1   # Climate-driven deterioration of future ozone pollution in

2   # Asia predicted by machine learning with multisource data

4   Huimin Li[1], Yang Yang[1*], Jianbing Jin[1], Hailong Wang[2], Ke Li[1], Pinya Wang[1],

5   Hong Liao[1]

9   [1]Jiangsu Key Laboratory of Atmospheric Environment Monitoring and

10   Pollution Control, Jiangsu Collaborative Innovation Center of Atmospheric

11   Environment and Equipment Technology, School of Environmental Science

12   and Engineering, Nanjing University of Information Science and Technology,

13   Nanjing, Jiangsu, China

14   [2]Atmospheric Sciences and Global Change Division, Pacific Northwest

15   National Laboratory, Richland, Washington, USA

19   *Correspondence to yang.yang@nuist.edu.cn

Abstract

21   Ozone (O$_3$) is a secondary pollutant in the atmosphere formed by

22 photochemical reactions that endangers human health and ecosystems. O$_3$ has

23 aggravated in Asia in recent decades and will vary in the future. In this study, to

24 quantify the impacts of future climate change on O$_3$ pollution, near-surface O$_3$

25 concentrations over Asia in 2020–2100 are projected using a machine learning

26 (ML) method along with multisource data. The ML model is trained with

27 combined O$_3$ data from a global atmospheric chemical transport model and real-

28 time observations. The ML model is then used to estimate future O$_3$ with

29 meteorological fields from multi-model simulations under various climate

30 scenarios. The near-surface O$_3$ concentrations are projected to increase by 5–

31 20% over South China, Southeast Asia, and South India and less than 10%

32 over North China and Gangetic Plains under the high forcing scenarios in the

33 last decade of 21st century, compared to the first decade of 2020–2100. The O$_3$

34 increases are primarily owing to the favorable meteorological conditions for O$_3$

35 photochemical formation in most Asian regions. We also find that the

36 summertime O$_3$ pollution over eastern China will expand from North China to

37 South China and extend into the cold season in a warmer future. Our results

38 demonstrate the important role of climate change penalty on Asian O$_3$ in the

39 future, which provides implications for environmental and climate strategies of

40 adaptation and mitigation.

1. Introduction

41

42   Tropospheric ozone (O_3) is a primary air pollutant, formed by

43 photochemical oxidation of nonmethane volatile organic compounds (NMVOCs)

44 and carbon monoxide (CO) in the presence of nitrogen oxides (NO_x = NO +

45 NO_2) and sunlight. It has adverse effects on human health (Malley et al., 2017;

46 Cakmak et al., 2018), vegetation growth (Yue et al. 2017; Mills et al., 2018) and

47 climate change (Checa-Garcia et al., 2018; Gaudel et al., 2018). A better

48 understanding of the causes of changes in O_3 concentrations is useful for

49 developing effective environment and climate strategies.

50   Since mid-1990s, Asian regions, including South Asia, East Asia and

51 Southeast Asia, have experienced the fastest O_3 increase rate of 2–8

52 ppb/decade at remote surface sites and in the lower free troposphere across

53 the world (IPCC, 2021). A number of air quality monitoring stations administered

54 by China National Environmental Monitoring Center (CNEMC) have been

55 established in China since 2013 to measure real-time near-surface particulate

56 matter, O_3, and other air pollutants. The measurements showed an increasing

57 trend of urban warm-season daily maximum 8-hour average (MDA8) O_3

58 concentrations of 2.4 ppb (5%) yr^{-1} that is faster than any other regions

59 worldwide during 2013–2019 (Lu et al., 2020). However, many regions in Asia

60 lack O_3 observations with sufficient spatial and temporal coverage. Also, most

61 of the present regional observations are collected only near population clusters,

62 which are not representative of the entire region (Zhou et al., 2022).

63   To supplement the limited near-surface O_3 measurements, many studies

64   utilized global and regional models with comprehensive physical and chemical

65   processes to simulate O_3 concentrations (Zhu et al., 2017; Gao et al., 2020;

66   Yang et al., 2022). Moreover, statistical models have also been used to estimate

67   O_3 concentrations (Chen et al., 2020; Zhang et al., 2020). In recent years,

68   machine learning (ML) approaches, such as random forest (Xue et al., 2020;

69   Wei et al., 2022), neural network (Di et al., 2017), support vector machine (Su

70   et al., 2020), extreme gradient boosting (Li et al., 2020), and ensemble learning

71   (Liu et al., 2022), were widely applied to estimate O_3 levels based on potential

72   influential factors (e.g., precursor emissions, meteorological conditions, land

73   use, surface elevation, gross domestic product, population density, and

74   geographical variables).

75   Meteorological factors and synoptic conditions play important roles in

76   affecting O_3 pollution (Fu and Tai, 2015; Gong and Liao, 2019; Yin et al., 2019;

77   Liu et al., 2020; Dang et al., 2021). Gong et al. (2019) illustrated that hot, dry,

78   and stagnant weather conditions are favorable for the formation and

79   persistence of severe O_3 pollution over northern China. High air temperature

80   along with intense incoming shortwave radiation accelerates both

81   photochemical reaction rates and natural precursor emissions for O_3 production

82   (Jacob and Winner, 2009). Under high relative humidity conditions, O_3

83   concentrations decrease due to many complex physical and chemical

84   mechanisms (Jeong and Park, 2013; Kavassalis and Murphy., 2017; Lu et al.,

85   2019; Li et al., 2021). Cloud and precipitation impact O_3 levels through reducing

86   the downwelling solar radiation and washout of pollutants (Toh et al., 2013).

87   Anomalous sea level pressure patterns can affect the long-range transport of

88   O_3 by influencing atmospheric circulation (Santurtún et al., 2015). By changing

89   the air stagnant condition and transport of pollutants, wind fields can also affect

90   O_3 concentrations in local and downwind areas of emission sources (Doherty

91   et al., 2013).

92   Future climate change corresponding to the different climate scenarios can

93   impact O_3 through altering meteorological conditions (Wang et al., 2013, Fu and

94   Tian et al., 2019). Using regional climate fields downscaled from general

95   circulation models to investigate potential O_3 variations in the U.S. due to

96   changing climate, Fann et al. (2015) projected the MDA8 O_3 to increase by 1–

97   5 ppb as daily maximum average temperature increases by 1–4°C in 2030

98   relative to 2000. Colette et al. (2015) estimated that the climate penalty for

99   future summertime surface O_3 reaches 0.99–1.5 ppb by the end of the 21st

100   century (2071–2100) in Europe compared to present-day levels using an

101   ensemble of eight global coupled climate-chemistry models under the RCP

102   (Representative Concentration Pathway) 8.5 scenario. Through fixing sea

103   surface temperature at present-day and future conditions in five atmospheric-

104   only models as part of the AerChemMIP (Aerosol Chemistry Model

105   Intercomparison Project), Zanis et al. (2022) projected the climate change

106   penalties and benefits on global surface O_3 concentrations from 2015 to 2100

107 under the SSP3-7.0 scenario. They found O_3 reductions in most regions of the

108 globe, except a robust O_3 climate penalty of 1–2 ppb °C^{-1} in South and East

109 Asia under global warming following the SSP3-7.0 pathway. However, SSP3-

110 7.0 is not a good representative scenario for both air quality and climate in Asia.

111 The emissions of greenhouse gases (GHGs) and air pollutants over East Asia

112 in SSP3-7.0 are assumed to significantly increase in the near future and keep

113 at high levels in the middle of the 21st century among all SSPs (Li et al., 2022),

114 while the emissions of air pollutants have been cut by a lot since 2010s in the

115 real world (Wang et al., 2021). The GHGs and pollutant emissions are very likely

116 to continually decline in the future related to the carbon neutrality commitment

117 (Cheng et al., 2021).

118   In this study, we aim to better characterize the impact from future climate

119 change on Asian O_3 pollution using multiple state-of-the-art modeling tools and

120 data. It is important for policy-makers that mitigating global climate change

121 potentially has positive benefits to surface air quality through meteorological

122 factors, not only the reduction in fossil fuel co-emissions. The near-surface O_3

123 concentrations covering 2020–2100 in Asia are projected using a ML method

124 integrated with multisource data, including assimilated O_3 data that combine

125 ground observations and simulations from a global 3-D chemical transport

126 model (GEOS-Chem), meteorological fields under various climate scenarios

127 from the latest Coupled Model Intercomparison Project Phase 6 (CMIP6) multi-

128 model simulations, and other auxiliary data (e.g., emissions, land use,

129   topography, population density, and spatiotemporal information). ML approach

130   gives the capacity to explore many scenarios more rapidly and for longer time

131   periods than the chemical transport model process-based modeling. Details of

132   the data and methodology used in this study are described in section 2. Section

133   3 analyzes the results of climate-driven O_3 variations over different key regions

134   of Asia. Section 4 summarizes the main conclusions and discusses potential

135   uncertainties in this study.

136   **2. Materials and Methods**

137   **2.1 GEOS-Chem model description**

138   To assimilate O_3 data for the ML model training, the near-surface O_3

139   concentrations from 2014 to 2019 are firstly simulated using the nested-grid

140   version of the 3-D GEOS-Chem model (version 12.9.3), driven by the Modern-

141   Era Retrospective analysis for Research and Applications, Version 2 (MERRA-

142   2) reanalysis meteorological data (Gelaro et al., 2017). The nested GEOS-

143   Chem has 47 vertical layers from the surface up to 0.01 hPa, with a horizontal

144   resolution of 0.5° latitude × 0.625° longitude over the Asia domain (11°S–55°N,

145   60–150°E). The lateral boundaries of chemical tracer concentrations are

146   provided by global simulations at 2° latitude × 2.5° longitude horizontal

147   resolution. The model includes fully coupled aerosol-O_3-NO_x-hydrocarbon

148   chemical mechanisms (Park et al., 2004; Pye et al., 2009; Mao et al., 2013),

149   with about 300 species participated in over 400 kinetic and photochemical

150   reactions (Bey et al., 2001). The stratospheric O_3 chemistry is simulated

151   through linearized O_3 parameterization scheme (LINOZ; Mclinden et al., 2000),

152   and the planetary boundary layer mixing is calculated by a nonlocal scheme

153   (Lin and McElroy, 2010). GEOS-Chem has shown a good performance in

154   reproducing spatiotemporal distributions of O_3 concentrations (e.g., Ni et al.,

155   2018; Li et al., 2019).

156   The historical (2014–2019) anthropogenic emissions of O_3 precursor

157   gases, including NO_x, NMVOCs, and CO, utilized in the nested domain are

158   obtained from the Community Emissions Data System (CEDS; Hoesly et al.,

159   2018) version 2021_04_21, which fully considered the recent emission

160   reductions in China related to clean air measures. The biomass burning

161   emissions are acquired from the Global Fire Emissions Database version 4

162   (GFED4; van der Werf et al., 2017). Biogenic emissions of NMVOCs from the

163   Model of Emissions of Gases and Aerosols from Nature (MEGAN) version 2.1

164   are employed, with updates from Guenther et al. (2012). Soil NO_x sources are

165   calculated with an updated version of the Berkeley-Dalhousie Soil NO_x

166   Parameterization scheme (Hudman et al., 2012). NO_x emissions from lightning

167   are as described by Murray et al. (2012), and the vertical distribution of

168   emissions follows Ott et al. (2010).

169   **2.2 Ground O_3 observations**

170   To improve the performance of the ML model in predicting O_3

171   concentrations, the nationwide hourly near-surface O_3 concentrations in China

172   during 2014–2019 are obtained from the CNEMC and used for O_3 data

173 assimilation. The observational network had about 500 monitoring sites in 2013,

174 and expanded to more than 1500 sites after 2019, covering 360 cities in

175 mainland China. The O_3 observations are averaged within each 0.5° latitude ×

176 0.625° longitude grid of the GEOS-Chem model.

177 **2.3 Data assimilation**

178 The assimilation system, which is used to combine the O_3 observations

179 with results from GEOS-Chem simulations, is based on a three-dimensional

180 variational (3DVar) data assimilation. The goal of the 3DVar is to find the

181 maximum likelihood estimation of a state vector x, which is the O_3

182 concentrations here in this study, given the available observations y through

183 minimizing the cost function:

184
$$J(x) = \frac{1}{2}(x - x^b)^\mathrm{T}\mathbf{B^{-1}}\,(x - x^b) \; + \; \frac{1}{2}(y - \mathrm{H}(x))^\mathrm{T}\mathbf{O^{-1}}\,(y - \mathrm{H}(x)\,)$$

185 Here x^b represents the priori simulation, **B** is the empirical background

186 covariance matrix representing the uncertainty. **H** denotes the operator that

187 converts the simulation results into the observation space, and **O** is a diagonal

188 covariance matrix storing the uncertainty of the measurements used.

189 Comparisons between observed and assimilated O_3 concentrations over

190 2014–2019 are shown in Figure 1. The overall correlation coefficient (R) is 0.94,

191 and the normalized mean bias (NMB) is –0.1%, suggesting that the assimilated

192 data have an excellent representation of O_3 observations and minimize the

193 uncertainties of GEOS-Chem simulations.

194 **2.4 Predicting O_3 using a machine learning method**

195 In this study, a random forest (RF) model is used to predict O_3

196 concentrations, similar to our previous studies (Li et al., 2021, 2022), with input

197 data of assimilated O_3 concentrations that combine observations and results

198 from GEOS-Chem model simulations, MERRA-2 meteorological variables, O_3

199 precursor emissions, land cover (LC), normalized difference vegetation index

200 (NDVI), topography (TOPO), population density (POP), and the month of the

201 year (MOY) and geographic location of each model grid as spatiotemporal

202 information. Details of the datasets are summarized in Table 1.

203 For predicting future climate-driven near-surface O_3 concentrations, the ML

204 model is trained with samples over 2014–2018 and the remaining 2019 data

205 are used for model validation. To obtain an optimal ML model, hyperparameters

206 are firstly tuned using the 10-fold cross-validation (Rodroguez et al., 2010).

207 Several statistical metrics, including coefficient of determination (R^2), mean

208 absolute error (MAE), root mean square error (RMSE) and mean relative error

209 (MRE) are used to evaluate the performance of ML model. Then the climate-

210 driven near-surface O_3 concentrations during 2020–2100 under four SSPs

211 (SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5) in Asia can be estimated using

212 the trained ML model with varying meteorological factors under the climate

213 change scenarios. Both anthropogenic and natural emissions of O_3 precursors

214 are fixed at the present-day levels for the prediction.

215 **2.5 Meteorological fields from CMIP6 multi-model simulations**

216 Monthly meteorological parameters under four different future climate

217   scenarios, including SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5, are fed to

218   a ML model to predict ground-level O_3 concentrations. The Scenario Model

219   Intercomparison Project (ScenarioMIP) as part of CMIP6 provides multi-model

220   projections of climate variables driven by future emission and land use changes

221   under different SSPs (O'Neill et al., 2016). In our study, meteorological fields,

222   such as air temperature (at 2m, 850 hPa, and 500 hPa), wind fields (at 850 and

223   500 hPa), surface relative humidity, incoming shortwave radiation at the surface,

224   total cloud cover, precipitation rate, and sea level pressure, are chosen as the

225   key meteorological predictors for ground-level O_3 concentrations, which are

226   obtained from 18 global climate models, i.e., ACCESS-CM2, ACCESS-ESM1-

227   5, CanESM5, CESM2-WACCM, CMCC-CM2-SR5, EC-Earth3-Veg, EC-Earth3,

228   FGOALS-f3-L, FGOALS-g3, GFDL-ESM4, INM-CM5-0, IPSL-CM6A-LR,

229   MIROC6, MPI-ESM1-2-HR, MPI-ESM1-2-LR, MRI-ESM2-0, NorESM2-LM,

230   and NorESM2-MM. Before being applied to the ML model, future meteorological

231   fields from ScenarioMIP are adjusted to minimize the inconsistencies in the

232   initial conditions in models and reanalysis data following Li et al. (2022).

233   **3. Results**

234   **3.1 Predictive capability of the machine learning model**

235   The ML predicted monthly O_3 concentrations over Asia in 2019 by the ML

236   model are in good agreement with the assimilated O_3 data constructed with

237   observations and GEOS-Chem model results (Fig. 2). The overall R^2 between

238   the predicted and assimilated O_3 concentrations is as high as 0.92 and the ML

239   model has a low MRE of 9% in predicting O_3 concentrations over the Asia

240   domain. Overall, these statistical indices indicate that the RF model is promising

241   for predicting the spatial distributions and temporal variations of near-surface

242   O_3 concentrations over Asia, which can provide a practical means for studying

243   long-term variations in O_3 under the future climate change.

244   Meanwhile, the ML model predictive capability for each grid cell over the

245   entire domain during 2014–2019 is further evaluated and demonstrated in

246   Figure 3. Regarding the spatial performance, the estimated O_3 concentrations

247   are highly correlated to the assimilated data in most regions of Asia with small

248   biases, indicating a strong spatial predictive ability of the RF model. More than

249   80% of land areas have a R^2 greater than 0.9. In terms of model uncertainties,

250   about 95% of land areas have a RMSE (MAE) less than 3 (2) parts per billion

251   (ppb). Furthermore, approximately 86% of land areas show small modeling bias

252   with MRE below 5%. Note that several grid cells show MRE over 5% but still

253   below 15%, which is related to the data assimilation using monitored and

254   simulated O_3 concentrations in China and the coarse resolution for coastal

255   areas and islands over Southeast Asia.

256   Figure 4 shows the importance score of independent variables that

257   contribute to the prediction of trained ML model. The results suggest that among

258   all the input predictors, relative humidity, incoming solar radiation at the surface,

259   and topography are the top-three most influential variables for the model

260   construction of near-surface O_3 in Asia, with importance scores of 15%, 12%

261 and 10%, respectively. The primary importance of relative humidity has also

262 been reported in previous studies (e.g., Han et al., 2020; Qian et al., 2022).

263 Other meteorological parameters, such as cloud cover, sea level pressure, air

264 temperature, precipitation, also have a substantial impact on the O_3 estimates,

265 with importance scores ranging from 4% to 8%. In the ML model, the emissions

266 of three primary O_3 precursors, including NMVOCs, NO_x, and CO, have a

267 relatively low importance score of 4–5% individually due to the spatiotemporal

268 diversity of O_3 production regimes.

269 **3.2 Predicted future climate-driven O_3 variations**

270 Figure 5 shows the predicted absolute and percentage changes in annual

271 mean surface O_3 concentrations in response to climate change between the

272 first and last decades of 2020–2100 based on the future meteorological fields

273 from the 18 CMIP6 models. Fig. 6 shows the time series of the regional

274 averaged values over six sub-regions of Asia during 2020–2100. Under the

275 global warming trends of all future scenarios, the climate-driven near-surface

276 O_3 concentrations increase constantly from 2020 to 2100 over many key

277 regions in Asia, such as North China (NC), South China (SC), Southeast Asia

278 (SEA), South India (SI) and Gangetic Plains (GP), except the Tibetan Plateau

279 (TP). The O_3 concentrations over SC, SEA, and SI are projected to increase

280 considerably with the maximum increase up to 5 ppb (20%) in 2095 (2091–

281 2100 mean) compared to 2025 (2020–2029 mean) under the SSP5-8.5

282 scenario, revealing a strong O3-climate penalty in most Asian regions. The

283   climate-driven changes in O_3 concentrations are smaller under the less warming

284   scenarios, especially in SSP1-2.6 that has O_3 changes less than 5% across

285   Asia. These suggest that future climate following low emissions and sustainable

286   pathways is more favorable for the mitigation of O_3 pollution in Asia than high

287   forcing scenarios.

288   The strong O_3-climate penalty over eastern China can be attributed to the

289   particularly high O_3 precursor emissions (Fig. S1), relative to western China,

290   which lead to a positive local net O_3 production close to sources in a warming

291   climate (Fig. S2) (Zanis et al., 2022). The absolute and percentage changes in

292   regional averaged surface O_3 concentrations between 2025 and 2095 under the

293   four scenarios are shown in Figure 7. The climate-driven changes in O_3

294   concentrations are gradually stronger from north (2–3%) to south (3–8%) of

295   China, which demonstrates that the changes in meteorology exert a greater

296   impact on ground-level O_3 concentrations over SC than NC under future climate

297   change. By the end of the 21st century, the relative humidity will decrease (Fig.

298   S3) and downward solar radiation will increase (Fig. S4) over SC compared to

299   those in 2025, which are conducive to the O_3 productions, while NC has the

300   opposite changes. Moreover, cloud cover will decrease more over SC than NC

301   (Fig. S5), contributing to the larger increase in O_3 productions and

302   concentrations over SC than NC in a warming climate.

303   In South Asia, climate change also enhances O_3 concentrations by <5%

304   over GP and SI (Fig. 7), due to the massive precursor emissions (Fig. S1) and

305   O_3 productions. Over SI, the decreases in relative humidity (Fig. S3) and cloud

306   amount (Fig. S5), and increases in downward solar radiation at the surface (Fig.

307   S4) favor photochemical production of O_3 and induce the large increases in O_3

308   concentrations in this region. Averaged over SEA, O_3 concentrations driven by

309   higher temperature (Fig. S2), more downward solar radiation (Fig. S4) and

310   lower relative humidity (Fig. S3) and cloud cover (Fig. S5) in 2095 are projected

311   to increase O_3 concentrations by 5–7% in SSP3-7.0 and SSP5-8.5 and 0–3%

312   in SSP1-2.6 and SSP2-4.5 scenarios, relative to 2025 (Fig. 7).

313   The Tibetan Plateau (TP), known as the highest topography in China with

314   more solar radiation at the surface, has strong stratosphere-troposphere

315   exchanges of O_3 compared with other regions leading to high O_3 concentrations

316   over this region (Fig. S6). Climate-driven O_3 concentrations are projected to

317   decline by less than 2% over TP from 2025 to 2095 (Fig. 7). It is likely because

318   less solar radiation (Fig. S4) and more frequent occurrence of rainy weather

319   (Fig. S7) in the future would reduce the local chemical production of O_3.

320   **3.3 The seasonality of future climate-driven O_3 variations**

321   Climate over Asia has obvious seasonal variation related to the Asian

322   monsoon system. Figure 8 shows the spatial distributions of percentage

323   changes in projected climate-driven O_3 concentrations in spring (March–April–

324   May, MAM), summer (June–July–August, JJA), autumn (September–October–

325   November, SON), and winter (December–January–February, DJF) between

326   2025 and 2095 under the four scenarios. In general, air quality in many regions

327   of Asia will deteriorate in all seasons associated with intensified O$_3$ pollution

328   under climate change.

329   In eastern China, O$_3$ pollution occurs most frequently in summer and is

330   more severe in NC than SC currently (Li et al., 2019). Under future climate

331   warming, JJA O$_3$ concentrations will increase by 5–20% in SC under the high

332   forcing scenarios, while the changes in NC are less than 5%. It suggests that

333   future climate change will expand the summertime O$_3$ pollution from NC to SC

334   over eastern China. Another feature is the strong increases in O$_3$

335   concentrations by 10–20% throughout eastern China and exceeding 20% over

336   Sichuan Basin in SON, which relate to the significant increases in temperature

337   (Fig. S8) and solar radiation (Fig. S9) in this season over central-eastern China

338   under the high forcing scenarios. It further indicates that future climate change

339   will extend the O$_3$ pollution from summer into autumn.

340   In South Asia, the climate-driven increases in O$_3$ concentrations vary from

341   JJA over SI to DJF over GP. Relative to 2025, in summer of 2095, anomalous

342   high pressure (Fig. S10) along with anticyclone (Figs. S11 and S12) dominates

343   South Asia, which is not conducive to O$_3$ diffusion, leading to increases in JJA

344   O$_3$ concentrations over SI. The intensified O$_3$ pollution across GP in DJF under

345   climate change is related to the strong surface warming (Fig. S8), decreases in

346   relative humidity (Fig. S13), cloud cover (Fig. S14) and rainfall (Fig. S15), as

347   well as increases in solar radiation at the surface (Fig. S9), favoring the

348   photochemical production of O$_3$. In north part of Southeast Asia, JJA has the

349 largest O$_3$ rise via the same mechanism as for SI, while O$_3$ increases by the

350 same magnitude in all seasons in south part of Southeast Asia driven by future

351 climate change.

4. Conclusions and discussion

353   Ground-level O$_3$ pollution has been increasing over Asia in recent decades,

354 which harms human health and vegetations. In the future warmer climate, O$_3$

355 pollution over Asia can be modulated by changes in meteorological fields. In

356 this study, to examine the variations in O$_3$ concentrations over Asia due to the

357 future climate change, monthly surface O$_3$ concentrations from 2020 to 2100

358 under four climate scenarios (SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5)

359 are predicted using a ML model with input data from assimilated O$_3$ combining

360 GEOS-Chem simulations and real-time observations, future meteorological

361 parameters from CMIP6 multi-model simulations, emissions of O$_3$ precursors,

362 land use, topography, population density and spatiotemporal information. Our

363 results suggest that the future O$_3$ pollution over Asia will be significantly

364 exacerbated in a warming climate, especially under high forcing scenarios.

365   Trained by the assimilated O$_3$ concentrations and reanalysis data, the ML

366 model can well predict O$_3$ over Asia with the coefficient determination of 0.92

367 between assimilated and predicted O$_3$ concentrations and relative error of 9%.

368 Then the future Asian O$_3$ concentrations from 2020 to 2100 driven by climate

369 change are projected in the ML model with varying meteorological fields from

370 18 CMIP6 models under four future climate scenarios.

371   The climate penalty on O_3 is robust over most regions of Asia. The annual

372   mean O_3 levels in 2095 are projected to increase by 5–20% relative to 2025

373   under the high forcing scenarios over South China, Southeast Asia, and South

374   India and less than 10% over North China and Gangetic Plains, due to more

375   favorable meteorological conditions for O_3 photochemical production, while

376   there is a decrease of <5% over the Tibetan Plateau. The climate-driven

377   changes in O_3 concentrations are smaller under the less warming scenarios,

378   suggesting that future climate following low emissions and sustainable

379   pathways would be more effective in the mitigation of O_3 pollution in Asia than

380   the high forcing scenarios. Seasonal variation analysis reveals that the

381   summertime O_3 pollution over eastern China will expand from North China to

382   South China and extend into the cold season under the future climate change.

383   In addition, South Asian O_3 pollution will increase over South India in summer

384   and over Gangetic Plains in winter.

385   Zanis et al. (2022) analyzed the global climate change benefit and penalty

386   on O_3 based on sensitivity simulations from five CMIP6 models under the SSP3-

387   7.0 scenario. They showed positive changes in JJA O_3 concentrations by less

388   than 1 ppb from 2010 to 2095 over East Asia and South Asia driven by climate

389   change, but with large uncertainties due to the model diversity. The ML method

390   in this study gives similar positive changes in O_3 as Zanis et al. (2022). Pommier

391   et al. (2018) applied the EMEP chemical transport model driven by the

392   downscaled meteorological data from the NorESM1-M to investigate the

393   impacts of regional climate change on surface O_3 over India. They showed that

394   near-surface O_3 would increase by up to 4% over Northern India and decrease

395   by 3% over Southern India from 2050 to 2100 under the RCP8.5 scenario. We

396   show that the climate-driven O_3 in this study would increase over both Gangetic

397   Plains (0.2%) and South India (3%) under the SSP5-8.5 scenario in 2050

398   relative to 2016 (2014–2019 mean). The discrepancies may rise from that the

399   results of Pommier et al. (2018) were based on NorESM1-M simulated climate

400   alone, while the climate change predicted by 18 CMIP6 models were applied in

401   this study and the ensemble mean O_3 concentrations were shown here.

402   There are a few uncertainties and limitations in the projected near-surface

403   O_3 concentrations over Asia in terms of input data for the ML model, GEOS-

404   Chem simulations, and CMIP6 multi-model simulations. First, only

405   observational data over 2014–2019 across China were used for the O_3

406   assimilation. Longer-term measurements with broader spatial coverage are

407   more desirable to improve the model performance. Land use data and

408   population density are fixed at present-day conditions when predicting the

409   future O_3 since we focus on the variations in meteorological parameters under

410   climate change, which will vary in the future. In addition, natural O_3 precursor

411   emissions such as biogenic emissions of NMVOCs, and NO_x from soil and

412   lightning sources are fixed at year-2016 levels in the future estimates, which

413   can induce biases in the O_3 projections since climate change can strongly

414   influence natural emissions of O_3 precursors (Liu et al., 2019). Although the

415   climate influence of methane is considered in the future predictions, its role in

416   the O_3 production is not included in the ML model. The GEOS-Chem model has

417   been demonstrated to well capture the magnitude of and spatiotemporal

418   variations in O_3, with an average bias of about 10% over China (Lou et al., 2014)

419   and Southeast Asia (Marvin et al., 2021), and less than 20% over India (David

420   et al., 2019). The future decrease in relative humidity will cause stomatal

421   closure and also increase surface O_3. The O_3-vegetation interactions are not

422   represented in the default GEOS-Chem model. A newly coupled global

423   atmospheric chemistry-vegetation model (Lei et., 2020) could be applied in the

424   future study. Additionally, the meteorological parameters characterizing future

425   climate change from the CMIP6 multi-model simulations can also give rise to

426   uncertainties in this study (Xu et al., 2021). Also, the good ability of the ML

427   model for the present-day condition may not imply a satisfactorily extrapolation

428   under the future warming condition, which can bias our results and deserves

429   further investigation in future studies.

430   Overall, our study provides a framework of combining real-time

431   observations, chemical transport model simulations and multi-climate model

432   predictions with data assimilation and machine learning methods to estimate

433   future climate driven near-surface O_3 concentrations. The emphasis of this work

434   is to quantify the impacts of future climate change on O_3 pollution in Asia, which

435   is of great significance for the future O_3 pollution mitigation strategies.

Author contributions

YY designed the research. HL performed the model simulations, analyzed data

and wrote the initial draft. JJ designed the data assimilation. YY, JJ, HW, and

KL helped edit and review the manuscript. All the authors discussed the results

and contributed to the final manuscript.

Code and data availability

The GEOS-Chem model is available at

https://zenodo.org/record/3974569#.YTD81NMzagR (last access: 1 August

2022). MERRA-2 reanalysis data can be downloaded at

https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/ (last access: 1 August 2022).

Multi-model projections of climate variables are from Scenario Model

Intercomparison Project in Phase 6 of the Coupled Model Intercomparison

Project https://esgf-node.llnl.gov/search/cmip6/ (last access: 1 August 2022).

Land cover is derived from http://maps.elie.ucl.ac.be/CCI/viewer/download.php

(last access: 1 August 2022). Hourly O_3 concentrations are obtained from the

public website of the China National Environmental Monitoring Centre

http://www.cnemc.cn/en/ (last access: 1 August 2022). Normalized difference

vegetation index is obtained from https://www.ncei.noaa.gov/data/avhrr-land-

normalized-difference-vegetation-index/access/ (last access: 1 August 2022).

Topography is collected from https://cgiarcsi.community/data/srtm-90m-digital-

elevation-database-v4-1/ (last access: 1 August 2022). Population density is

457   acquired from https://landscan.ornl.gov/landscan-datasets (last access: 1

458   August 2022).

459   ***Acknowledgments***

460   H.W. acknowledges the support by the U.S. Department of Energy (DOE),

461   Office of Science, Office of Biological and Environmental Research (BER), as

462   part of the Earth and Environmental System Modeling program. The Pacific

463   Northwest National Laboratory (PNNL) is operated for DOE by the Battelle

464   Memorial Institute under contract DE-AC05-76RLO1830. The projected O_3

465   concentrations in this study are available upon request.

466   ***Competing Interest***

467   The contact author has declared that neither they nor their co-authors have any

468   competing interests.

469   ***Financial support.***

470   This study was supported by the National Key Research and Development

471   Program of China (grant 2019YFA0606800 and 2020YFA0607803) and the

472   National Natural Science Foundation of China (grant 41975159) and Jiangsu

473   Science Fund for Distinguished Young Scholars (grant BK20211541).

474 **Reference**

475 Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li,
476   Q., Liu, H., Mickley, L. J., and Schultz, M. G.: Global modeling of
477   tropospheric chemistry with assimilated meteorology: Model description
478   and evaluation, J. Geophys. Res. Atmos., 106, 23073–23095,
479   https://doi.org/10.1029/2001JD000807, 2001.
480

481 Cakmak, S., Hebbern, C., Pinault, L., Lavigne, E., Vanos, J., Crouse, D. L., and
482   Tjepkema, M.: Associations between long-term $PM_{2.5}$ and ozone exposure
483   and mortality in the Canadian Census Health and Environment Cohort
484   (CANCHEC), by spatial synoptic classification zone, Environ. Int., 111, 200–
485   211, https://doi.org/10.1016/j.envint.2017.11.030, 2018.
486

487 Checa-Garcia, R., Hegglin, M. I., Kinnison, D., Plummer, D. A., and Shine, K.
488   P.: Historical tropospheric and stratospheric ozone radiative forcing using
489   the CMIP6 database, Geophys. Res. Lett., 45, 3264–3273,
490   https://doi.org/10.1002/2017GL076770, 2018.
491

492 Chen, L., Liang, S., Li, X., Mao, J., Gao, S., Zhang, H., Sun, Y., Vedal, S., Bai,
493   Z., Ma, Z., Haiyu., and Azzi, M.: A hybrid approach to estimating long-term
494   and short-term exposure levels of ozone at the national scale in China using
495   land use regression and Bayesian maximum entropy, Sci. Total Environ.,
496   752, 141780, https://doi.org/10.1016/j.scitotenv.2020.141780, 2020.
497

498 Cheng, J., Tong, D., Zhang, Q., Liu, Y., Lei, Y., Yan, G., Yan, L., Yu, S., Cui, R.
499   Y., Clarke, L., Geng, G., Zheng, B., Zhang, X., Davis, S. J., and He, K.:
500   Pathways of China's $PM_{2.5}$ air quality 2015−2060 in the context of carbon
501   neutrality, Natl. Sci. Rev., 8, nwab078, https://doi.org/10.1093/nsr/nwab078,
502   2021.
503

504 Colette, A., Andersson, C., Baklanov, A., Bessagnet, B., Brandt, J., Christensen,
505   J. H., Doherty, R., Engardt, M., Geels, C., Giannakopoulos, C., Hedegaard,
506   G. B., Katragkou, E., Langner, J., Lei, H., Manders, A., Melas, D., Meleux,
507   F., Rouïl, L., Sofiev, M., Soares, J., Stevenson, D. S., Tombrou-Tzella, M.,
508   Varotsos, K. V., and Young, P.: Is the ozone climate penalty robust in
509   Europe? Environ. Res. Lett., 10, 084015, http://dx.doi.org/10.1088/1748-
510   9326/10/8/084, 2015.
511

512 Dang, R., Liao, H., and Fu, Y.: Quantifying the anthropogenic and
513   meteorological influences on summertime surface ozone in China over
514   2012–2017, Sci. Total Environ., 754, 142394,
515   https://doi.org/10.1016/j.scitot, 2021.
516

David, L. M., Ravishankara, A., Brewer, J. F., Sauvage, B., Thouret, V., Venkataramani, S., and Sinha, V.: Tropospheric ozone over the Indian subcontinent from 2000 to 2015: Data set and simulation using GEOS-Chem chemical transport model, Atmos. Environ., 219, 117039, https://doi.org/10.1016/j.atmosenv.2019.117039, 2019.

Di, Q., Rowland, S., Koutrakis, P., and Schwartz, J.: A hybrid model for spatially and temporally resolved ozone exposures in the continental United States, J. Air Waste Manage. Assoc., 67, 39–52, https://doi.org/10.1080/10962247.2016.1200159, 2017.

Doherty, R. M., Wild, O., Shindell, D. T., Zeng, G., MacKenzie, I. A., Collins, W. J., Fiore, A. M., Stevenson, D. S., Dentener, F. J., Schultz, M. G., Hess, P., Derwent, R. G., and Keating, T. J.: Impacts of climate change on surface ozone and intercontinental ozone pollution: A multi-model study, J. Geophys. Res., 118, 3744–3763, https://doi.org/10.1002/jgrd.50266, 2013.

Fann, N., Nolte, C. G., Dolwick, P., Spero, T. L., Brown, A. C., Phillips, S., and Anenberg, S.: The geographic distribution and economic value of climate change-related ozone health impacts in the United States in 2030, J. Air Waste Manag. Assoc., 65, 570–580, https://doi.org/10.1080/10962247.2014.996270, 2015.

Fu, T.-M., and Tian, H.: Climate Change Penalty to Ozone Air Quality: Review of Current Understandings and Knowledge Gaps, Curr. Pollut. Rep., 5, 159–171, https://doi.org/10.1007/s40726-019-00115-6, 2019.

Fu, Y., and Tai, A. P. K.: Impact of climate and land cover changes on tropospheric ozone air quality and public health in East Asia between 1980 and 2010, Atmos. Chem. Phys., 15, 10093–10106, https://doi.org/10.5194/acp-15-10093-2015, 2015.

Gao, M., Gao, J., Zhu, B., Kumar, R., Lu, X., Song, S., Zhang, Y., Jia, B., Wang, P., Beig, G., Hu, J., Ying, Q., Zhang, H., Sherman, P., and McElroy, M. B.: Ozone pollution over China and India: seasonality and sources, Atmos. Chem. Phys., 20, 4399–4414, https://doi.org/10.5194/acp-20-4399-2020, 2020.

Gaudel, A., Cooper, O. R., Ancellet, G., Barret, B., Boynard, A., Burrows, J. P., Clerbaux, C., Coheur, P. F., Cuesta, J., Cuevas, E., Doniki, S., Dufour, G., Ebojie, F., Foret, G., Garcia, O., Granados-Muñoz, M. J., Hannigan, J. W., Hase, F., Hassler, B., Huang, G., Hurtmans, D., Jaffe, D., Jones, N., Kalabokas, P., Kerridge, B., Kulawik, S., Latter, B., Leblanc, T., Le Flochmoën, E., Lin, W., Liu, J., Liu, X., Mahieu, E., McClure-Begley, A.,

561   Neu, J. L., Osman, M., Palm, M., Petetin, H., Petropavlovskikh, I., Querel,
562   R., Rahpoe, N., Rozanov, A., Schultz, M. G., Schwab, J., Siddans, R.,
563   Smale, D., Steinbacher, M., Tanimoto, H., Tarasick, D. W., Thouret, V.,
564   Thompson, A. M., Trickl, T., Weatherhead, E., Wespes, C., Worden, H. M.,
565   Vigouroux, C., Xu, X., Zeng, G., and Ziemke, J.: Tropospheric Ozone
566   Assessment Report: Present-day distribution and trends of tropospheric
567   ozone relevant to climate and global atmospheric chemistry model
568   evaluation, Elem. Sci. Anth., 6, 39, https://doi.org/10.1525/elementa.291,
569   2018.

571 Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L.,
572   Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K.,
573   Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da
574   Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D.,
575   Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M.,
576   Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era
577   Retrospective Analysis for Research and Applications, Version 2 (MERRA-
578   2), J. Clim., 30, 5419−5454, https://doi.org/10.1175/JCLI-D-16-0758.1,
579   2017.

581 Gong, C., and Liao, H.: A typical weather pattern for ozone pollution events in
582   North China, Atmos. Chem. Phys., 19, 13725–13740,
583   https://doi.org/10.5194/acp-19-13725-2019, 2019.

585 Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T.,
586   Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and
587   Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated
588   framework for modeling biogenic emissions, Geosci. Model Dev., 5, 1471–
589   1492, https://doi.org/10.5194/gmd-5-1471-2012, 2012.

591 Han, H., Liu, J., Shu, L., Wang, T., and Yuan, H.: Local and synoptic
592   meteorological influences on daily variability in summertime surface ozone
593   in eastern China, Atmos. Chem. Phys., 20, 203–222,
594   https://doi.org/10.5194/acp-20-203-2020, 2020.

596 Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G.,
597   Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C.,
598   Dawidowski, L., Kholod, N., Kurokawa, J.-I., Li, M., Liu, L., Lu, Z., Moura,
599   M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750–2014)
600   anthropogenic emissions of reactive gases and aerosols from the
601   Community Emissions Data System (CEDS), Geosci. Model Dev., 11,
602   369−408, https://doi.org/10.5194/gmd-11-369-2018, 2018.

604 Hudman, R. C., Moore, N. E., Mebust, A. K., Martin, R. V., Russell, A. R., Valin,

605   L. C., and Cohen, R. C.: Steps towards a mechanistic model of global soil
606   nitric oxide emissions: implementation and space based-constraints, Atmos.
607   Chem. Phys., 12, 7779–7795, https://doi.org/10.5194/acp-12-7779-2012,
608   2012.

610   IPCC: Climate change 2021: The physical science basis. Contribution of
611   working group I to the sixth assessment report of the intergovernmental
612   panel on climate change. Cambridge, UK: Cambridge University Press,
613   2021.

615   Jacob, D. J., and Winner, D. A.: Effect of climate change on air quality, Atmos.
616   Environ., 43, 51–63, https://doi.org/10.1016/j.atmosenv.2008.09.051, 2009.

618   Jeong, J. I., and Park, R. J.: Effects of the meteorological variability on regional
619   air quality in East Asia, Atmos. Environ., 69, 46–55,
620   https://doi.org/10.1016/J.Atmosenv.2012.11.061, 2013.

622   Kavassalis, S. C., and Murphy, J. G.: Understanding ozone-meteorology
623   correlations: A role for dry deposition, Geophys. Res. Lett., 44, 2922–2931,
624   https://doi.org/10.1002/2016gl071791, 2017.

626   Lei, Y., Yue, X., Liao, H., Gong, C., and Zhang, L.: Implementation of Yale
627   Interactive terrestrial Biosphere model v1.0 into GEOS-Chem v12.0.0: a
628   tool for biosphere– chemistry interactions, Geosci. Model Dev., 13, 1137–
629   1153, https://doi.org/10.5194/gmd-13-1137-2020, 2020.

631   Li, H., Yang, Y., Wang, H., Li, B., Wang, P., Li, J., and Liao, H.: Constructing a
632   spatiotemporally coherent long-term $PM_{2.5}$ concentration dataset over
633   China during 1980–2019 using a machine learning approach, Sci. Total
634   Environ., 765, 144263, https://doi.org/10.1016/j.scitotenv.2020.144263,
635   2021.

637   Li, H., Yang, Y., Wang, H., Wang, P., Yue, X., and Liao, H.: Projected Aerosol
638   Changes Driven by Emissions and Climate Change Using a Machine
639   Learning Method, Environ. Sci. Technol., 56, 7, 3884–3893,
640   https://doi.org/10.1021/acs.est.1c04380, 2022.

642   Li, K., Jacob, D. J., Liao, H., Shen, L., Zhang, Q., and Bates, K. H.:
643   Anthropogenic Drivers of 2013–2017 Trends in Summer Surface Ozone in
644   China, P. Natl. Acad. Sci. USA., 116, 422–427,
645   https://doi.org/10.1073/pnas.1812168116, 2019.

647   Li, M., Yu, S., Chen, X., Li, Z., Zhang, Y., Wang, L., Liu, W., Li, P., Lichtfouse,
648   E., Rosenfeld, D., and Seinfeld, J. H.: Large scale control of surface ozone

649 by relative humidity observed during warm seasons in China, Environ.
650 Chem. Lett., 19, 3981–3989, https://doi.org/10.1007/s10311-021-01265-0,
651 2021.
652

653 Li, R., Zhao, Y., Zhou, W., Meng, Y., Zhang, Z., and Fu, H.: Developing a novel
654 hybrid model for the estimation of surface 8h ozone (O_3) across the remote
655 Tibetan Plateau during 2005–2018, Atmos. Chem. Phys., 20, 6159–6175,
656 https://doi.org/10.5194/acp-20-6159-20, 2020.
657

658 Lin, J.-T., and McElroy, M. B.: Impacts of boundary layer mixing on pollutant
659 vertical profiles in the lower troposphere: Implications to satellite remote
660 sensing, Atmos. Environ., 44, 1726–1739,
661 https://doi.org/10.1016/j.atmosenv.2010.02.009, 2010.
662

663 Liu, S., Xing, J., Zhang, H., Ding, D., Zhang, F., Zhao, B., Sahu, S. K., and
664 Wang, S.: Climate-driven trends of biogenic volatile organic compound
665 emissions and their impacts on summertime ozone and secondary organic
666 aerosol in China in the 2050s, Atmos. Environ., 218, 117020,
667 https://doi.org/10.1016/j.atmosenv.2019.117020, 2019.
668

669 Liu, X., Zhu, Y., Xue, L., Desai, A. R., and Wang, H.: Cluster-enhanced
670 ensemble learning for mapping global monthly surface ozone from 2003 to
671 2019, Geophys. Res. Lett., 49, e2022GL097947,
672 https://doi.org/10.1029/2022GL097947, 2022.
673

674 Liu, Y., and Wang, T.: Worsening urban ozone pollution in China from 2013 to
675 2017–Part 1: The complex and varying roles of meteorology, Atmos. Chem.
676 Phys., 20, 6305–6321, https://doi.org/10.5194/acp-20-6305-2020, 2020.
677

678 Lou, S., Liao, H., and Zhu, B.: Impacts of aerosols on surface-layer ozone
679 concentrations in China through heterogeneous reactions and changes in
680 photolysis rates, Atmos. Environ., 85, 123–138,
681 http://dx.doi.org/10.1016/j.atmosenv.2013.12.004, 2014.
682

683 Lu, X., Zhang, L., Chen, Y., Zhou, M., Zheng, B., Li, K., Liu, Y., Lin, J., Fu, T.-
684 M., and Zhang, Q.: Exploring 2016–2017 surface ozone pollution over
685 China: source contributions and meteorological influences, Atmos. Chem.
686 Phys., 19, 8339–8361, https://doi.org/10.5194/acp-19-8339-2019, 2019.
687

688 Lu, X., Zhang, L., Wang, X., Gao, M., Li, K., Zhang, Y., Yue, X., and Zhang, Y.:
689 Rapid Increases in Warm-Season Surface Ozone and Resulting Health
690 Impact in China since 2013, Environ. Sci. Technol. Lett., 7, 240–247,
691 https://doi.org/10.1021/acs.estlett.0c00171, 2020.
692

693   Malley, C. S., Henze, D. K., Kuylenstierna, J. C. I., Vallack, H., Davila, Y.,
694   Anenberg, S. C., Turner, M. C., and Ashmore, M.: Updated Global
695   Estimates of Respiratory Mortality in Adults ≥ 30 Years of Age Attributable
696   to Long-Term Ozone Exposure, Environ. Health Perspect., 125, 087021,
697   https://doi.org/10.1289/EHP1390, 2017.

699   Mao, J., Paulot, F., Jacob, D. J., Cohen, R. C., Crounse, J. D., Wennberg, P.
700   O., Keller, C. A., Hudman, R. C., Barkley, M. P., and Horowitz, L. W.: Ozone
701   and organic nitrates over the eastern United States: sensitivity to isoprene
702   chemistry, J. Geophys. Res. Atmos, 118, 11256–68,
703   https://doi.org/10.1002/jgrd.50817, 2013.

705   Marvin, M. R., Palmer, P. I., Latter, B. G., Siddans, R., Kerridge, B. J., Latif, M.
706   T., and Khan, M. F.: Photochemical environment over Southeast Asia
707   primed for hazardous ozone levels with influx of nitrogen oxides from
708   seasonal biomass burning, Atmos. Chem. Phys., 21, 1917–1935,
709   https://doi.org/10.5194/acp-21-1917-2021, 2021.

711   McLinden, C. A., Olsen, S. C., Hannegan, B., Wild, O., Prather, M. J., and
712   Sundet, J.: Stratospheric ozone in 3-D models: A simple chemistry and the
713   cross-tropopause flux, J. Geophys. Res. Atmos., 105, 14653–14665,
714   https://doi.org/10.1029/2000jd900124, 2000.

716   Mills, G., Pleijel, H., Malley, C. S., Sinha, B., Cooper, O. R., Schultz, M. G.,
717   Neufeld, H. S., Simpson, D., Sharps, K., Feng, Z., Gerosa, G., Harmens,
718   H., Kobayashi, K., Saxena, P., Paoletti, E., Sinha, V., and Xu, X.:
719   Tropospheric ozone assessment report: Present-day tropospheric ozone
720   distribution and trends relevant to vegetation, Elem. Sci. Anth., 6,
721   47, https://doi.org/10.1525/elementa.302, 2018.

723   Murray, L. T., Jacob, D. J., Logan, J. A., Hudman, R. C., and Koshak, W. J.:
724   Optimized regional and interannual variability of lightning in a global
725   chemical transport model constrained by LIS/OTD satellite data, J.
726   Geophys. Res. Atmos., D20307, https://doi.org/10.1029/2012jd017934,
727   2012.

729   Ni, R., Lin, J., Yan, Y., and Lin, W.: Foreign and domestic contributions to
730   springtime ozone over China, Atmos. Chem. Phys., 18, 11447–11469,
731   https://doi.org/10.5194/acp-18-11447-2018, 2018.

733   O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt,
734   G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss,
735   R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison
736   Project (ScenarioMIP) for CMIP6, Geosci. Model Dev., 9, 3461–3482,

737   https://doi.org/10.5194/gmd-9-3461-2016, 2016.

739   Ott, L. E., Pickering, K. E., Stenchikov, G. L., Allen, D. J., DeCaria, A. J., Ridley,
740   B., Lin, R.-F., Lang, S., and Tao, W.-K.: Production of lightning NO_x and its
741   vertical distribution calculated from three-dimensional cloud-scale chemical
742   transport model simulations, J. Geophys. Res., 115, D04301,
743   https://doi.org/10.1029/2009JD011880, 2010.

745   Park, R. J., Jacob, D. J., Field, B. D., Yantosca, R. M., and Chin, M.: Natural
746   and transboundary pollution influences on sulfate-nitrate-ammonium
747   aerosols in the United States: Implications for policy, J. Geophys. Res.
748   Atmos., 109, 20, https://doi.org/10.1029/2003jd004473, 2004.

750   Pommier, M., Fagerli, H., Gauss, M., Simpson, D., Sharma, S., Sinha, V.,
751   Ghude, D. S., Landgren, O., Nyiri, A., and Wind, P.: Impact of regional
752   climate change and future emission scenarios on surface O_3 and $PM_{2.5}$ over
753   India, Atmos. Chem. Phys., 18, 103–27, https://doi.org/10.5194/acp-18-
754   103-2018, 2018.

756   Pye, H. O. T., Liao, H., Wu, S., Mickley, L. J., Jacob, D. J., Henze, D. K., and
757   Seinfeld, J. H.: Effect of changes in climate and emissions on future sulfate-
758   nitrate-ammonium aerosol levels in the United States, J. Geophys. Res.
759   Atmos., 114, D01205, https://doi.org/10.1029/2008jd010701, 2009.

761   Qian, J., Liao, H., Yang, Y., Li, K., Chen, L., and Zhu, J.: Meteorological
762   influences on daily variation and trend of summertime surface ozone over
763   years of 2015–2020: Quantification for cities in the Yangtze River Delta, Sci.
764   Total   Environ.,   834,   155107,
765   https://doi.org/10.1016/j.scitotenv.2022.155107, 2022.

767   Rodriguez, J. D., Perez, A., and Lozano, J. A.: Sensitivity analysis of k-fold
768   cross validation in prediction error estimation, IEEE T. Pattern Anal., 32,
769   569–575, https://doi.org/10.1109/TPAMI.2009.187, 2010.

771   Santurtún, A., González-Hidalgo, J. C., Sanchez-Lorenzo, A., and Zarrabeitia,
772   M. T.: Surface ozone concentration trends and its relationship with weather
773   types in Spain (2001–2010), Atmos. Environ., 101, 10–22,
774   https://doi.org/10.1016/j.atmosenv.2014.11.005, 2015.

776   Su, X., An, J., Zhang, Y., Zhu, P., and Zhu, B.: Prediction of ozone hourly
777   concentrations by support vector machine and kernel extreme learning
778   machine using wavelet transformation and partial least squares methods,
779   Atmos. Pollut. Res., 6, 51–60, https://doi.org/10.1016/j.apr.2020.02.024,
780   2020.

Toh, Y. Y., Lim, S. F., and von Glasow, R.: The influence of meteorological factors and biomass burning on surface ozone concentrations at Tanah Rata, Malaysia, Atmos. Environ., 70, 435–446, https://doi.org/10.1016/j.atmosenv.2013.01.018, 2013.

van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global fire emissions estimates during 1997–2016, Earth Syst. Sci. Data, 9, 697–720, https://doi.org/10.5194/essd-9-697-2017, 2017.

Wang, Y., Shen, L., Wu, S., Mickley, L. J., He, J., and Hao, J.: Sensitivity of surface ozone over China to 2000–2050 global changes of climate and emissions, Atmos. Environ., 75, 374–382, https://doi.org/10.1016/j.atmosenv.2013.04.045, 2013.

Wang, Z., Lin, L., Xu, Y., Che, H., Zhang, X., Dong, W., Wang, C., Gui, K., and Xie, B.: Incorrect Asian aerosols affecting the attribution and projection of regional climate change in CMIP6 models, npj Clim. Atmos. Sci., 4, 2, https://doi.org/10.1038/s41612-020-00159-2, 2021.

Wei, J., Li, Z., Li, K., Dickerson, R., Pinker, R., Wang, J., Liu, X., Sun, L., Xue, W., and Cribb, M.: Full-coverage mapping and spatiotemporal variations of ground-level ozone (O_3) pollution from 2013 to 2020 across China, Remote Sens. Environ., 270, 112775, https://doi.org/10.1016/j.rse.2, 2022.

Xu, Z., Han, Y., Tam, C. Y., Yang, Z., and Fu, C.: Bias-corrected CMIP6 global dataset for dynamical downscaling of the historical and future climate (1979–2100), Sci. Data, 8, 293, https://doi.org/10.1038/s41597-021-01079-3, 2021.

Xue, T., Zheng, Y., Geng, G., Xiao, Q., Meng, X., Wang, M., Li, X., Wu, N., Zhang, Q., and Zhu, T.: Estimating Spatiotemporal Variation in Ambient Ozone Exposure during 2013–2017 Using a Data-Fusion Model, Environ. Sci. Technol., 54, 14877–14888, https://dx.doi.org/10.1021/acs.est.0c03098, 2020.

Yang, Y., Li, M., Wang, H., Li, H., Wang, P., Li, K., Gao, M., and Liao, H.: ENSO modulation of summertime tropospheric ozone over China, Environ. Res. Lett., 17, 034020, https://doi.org/10.1088/1748-9326/ac54cd, 2022.

Yin, Z., Cao, B., and Wang, H.: Dominant patterns of summer ozone pollution in eastern China and associated atmospheric circulations, Atmos. Chem.

825 Phys., 19, 13933–13943, https://doi.org/10.5194/acp-19-13933-2019, 2019.
826
827 Yue, X., Unger, N., Harper, K., Xia, X., Liao, H., Zhu, T., Xiao, J., Feng, Z., and
828   Li, J.: Ozone and haze pollution weakens net primary productivity in China,
829   Atmos. Chem. Phys., 17, 6073–6089, https://doi.org/10.5194/acp-17-6073-
830   2017, 2017.
831
832 Zanis, P., Akritidis, D., Turnock, S., Naik, V., Szopa, S., Georgoulias, A. K.,
833   Bauer, S. E., Deushi, M., Horowitz, L. W., Keeble, J., Le Sager, P.,
834   O'Connor, F. M., Oshima, N., Tsigaridis, K., and van Noije., T.: Climate
835   change penalty and benefit on surface ozone: a global perspective based
836   on CMIP6 earth system models, Environ. Res. Lett., 17, 024014,
837   https://doi.org/10.1088/1748-9326/ac4a34, 2022.
838
839 Zhang, X., Zhao, L., Cheng, M., and Chen, D.: Estimating ground-level ozone
840   concentrations in eastern China using satellite-based precursors, IEEE
841   Trans. Geosci. Remote Sens., 58, 4754–4763, https://doi.org/
842   10.1109/TGRS.2020.2966780, 2020.
843
844 Zhou, C., Gao, M., Li, J., Bai, K., Tang, X., Lu, X., Liu, C., Wang, Z., and Guo,
845   Y.: Optimal Planning of Air Quality-Monitoring Sites for Better Depiction of
846   PM$_{2.5}$ Pollution across China, Environ. Au., 2, 314–323,
847   https://doi.org/10.1021/acsenvironau.1c00051, 2022.
848
849 Zhu, J., Liao, H., Mao, Y., Yang, Y., and Jiang, H.: Interannual variation,
850   decadal trend, and future change in ozone outflow from East Asia, Atmos.
851   Chem. Phys., 17, 3729–3747, https://doi.org/10.5194/acp-17-3729-2017,
852   2017.
853

854 **Table 1.** Summary of detailed datasets used in this study.

Dataset type	Variable	Description	Spatial resolution	Temporal resolution	Time period	Data source
O₃	O₃	Near-surface ozone concentrations	0.5°×0.625°	Monthly (historical)	2014–2019 (historical)	Assimilated GEOS-Chem simulations and Observations
Meteorology	T_2m	Air temperature at 2 meters	0.5°×0.625°	Monthly (historical) Monthly (future)	2014–2019 (historical) 2020–2100 (future)	MERRA-2 (historical) Adjusted CMIP6 (future)
	T_850	Air temperature at 850 hPa				
	T_500	Air temperature at 500 hPa				
	U_850	Zonal wind at 850 hPa				
	U_500	Zonal wind at 500 hPa				
	V_850	Meridional wind at 850 hPa				
	V_500	Meridional wind at 500 hPa				
	RH	Relative humidity				
	PRECP	Precipitation rate				
	CLT	Total cloud cover				
	RSDS	Incoming shortwave radiation at the surface				
	SLP	Sea level pressure				
Emission	NOₓ	Nitric oxide from anthropogenic sources	0.5°×0.625°	Monthly (historical) Monthly (future)	2014–2019 (historical) 2019 (future) 2016 2016	CEDS (Anthropogenic) GFED4 (Biomass burning) MEGAN V-2.1 (Biogenic)
		Nitric oxide from biomass burning				
		Nitric oxide from soil sources				
		Nitric oxide from lightning sources				
	CO	Carbon monoxide from anthropogenic sources			2014–2019 (historical) 2019 (future)	
		Carbon monoxide from biomass burning				
	NMVOC	Non-methane volatile organic compounds from anthropogenic sources				
		Non-methane volatile organic compounds from biomass burning				
		Non-methane volatile organic compounds from biogenic sources			2016	
Land use	LC	Land cover	300 m×300 m	Monthly	2014–2019 (historical) 2019 (future)	ESA CCI
	NDVI	Normalized Difference Vegetation Index	0.05°×0.05°			AVHRR
Topography	TOPO	Digital elevation model	90 m×90 m	-	2010	SRTM
Population	POP	Population density	1 km×1 km	-	2010	Land Scan

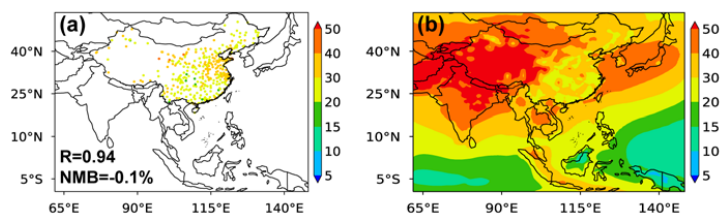
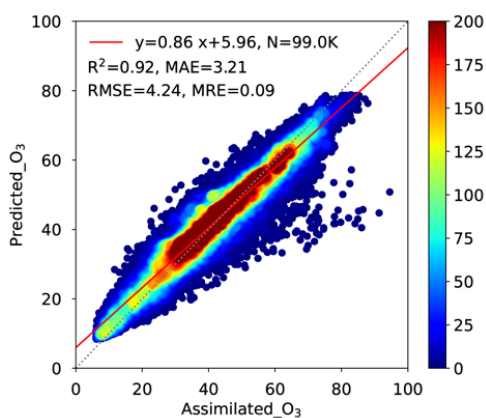



Figure 1. Spatial distributions of observed (a) and assimilated (b) O_3 concentrations in 2014–2019 over Asia. Correlation coefficient (R) between observed and assimilated O_3 and the normalized mean bias (NMB = $\sum$ (Observed – Assimilated) / $\sum$ Assimilated $\times$ 100%) are given at the bottom left of panel (a).

862

863

864 **Figure 2.** Density scatterplots of 10-fold cross-validation results for monthly O_3
865 concentrations (ppb) in 2019 over Asia. The gray and red lines are the 1:1 line
866 and linear regression line, respectively. Statistical metrics including the number
867 of samples (N), correlation of determination (R^2, unitless), root mean square
868 error (RMSE, ppb), mean absolute error (MAE, ppb), and mean relative error
869 (MRE, %) are shown at the top left.

870

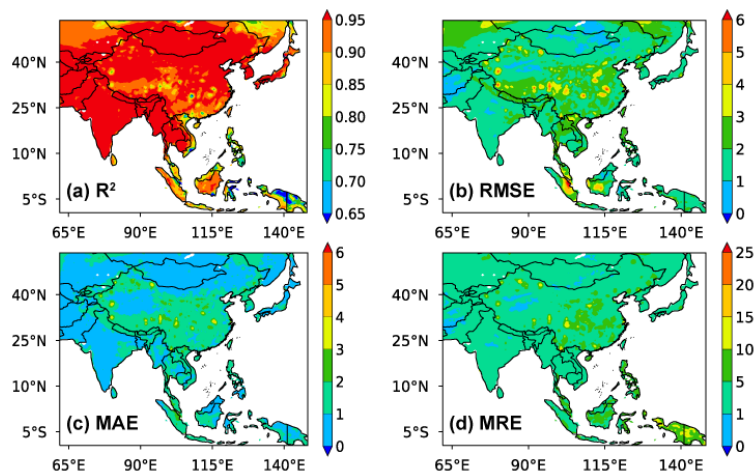



Figure 3. Spatial distributions of the performance statistics of the random forest (RF) model with regard to (a) R^2 (unitless), (b) RMSE (ppb), (c) MAE (ppb), and (d) MRE (%) during 2014–2019 over Asia.

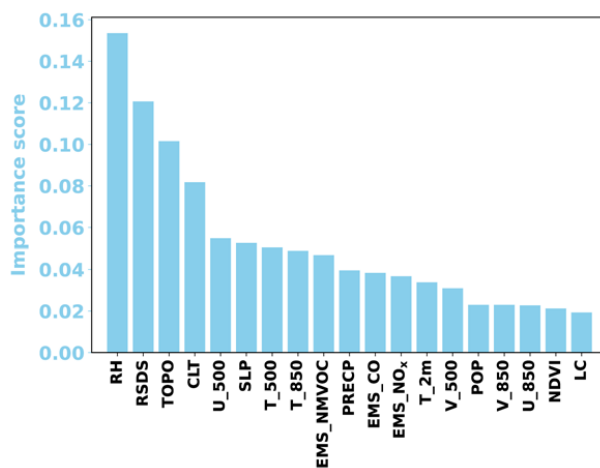



Figure 4. Importance scores of independent variables (meteorological parameters, emissions, land use, topography, and population density) used in the ML model for predicting future near-surface O_3 concentrations over Asia.

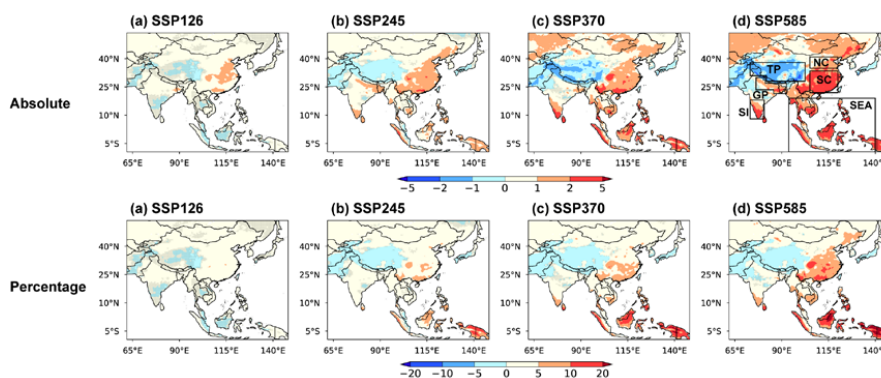



Figure 5. The spatial distributions of absolute (ppb) and percentage difference (%) of surface O_3 level between 2025 (2020–2029 mean) and 2095 (2091–2100 mean) driven by climate change under four scenarios (a, e) SSP1-2.6, (b, f) SSP2-4.5, (c, g) SSP3-7.0 and (d, h) SSP5-8.5. The box-outlined areas in (d) are North China (NC, 35°–41°N, 105°–120°E), South China (SC, 22°–33.5°N, 105°–120°E), Southeast Asia (SEA, -9.5°S–19°N, 93.75°–140°E), South India (SI, 8°–18°N, 73.125°–80.625°E), Gangetic Plains (GP, 21.5°–23.5°N, 85.625°–92.5°E, 23.5°–27°N, 76.25°–92.5°E, and 27°–30°N, 76.25°–81.25°E), and Tibetan Plateau (TP, 28°–31°N, 81.875°–102.5°E and 31°–38°N, 73.125°–102.5°E). No overlaying hatch pattern indicates statistical significance with 95% confidence from a two-tailed t test.

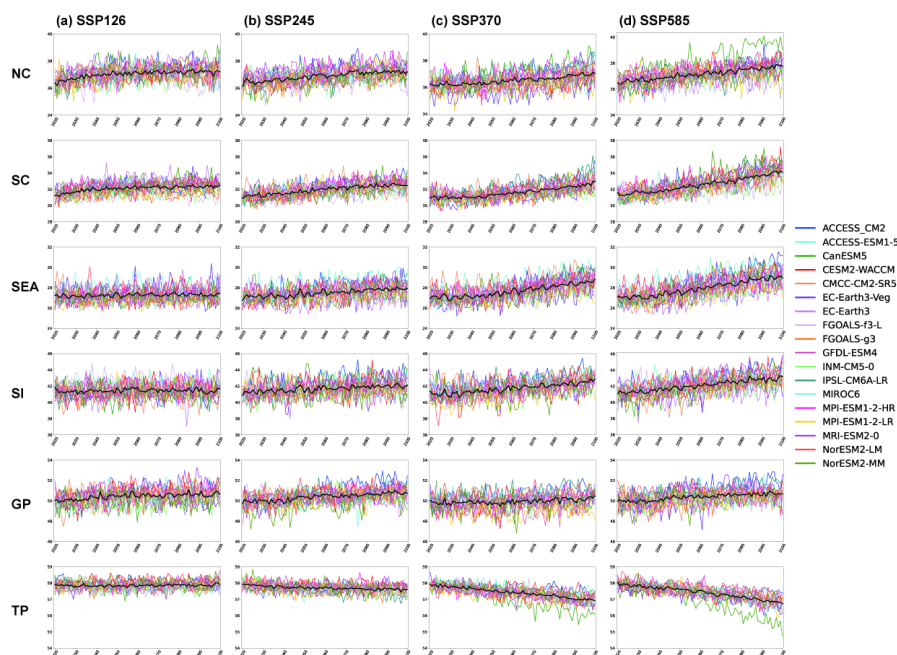
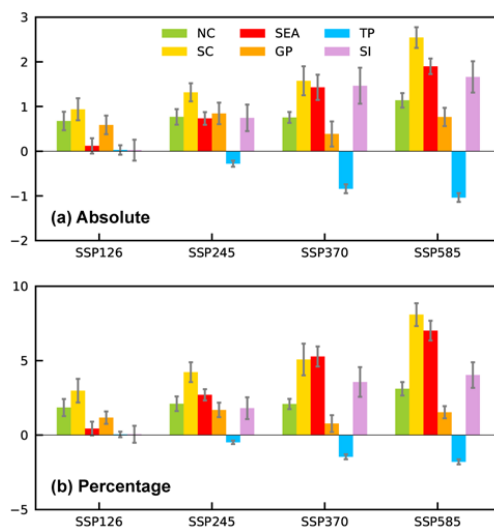



Figure 6. Time series (2020–2100) of annual mean near-surface O$_3$ concentrations (ppb) driven by climate change under the four scenarios (SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5) over North China (NC, 35°–41°N, 105°–120°E), South China (SC, 22°–33.5°N, 105°–120°E), Southeast Asia (SEA, -9.5°S–19°N, 93.75°–140°E), South India (SI, 8°–18°N, 73.125°–80.625°E), Gangetic Plains (GP, 21.5°–23.5°N, 85.625°–92.5°E, 23.5°–27°N, 76.25°–92.5°E, and 27°–30°N, 76.25°–81.25°E), and Tibetan Plateau (TP, 28°–31°N, 81.875°–102.5°E and 31°–38°N, 73.125°–102.5°E). The black lines are the averages of the predicted O$_3$ based on meteorological fields from 18 CMIP6 models.

x

y

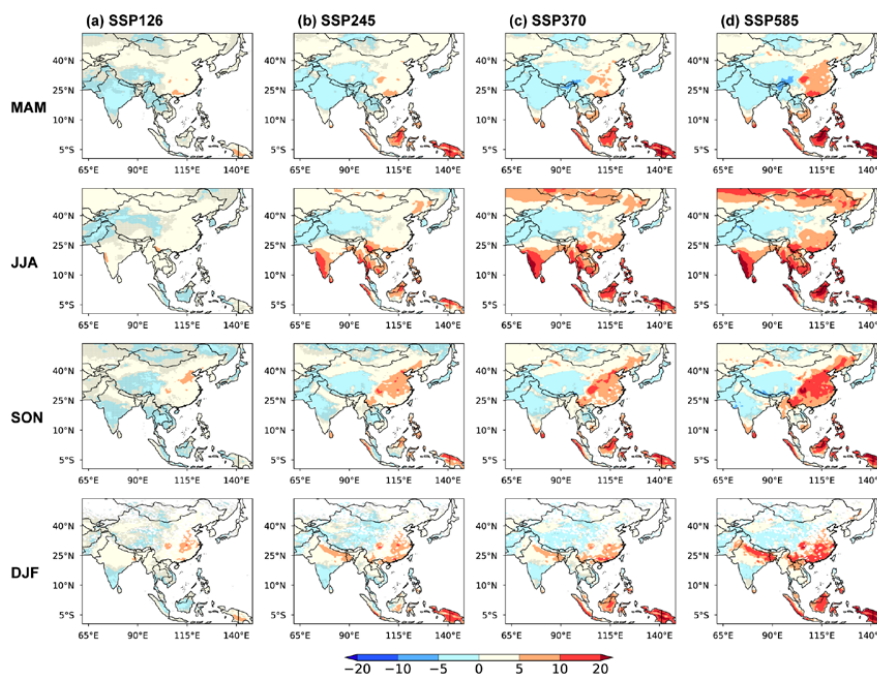



Figure 8 The spatial distributions of percentage differences (%) in near-surface O₃ concentrations between 2025 (2020–2029 mean) and 2095 (2091–2100 mean) driven by climate change under four scenarios (SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5, from left to right) averaged in MAM (March–April–May), JJA (June–July–August), SON (September–October–November), and DJF (December–January–February) (from top to bottom). No overlaying hatch pattern indicates statistical significance with 95% confidence from a two-tailed t test.