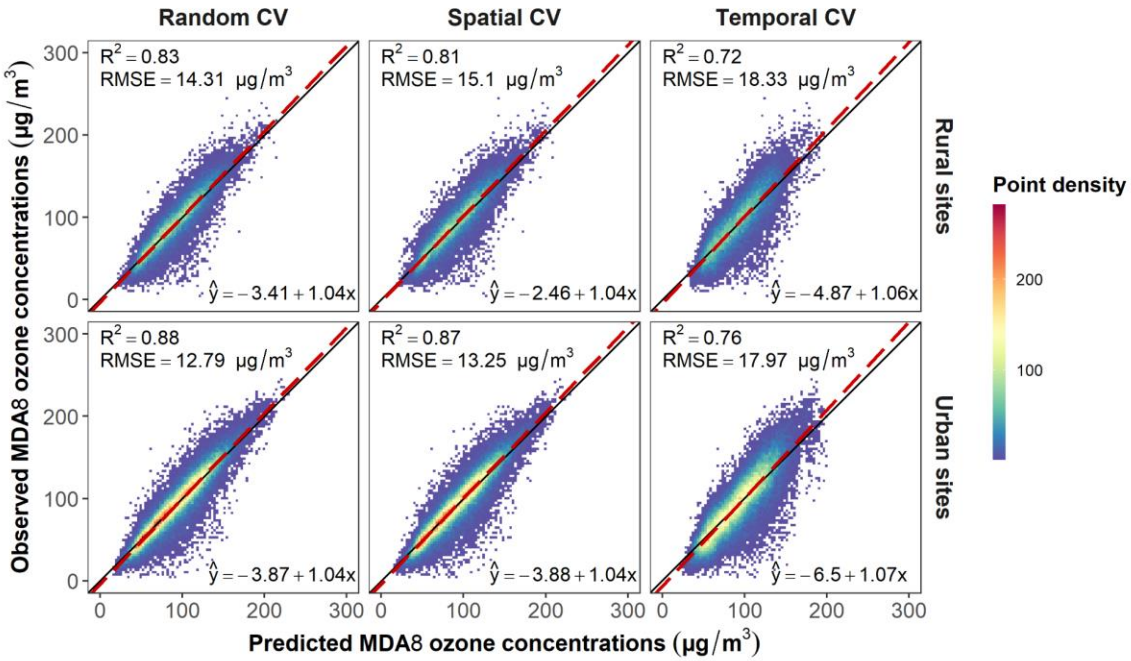

Fig.S6. Comparison of model performance in rural/urban areas using OMI L24 retrieval. Left column: random CV; middle column: spatial CV; right column: temporal CV. Upper row: model performance at rural CNEMC sites; Lower row: model performance in urban CNEMC sites. The functions on the bottom-right corners are the regression functions between the predicted and observed MDA8 ozone concentrations. Red dashed lines: the regression line between the predictions and observations; black solid lines: the x = y line; The color scale represents the density of the points. Abbreviations: CNEMC, China National Environmental Monitoring Centre; MDA8, daily maximum 8-hour average; OMI, Ozone Monitoring Instrument.

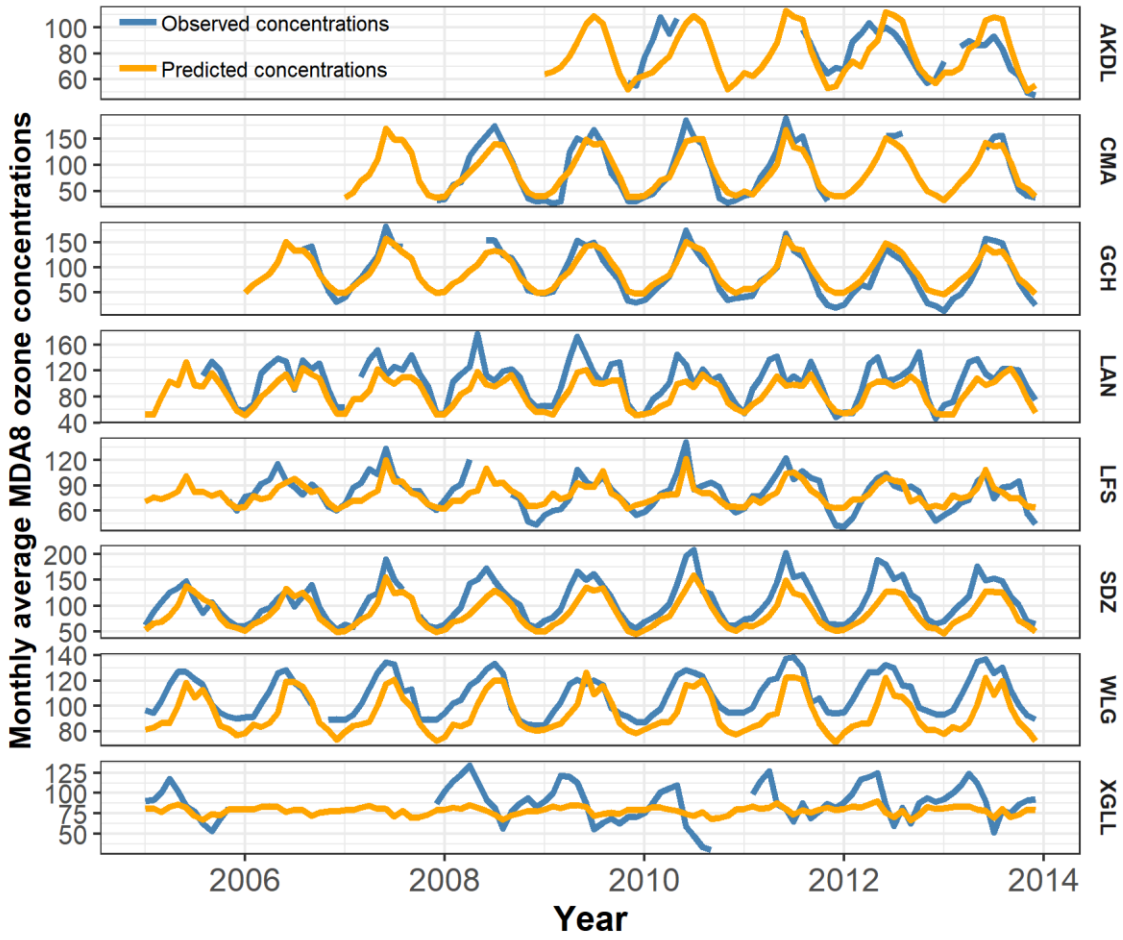

Fig.S7. Trend comparison between model predictions and the TOAR historical monitoring data (original model using OMI L24 retrieval). Each row represents a TOAR monitoring site. Blue lines: the observed monthly mean MDA8 ozone concentrations. Orange lines: the model predicted monthly MDA8 ozone concentrations. Abbreviations: MDA8, daily maximum 8-hour average; OMI, Ozone Monitoring Instrument. TOAR, Tropospheric Ozone Assessment Report.

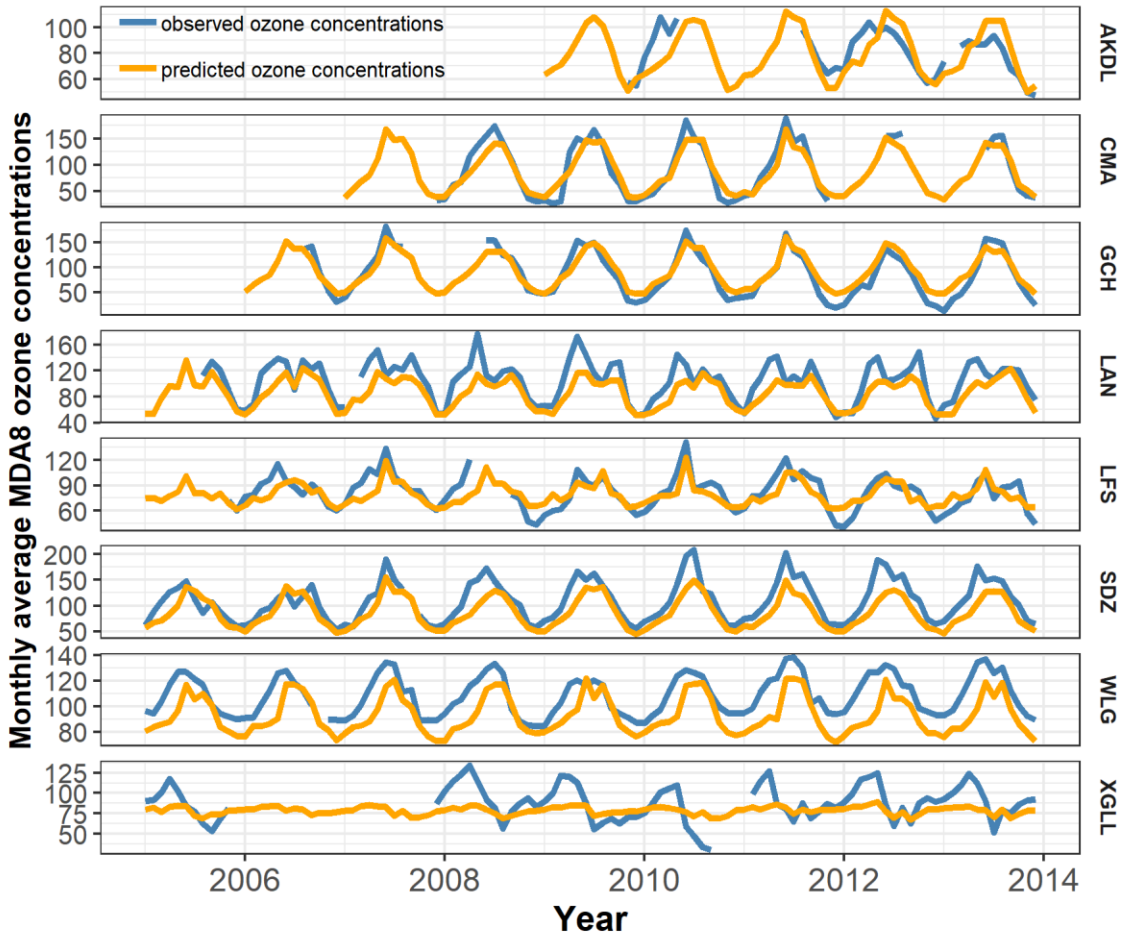

Fig.S8. Trend comparison between model predictions and the TOAR historical monitoring data (model trained with OMI L24 retrieval + MODIS FRP). Each row represents a TOAR monitoring site. Blue lines: the observed monthly mean MDA8 ozone concentrations. Orange lines: the model predicted monthly MDA8 ozone concentrations. Abbreviations: FRP, fire radiative power; MDA8, daily maximum 8-hour average; MODIS, Moderate Resolution Imaging Spectrometer; OMI, Ozone Monitoring Instrument. TOAR, Tropospheric Ozone Assessment Report.

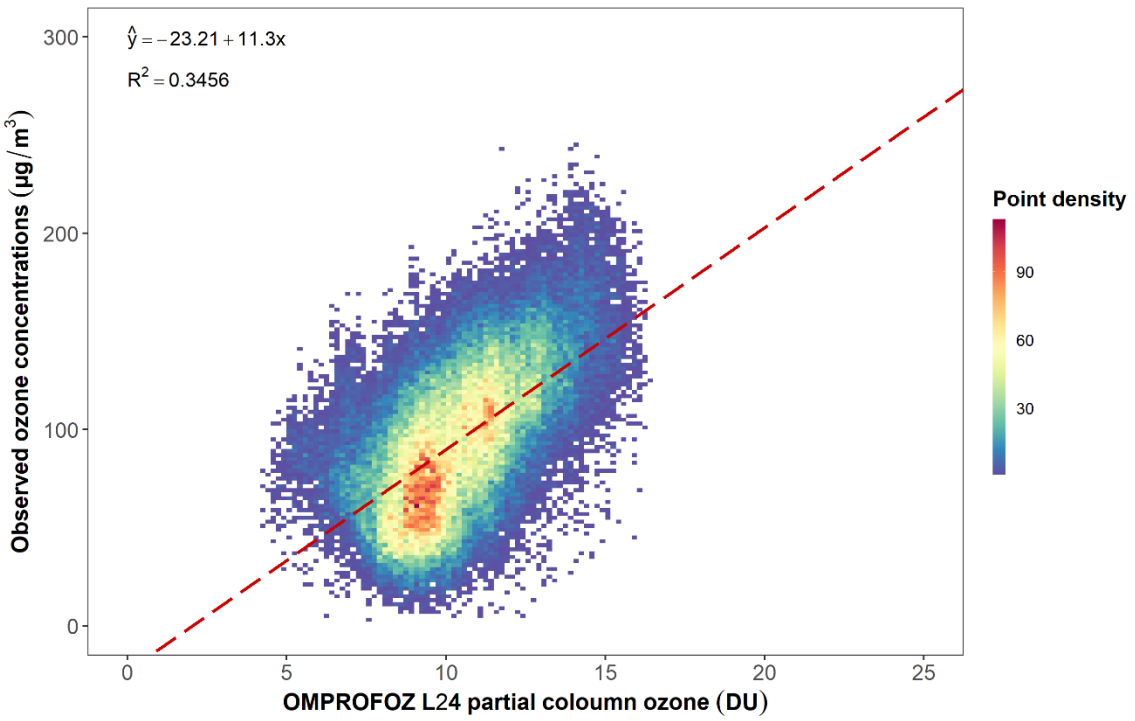

Fig.S9. Raw correlation between OMPROFOZ L24 partial column ozone and ground-level monitoring data. The function on the upper-left corner is the regression function between the OMPROFOZ L24 partial column ozone and the CNEMC surface MDA8 ozone concentrations. Red dashed line: the regression line as the regression function represents; black solid line: the x = y line; The color scale represents the density of the points. Abbreviations: CNEMC, China National Environmental Monitoring Centre; MDA8, daily maximum 8-hour average;

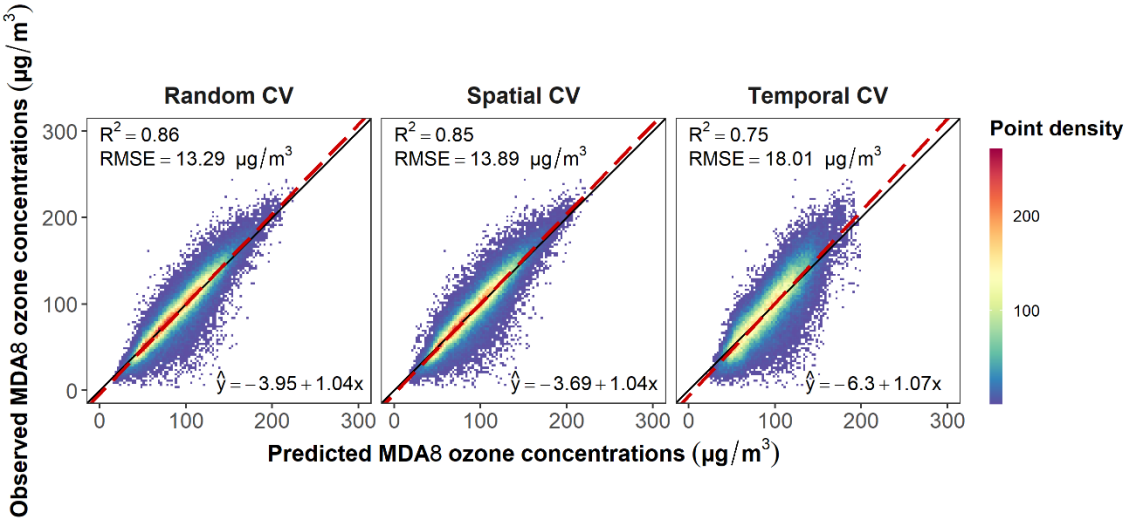

Fig.S10. model performance without the OMI L24 retrieval. Left panel: random CV; middle panel: spatial CV; right panel: temporal CV. The functions on the bottom-right corners are the regression functions between the predicted and observed monthly mean MDA8 ozone concentrations. Red dashed lines: the regression line between the predictions and observations; black solid lines: the x = y line; The color scale represents the density of the points. Abbreviations: CV, cross-validation; MDA8, daily maximum 8-hour average;

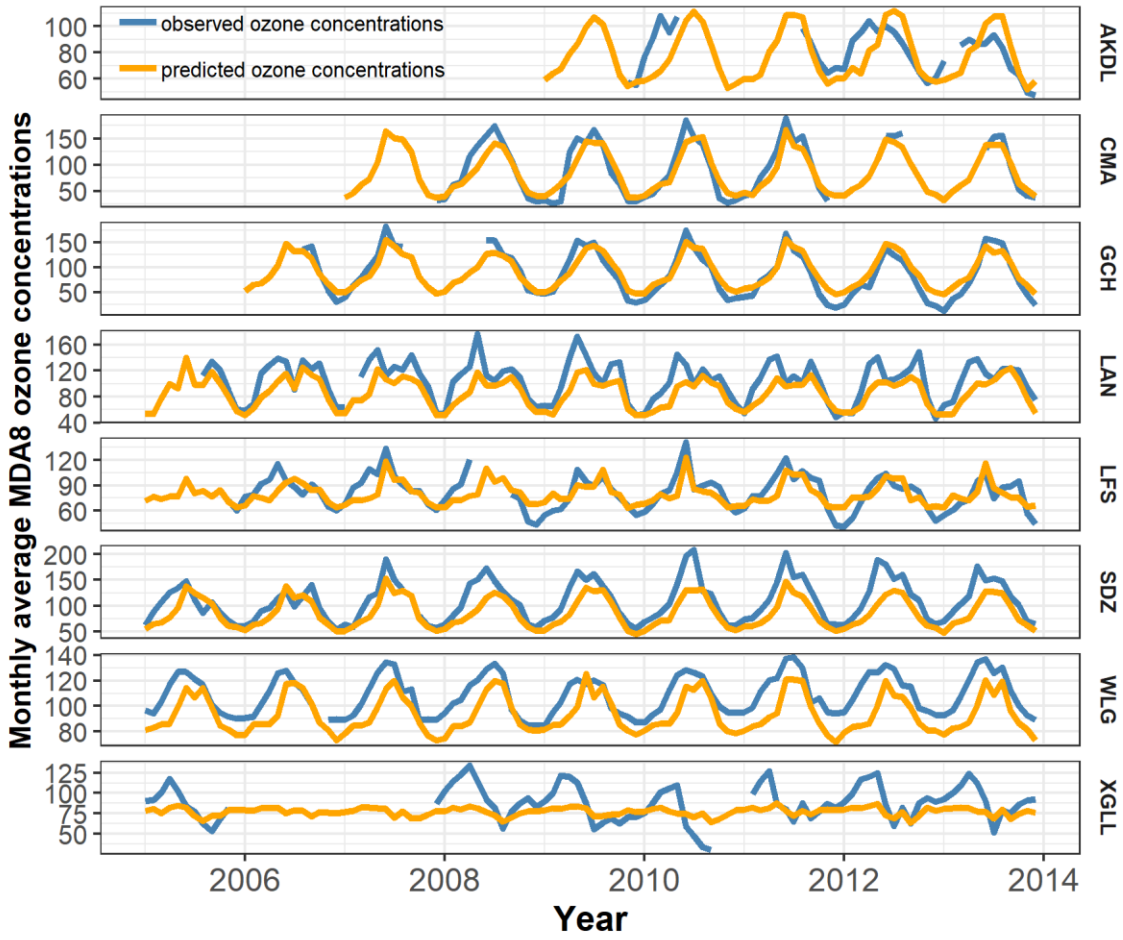

Fig.S11. Trend comparison between model predictions and the TOAR historical monitoring data (OMI-excluded model). Each row represents a TOAR monitoring site. Blue lines: the observed monthly mean MDA8 ozone concentrations. Orange lines: the model predicted monthly MDA8 ozone concentrations. Abbreviations: MDA8, daily maximum 8-hour average; OMI, Ozone Monitoring Instrument. TOAR, Tropospheric Ozone Assessment Report.

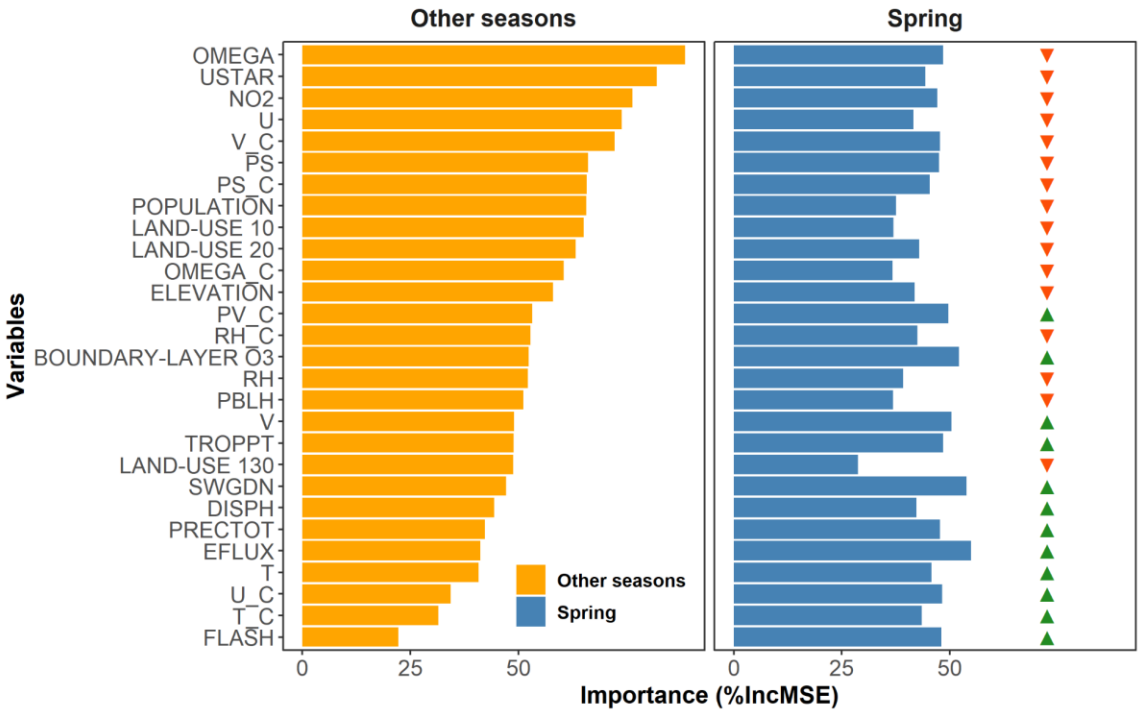

Fig. S12. The importance ranking of the input predictors. Orange bars in the left panel: the spring-excluded (all months except for March-April-May) model; blue bars in the right panel: the spring (March-April-May) model. The red downward arrows denoted that the relative ranking of the variable dropped in the spring model compared to other seasons; the blue upward arrows denoted that the relative ranking of a variable increased in spring. Variables' importance was generated with a permutation method.

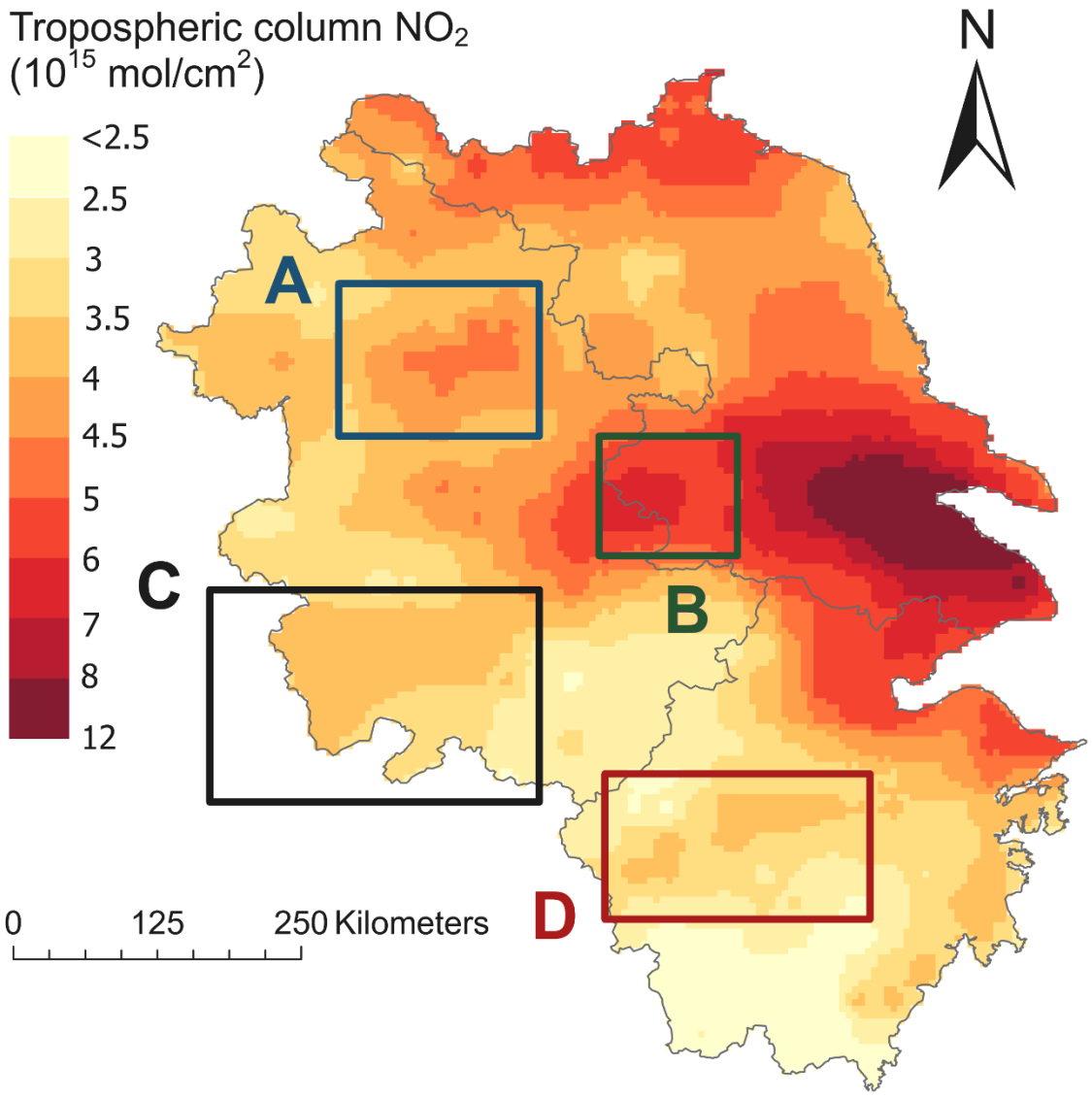

Fig.S13. The OMI tropospheric column NO_2 concentrations in the Yangtze River Delta region (August 2019, the marked regions were consistent to Fig. 5) The boxes represent some YRD cities and their surrounding area; A: Bengbu and Huainan City; B: Nanjing Metropolitan Area; C: Anqing City; D: Quzhou and Jinhua City. Abbreviations: MDA8, daily maximum 8-hour average.

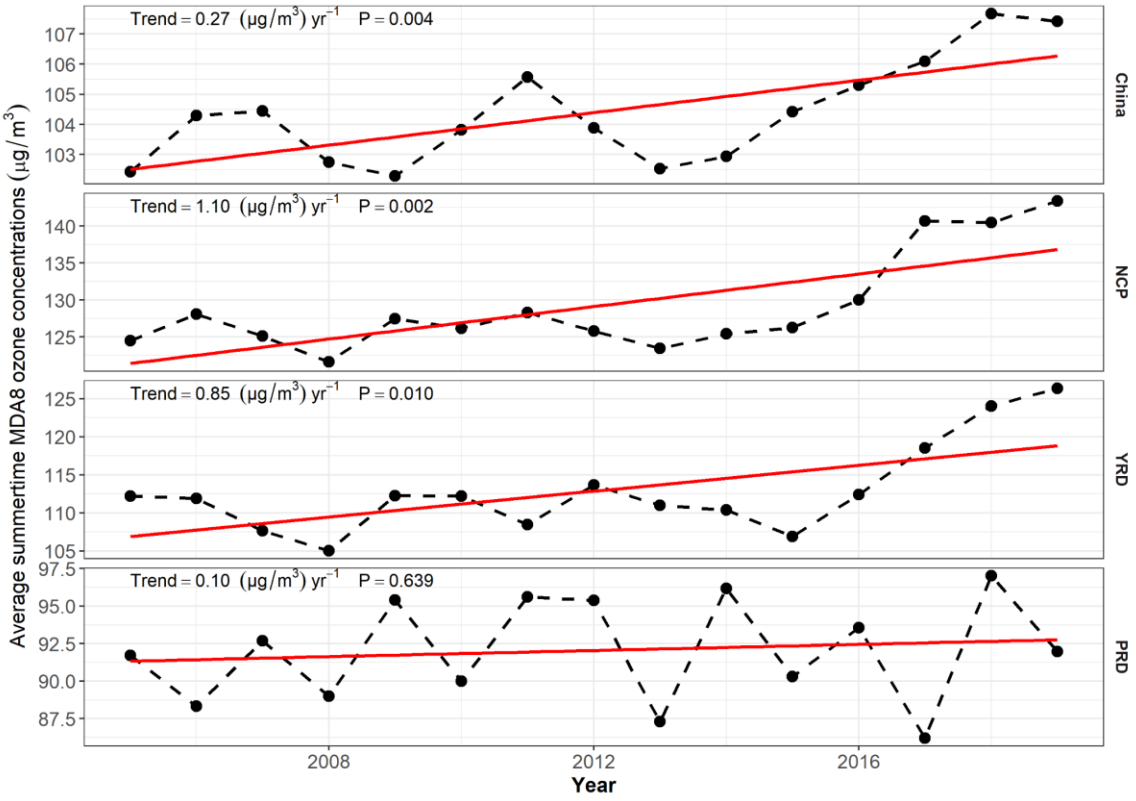

Fig.S14. The predicted long-term trend of summertime (June-July-August) mean MDA8 ozone concentrations over 2005-2019. The dots represented the predicted summertime mean MDA8 ozone concentrations; red lines are the linear smooth functions; P indicates the p-value of the linear trends. Abbreviations: MDA8, daily maximum 8-hour average; NCP: the North China Plain; PRD: the Pearl River Delta; YRD: the Yangtze River Delta.

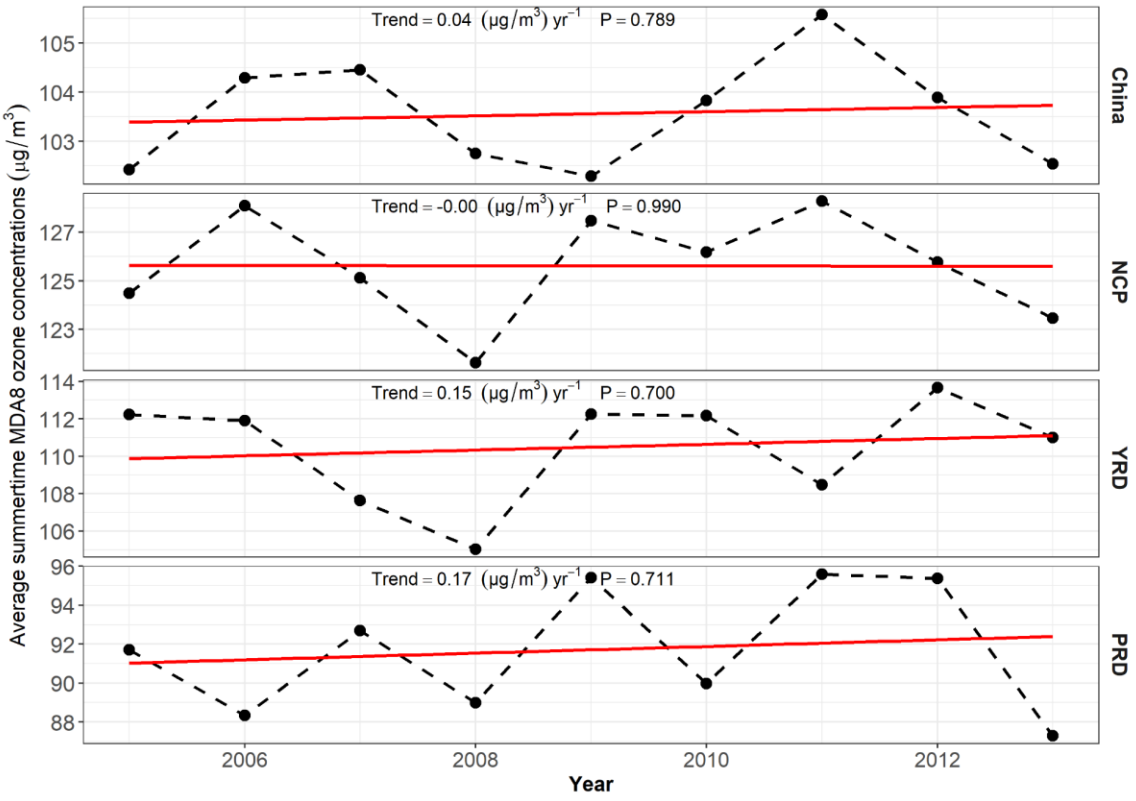

Fig.S15. The predicted long-term trend of summertime (June-July-August) mean MDA8 ozone concentrations over 2005-2013. The dots represented the predicted summertime mean MDA8 ozone concentrations; red lines are the linear smooth functions; P indicates the p-value of the linear trends. Abbreviations: MDA8, daily maximum 8-hour average; NCP: the North China Plain; PRD: the Pearl River Delta; YRD: the Yangtze River Delta.

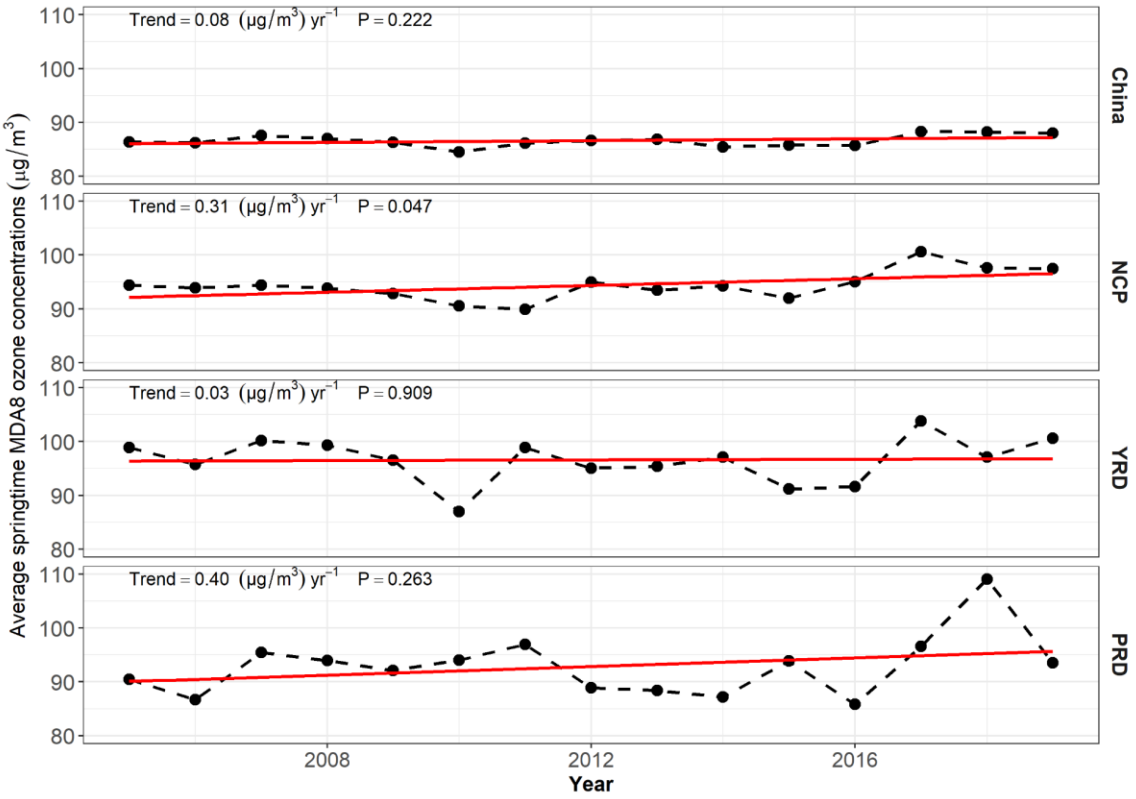

Fig.S16. The predicted long-term trend of springtime (March-April-May) mean MDA8 ozone concentrations over 2005-2019. The dots represented the predicted springtime mean MDA8 ozone concentrations; red lines are the linear smooth functions; P indicates the p-value of the linear trends. Abbreviations: MDA8, daily maximum 8-hour average; NCP: the North China Plain; PRD: the Pearl River Delta; YRD: the Yangtze River Delta.

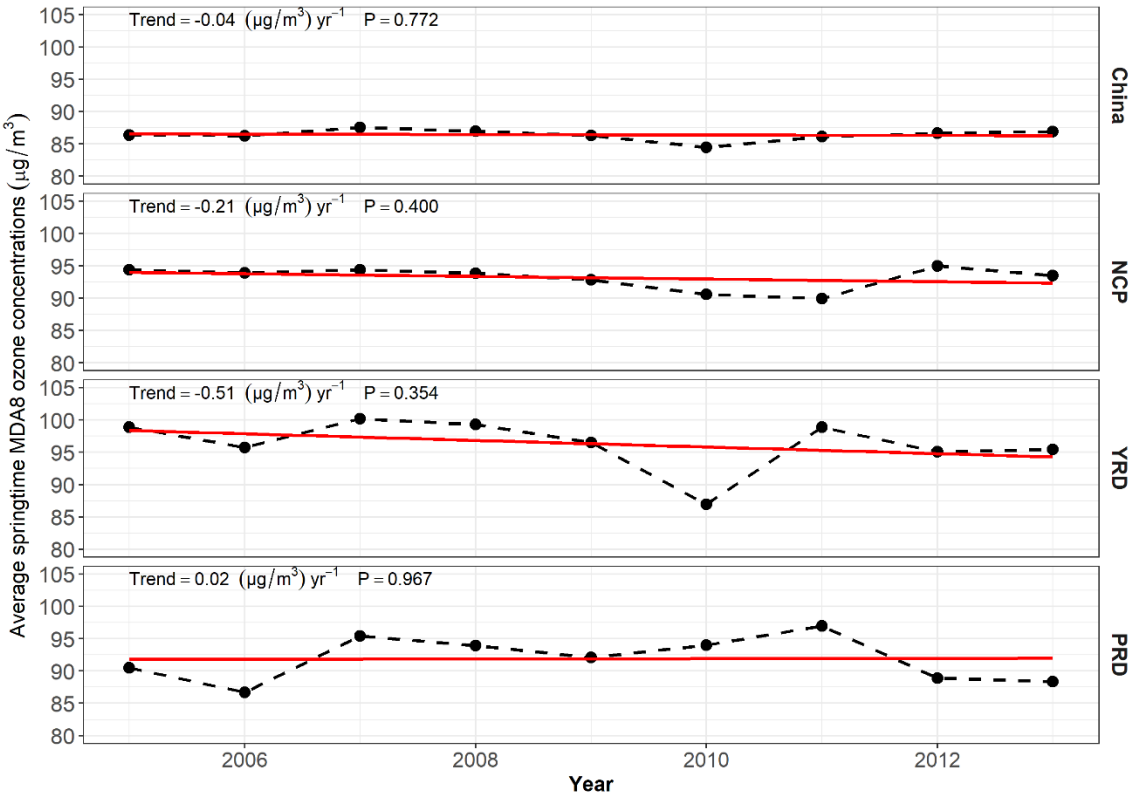

Fig.S17. The predicted long-term trend of springtime (March-April-May) mean MDA8 ozone concentrations over 2005-2013. The dots represented the predicted springtime mean MDA8 ozone concentrations; red lines are the linear smooth functions; P indicates the p-value of the linear trends. Abbreviations: MDA8, daily maximum 8-hour average; NCP: the North China Plain; PRD: the Pearl River Delta; YRD: the Yangtze River Delta.

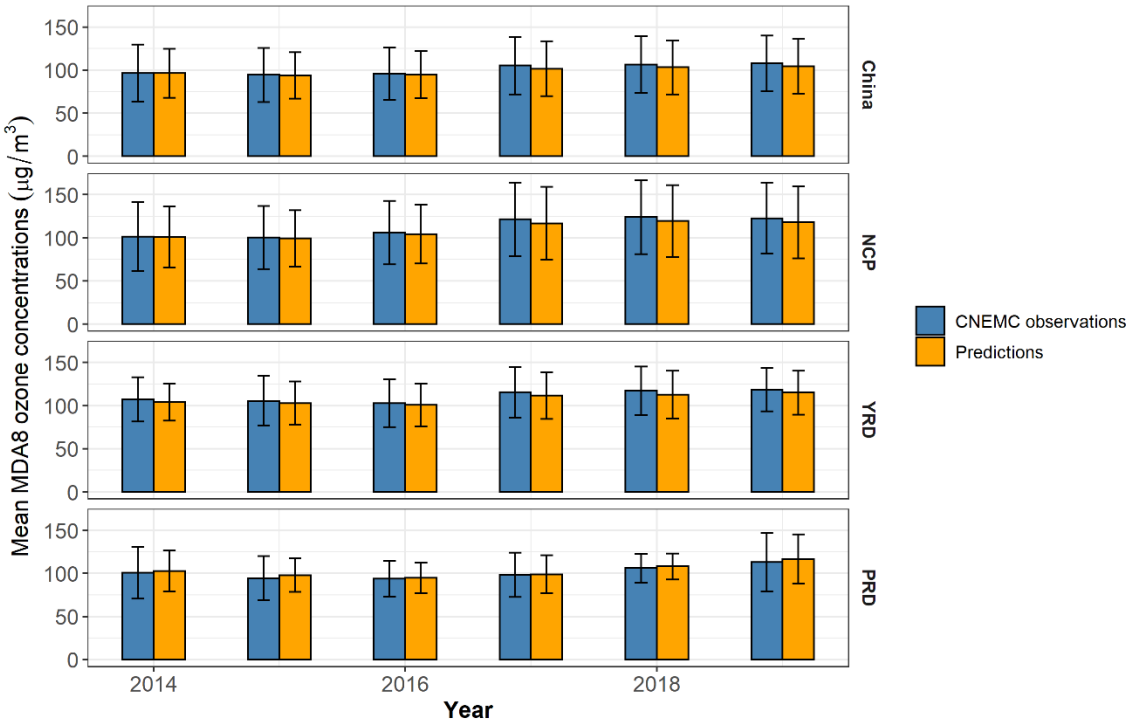

Fig. S18. The comparison of the mean ozone season (March-November) MDA8 ozone concentrations between our model predictions and the CNEMC monitoring data over 2014-2019. The blue columns on the left represent the CNEMC observations; the orange column on the right represents our model predictions. The height of the columns and the error bars represent the mean MDA8 ozone concentrations and the standard error. Abbreviations: CNEMC, China National Environmental Monitoring Centre; MDA8, daily maximum 8-hour average; NCP: the North China Plain; PRD: the Pearl River Delta; YRD: the Yangtze River Delta.

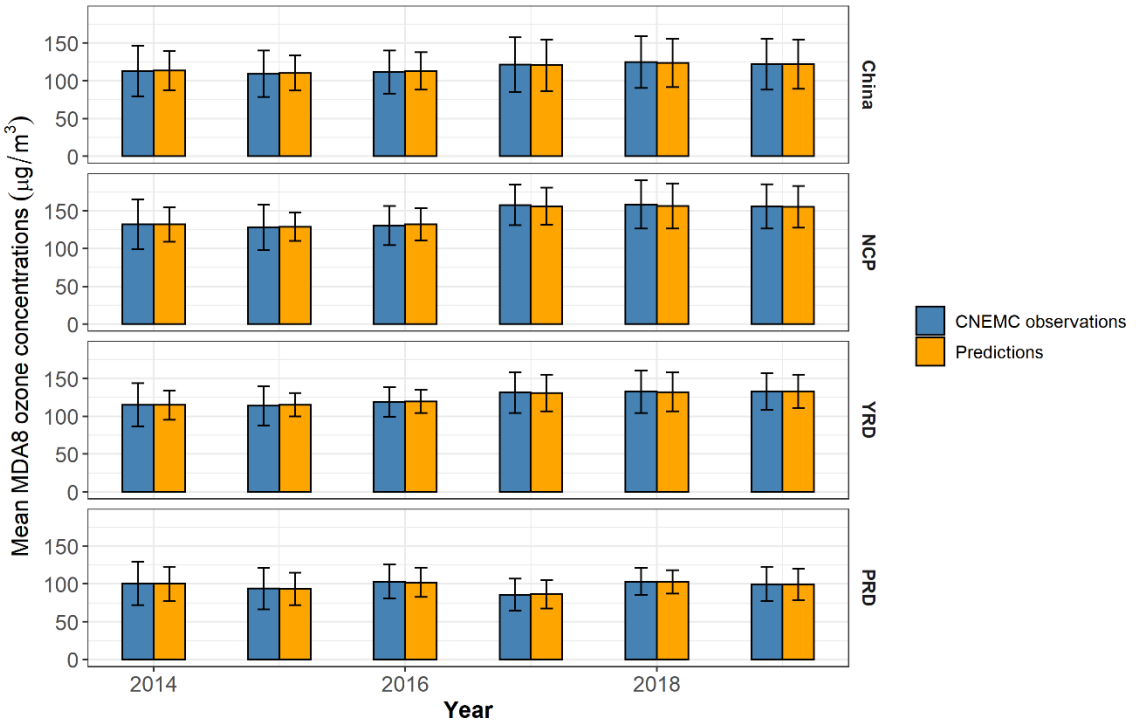

Fig. S19. The comparison of the summertime (June-July-August) mean MDA8 ozone concentrations between our model predictions and the CNEMC monitoring data over 2014-2019. The blue columns on the left represent the CNEMC observations; the orange column on the right represents our model predictions. The height of the columns and the error bars represent the mean MDA8 ozone concentrations and the standard error. Abbreviations: CNEMC, China National Environmental Monitoring Centre; MDA8, daily maximum 8-hour average; NCP: the North China Plain; PRD: the Pearl River Delta; YRD: the Yangtze River Delta.

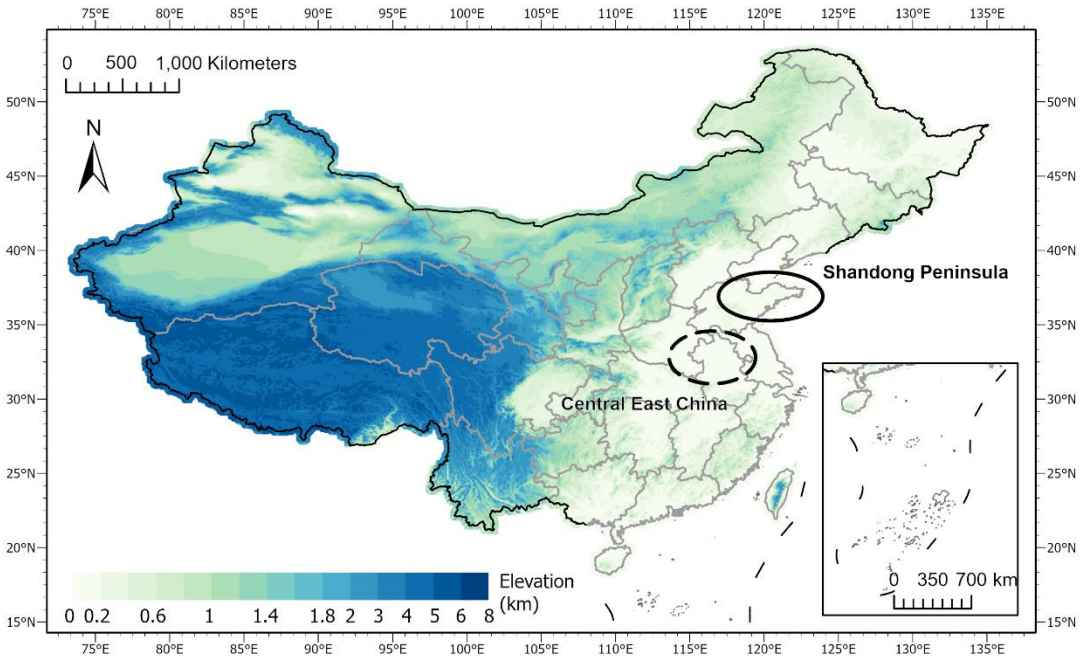

Fig. S20 Locations of the Shandong Peninsula and Central East China. The solid line encircles the Shandong Peninsula, while the dashed line encircles the Central East China area.