Estimating ground-level CO concentrations across China based on national monitoring network and MOPITT: Potentially overlooked CO hotspots in the Tibetan Plateau

Dongren Liu [a], Baofeng Di [a,b], Yuzhou Luo [c], Xunfei Deng [d], Hanyue Zhang [a], Fumo Yang [a,e],
Michael L. Grieneisen [c], Yu Zhan [a,e,f,g] *

[a] Department of Environmental Science and Engineering, Sichuan University, Chengdu 610065, China

[b] Institute for Disaster Management and Reconstruction, Sichuan University, Chengdu 610200, China

[c] Department of Land, Air, and Water Resources, University of California, Davis, CA 95616, United States

[d] Institute of Digital Agriculture, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China

[e] National Engineering Research Center for Flue Gas Desulfurization, Chengdu 610065, China

[f] Sino-German Centre for Water and Health Research, Sichuan University, Chengdu 610065, China

[g] Medical Big Data Center, Sichuan University, Chengdu 610041, China

Corresponding to: Yu Zhan (yzhan@scu.edu.cn)

Abstract. Given its relatively long lifetime in the troposphere, carbon monoxide (CO) is commonly employed as a tracer for characterizing airborne pollutant distributions. The present study aims to estimate the spatiotemporal distributions of ground-level CO concentrations across China during 2013-2016. A refined random-forest-spatiotemporal-kriging model (RF-STK) is developed to simulate daily gridded CO concentrations (0.1° grid with 98341 cells) based on the extensive CO monitoring data and the Measurements of Pollution in the Troposphere CO retrievals (MOPITT-CO). The refined RF-STK model alleviates the negative effects of sampling bias and variance heterogeneity on the model training, resulting in cross-validation R^2 of 0.51 and 0.71 for predicting daily and spatial CO concentrations, respectively. The national population-weighted CO concentrations were predicted to be (0.99 ± 0.30) mg m^{-3} ($\mu \pm \sigma$) and showed decreasing trends over all regions of China at a rate of (-0.021 ± 0.004) mg m^{-3} per year. The CO pollution was more severe in North China (1.19 ± 0.30) mg m^{-3}, and the predicted spatial pattern was roughly consistent with the MOPITT-CO. The hotspots in the Central Tibetan Plateau which were overlooked by the MOPITT were revealed by the refined RF-STK predictions. This information has an implication for improving the MOPITT-CO derivation procedure and air quality management.

1 Introduction

Ground-level carbon monoxide is a worldwide atmospheric pollutant posing risks to human health and the environment (Reeves et al., 2002; White et al., 1990). While CO is formed naturally from the oxidation of methane and non-methane volatile organic compounds, anthropogenic emissions from incomplete combustion of fossil fuels and biofuels contribute approximately 42% of the total atmospheric CO (Holloway et al., 2000; Pommier et al., 2013). China is one of the countries with the most severe CO pollution in the world, and the combustion of fossil fuels is the dominant source of anthropogenic CO emissions (Wang et al., 2004; Zhang et al., 2003; Duncan et al., 2007a). Due to its relatively long lifetime in the troposphere (i.e., one to two months), CO is commonly employed as a tracer for characterizing pollution transport (Goldan et al., 2000; Pommier et al., 2010). It is therefore essential to obtain the spatiotemporal distribution of CO throughout China for air quality management. The national air pollution surveillance network in China has been regularly monitoring ground-level CO concentrations since 2013, but these site-based measurements are inadequate to represent the spatially continuous distributions of CO.

Chemical Transport Models (CTM) have been employed to estimate ground-level CO concentrations (Hu et al., 2016; Arellano and Hess, 2006). On the basis of meteorological conditions generated by climate models, CTM simulates reactions, transport, and deposition of chemicals in the atmosphere, which generally requires high computational cost and a large amount of data inputs such as emission inventories. The prediction performance of CTM tends to be affected by uncertainties in the simulation algorithms and emission inventories. A CTM comparison study found that the difference in transport simulation resulted in considerable discrepancies between inter-model CO predictions (Arellano and Hess, 2006; Duncan et al., 2007b). It has been reported that a certain CTM underpredicted the monthly average CO concentrations in China by more than 60% (Hu et al., 2016). Although the emission inventories for China have been refined in recent years, high uncertainties still exist (Li et al., 2017). For instance, biomass combustion, residential biofuel consumption, and transient fire events tend to be underreported, consequently leading to underestimation of CO emissions in the emission inventories (Wang et al., 2002; Streets et al., 2003). Despite underestimation by CTM, the general patterns of CO concentrations are

captured, and they can be used as prior states for deriving posterior estimates based on satellite retrievals (Deeter et al., 2014).

Multiple satellite instruments have been operating to measure atmospheric CO for more than a decade, including MOPITT (Deeter, 2003; Deeter et al., 2010; Emmons et al., 2008), the Atmospheric Infrared Sounder (AIRS) (McMillan, 2005), the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) (Bovensmann et al., 1999; von Hoyningen-Huene et al., 2006), and the Infrared Atmospheric Sounding Interferometer (IASI) (Fortems-Cheiney et al., 2009). Strong absorption lines of CO occur in the thermal infrared (4.7 µm) and solar infrared (2.3 µm) spectral regions. Among the abovementioned satellite instruments, MOPITT is one of few sensors that are capable of measuring ground-level CO based on the instantaneous multispectral retrievals (Streets et al., 2013). The prior state used in MOPITT is simulated by the Community Atmosphere Model with Chemistry (CAM-Chem), which is an example of a CTM. The MOPITT-CO product plays an important role in analyzing spatiotemporal patterns of ground-level CO at large scales (Drummond et al., 2010; Worden et al., 2013). Compared with the site-based in-situ monitoring data, MOPITT-CO provides repeated measures with more extensive spatial coverages. Nevertheless, the sensitivity of MOPITT signals to ground-level CO is affected by the thermal contrast between the ground and atmosphere (Warner et al., 2007; Clerbaux et al., 2009). High uncertainties in CO estimations retrieved from MOPITT have been reported in previous studies, and more efforts are required to improve the estimation accuracy of CO (Peng et al., 2007; Zhao et al., 2006; Li and Liu, 2011).

Machine learning models are applied to predict the spatiotemporal distributions of atmospheric pollutants, such as fine particulate matter ($PM_{2.5}$) and nitrogen dioxide (NO_2), based on the associated satellite retrievals and ground measurements (Zhan et al., 2018; Reid et al., 2015). Complex structures are built to capture nonlinear and high-order interactions between the response and predictor variables. Machine learning models generally show superior prediction performance in the presence of abundant training data (Hastie et al., 2009). In the comparisons of models predicting $PM_{2.5}$ concentrations, random forests and gradient boosting machine, which incorporate the satellite-retrieved aerosol optical depth (AOD), presented conspicuously good prediction performance (Reid et al., 2015). In addition, the random forest and spatiotemporal kriging (RF-STK) model is proposed to predict the ground-level nitrogen dioxide (NO_2) concentrations across China based on the satellite-retrieved NO_2 density (Zhan et al., 2018). To the authors' knowledge, machine learning models have never been employed to estimate nationwide ground-level CO concentrations in China based on the satellite retrievals.

The present study aims to estimate the spatiotemporal distributions of ground-level CO concentrations across China during 2013-2016. A refined RF-STK model is developed to simulate daily gridded CO concentrations (0.1° grid with 98341 cells) based on the CO monitoring data and the MOPITT-CO with the consideration of extensive geographic factors. The strategy of inversely weighting training data by the local population densities is proposed to mitigate the effect of sampling bias towards populous areas. The spatial unit of 0.1° is chosen to be consistent with the series of our previous studies (Zhan et al., 2017; Zhan et al., 2018). The MOPITT-CO are thoroughly evaluated against the CO monitoring data. The discrepancies between the MOPITT-CO and the estimated CO distributions are highlighted both spatially and temporally. The results of this study are valuable for improving the MOPITT retrieval algorithm and air quality management.

2 Materials and methods

2.1 MOPITT-CO retrievals

The MOPITT operational gas correlation spectroscopy CO product (MOP02J.007), containing retrievals of surface CO mixing ratios, is obtained from the Atmospheric Science Data Center (ASDC, 2017). The MOPITT onboard the Terra satellite provides tropospheric CO density with global coverage every three days (Edwards, 2004). The CO surface mixing ratios from the Level-2 data product have a spatial resolution of 22 km at nadir. The Level-2 product has daytime and nighttime data fields, which are highly correlated ($R = 0.99$). This study chose the daytime data over the nighttime data, as the former exhibit higher correlations with the ground-level CO observations than the latter (Table 1). The overall bias of Version 7 is a few percent lower than Version 6 for the thermal infrared (TIR)-only, near infrared (NIR)-only, and TIR/NIR products at all levels (Deeter et al., 2014; Deeter et al., 2017). The TIR/NIR product, which features the maximum sensitivity to near-surface CO, is used throughout the study. Through the Gaussian kernel (Goodfellow et al., 2016), the MOPITT noise was filtered and gaps were filled, and then the CO retrievals were resampled to 0.1 degree grids.

2.2 Ground-level CO observations

Figure 1 shows the locations of the 1656 state-managed sites spread out over all of China, which monitor the ground-level CO concentrations (MEPC, 2017; EPAROC, 2017; EPDHK, 2017). Most of the sites were in the cities of East China, leading to nonnegligible sampling biases. The CO concentrations were measured by the non-dispersive infrared absorption method and the gas filter correlation infrared absorption method (Zhao et al., 2016). Hourly average CO concentrations were collected and cleaned by employing the "three sigma rule" that the values falling outside of ($\mu \pm 3\sigma$) were considered outliers (Kazmier, 2003). Less than 0.01% of the hourly data (values higher than 20.2 mg m^{-3}) were excluded. The days with more than 12-hour observations were included as representative days, and approximately 1.67 million records of daily average CO concentrations were obtained for the subsequent analyses.

2.3 Refined RF-STK model

The refined RF-STK model, consisting of a random forest (RF) submodel and a spatiotemporal Kriging (STK), was refined to predict the ground-level CO concentrations across China. The RF-STK model utilized the strengths of both RF and STK, which had shown the capability of predicting NO_2 concentrations in our previous study (Zhan et al., 2018). In brief, a RF submodel was first trained with the log-transformed ground-level CO observations. As the CO concentrations approximated a lognormal distribution, they were log transformed for variance stabilization (De'Ath and Fabricius, 2000). Leveraging variable selection was conducted based on the pre-experiments. Then, the out-of-bag (OOB) errors of the back-transformed RF predictions were filtered with the "three-sigma-rule", and interpolated with the STK afterwards. Finally, the CO concentrations were predicted as the sums of the STK interpolations and back-transformed RF predictions. It is worth mentioning that the RF algorithm was refined in the present study by inversely weighting each training sample with the surrounding population density to alleviate the effects of sampling bias towards populous areas. The loss function (L) of the refined RF is as follows:

$$L(y, f(x)) = \sum_{n=1}^{N} w_n [y_n - f(x_n)]^2 / \sum_{n=1}^{N} w_n \qquad (1)$$

where w_n is the weight of observation y_n (N observations in total), and $f(x_n)$ is the model prediction.

2.4 Model input data

The predictors of environmental conditions for the random forest submodel covered the meteorological conditions,

5 land uses, emission inventories, elevation, population densities, normalized difference vegetation index (NDVI), and road densities. The meteorological conditions included the atmospheric pressure, air temperature, precipitation, evaporation, relative humidity, insolation duration, wind speed, and planetary boundary layer height (PBLH). Land uses mainly recorded the areas of forests, grasslands, wetlands, artificial surfaces, and waterbodies. The emission inventories comprised emission distributions of 10 major atmospheric chemical constituents, such

10 as CO, NO_X, SO_2, $PM_{2.5}$, and black carbon (BC). The meteorological conditions, except for PBLH, were interpolated to the $0.1°$ grid by using co-kriging with elevation. The elevation, land uses, population densities, NDVI, PBLH, and emission inventories were resampled to the $0.1°$ grid by calculating area-weighted means, for which additional predictors were generated by applying spatial convolution with Gaussian kernels. The spatial convolution smoothed spatial transition and took into account neighboring effects (Goodfellow et al., 2016).

15 Please refer to section S.1 and Table S8 in the Supporting Information (SI) for the detailed descriptions and data sources of the environmental conditions.

2.5 Model evaluation

The prediction performance and the predictor effects of the refined RF-STK model were investigated. The 10-fold site-based cross-validation was applied to evaluate the performance of the refined RF-STK model in

20 predicting the CO concentrations. All the monitoring sites were approximately evenly divided into ten groups. In each iteration, nine groups were used to develop a model, and the remaining group was used for validation. The training and prediction steps were duplicated 10 times in order to make ground-level CO observations had paired predictions. Various statistical metrics, such as the coefficient of determination (R^2), root mean square error (RMSE), and mean normalized error (MNE), were used to reflect the prediction performance. In addition, the

25 measures of variable importance and partial dependence plots were employed to evaluate the predictor effects. The improvement in the split-criterion attributed to a predictor variable measures its relative importance in the model. A partial dependence plot illustrated the effect of a predictor on the CO concentrations after accounting for the average effects of all the other predictors (Friedman, 2001; Hastie et al., 2009).

2.6 Spatiotemporal analyses

30 Detailed spatiotemporal analyses were performed to investigate the correlation strength between the MOPITT-CO and ground-level CO observations, as well as the distributions of the ground-level CO predictions. The whole nation was divided into seven conventional regions, including Central, East, North, Northeast, Northwest, South, and Southwest China (Fig. 1). For each region, the effectiveness of the MOPITT-CO was evaluated by estimating its correlation with the ground-level CO observations at daily, seasonal, and annual scales. In addition, the

35 seasonal/annual average concentrations maps were delineated based on the full-coverage CO predictions. The population-weighted averages of MOPITT-CO (M_{PW}) and ground-level CO predictions (C_{PW}) were summarized for the whole nation and by regions. The temporal trends of the national and regional M_{PW} and C_{PW} were evaluated

by conducting linear regression on the time series of monthly averages that were deseasonalized by the loess smoothers (Cleveland, 1990). More detailed analyses were conducted for the North China Plain (NCP) and the Central Tibetan Plateau (CTP). While the air pollution in NCP has been well recognized, the air quality in CTP is usually considered to be pristine. Nevertheless, CTP was identified as a potentially overlooked CO hotspot in the present study.

2.7 Computing environment

The data processing and modeling were mainly performed using python and R (Team, 2018). The *scikit-learn* python package was used to develop random forests (Pedregosa et al., 2012). The spatial operations, such as spatiotemporal kriging were conducted by using the R packages of *gstat* (Gräler et al., 2016), *rgdal* (Bivand et al., 2017), and *sp* (Pebesma and Bivand, 2005).

3 Results and discussion

3.1 Descriptive statistics of MOPITT-CO and ground-level CO observations

The results of MOPITT-CO, with an overall coverage rate of (3.5 ± 0.5) % (mean ± standard deviation), show that the surface CO level for China was (0.25 ± 0.18) mg m^{-3} during 2013-2016 (Table S1). The MOPITT-CO values approximate a lognormal distribution, with a median of 0.20 mg m^{-3} and an interquartile range (IQR) of 0.17 mg m^{-3}. The MOPITT-CO has the highest coverage in fall $(4.2 \pm 1.9\%)$ and lowest in summer $(2.9 \pm 1.5\%)$ (Table S2). Southwest China, especially the Sichuan Basin, has the lowest coverage ($< 1\%$) among all the regions (Fig. S1). The sparse coverages of MOPITT-CO limit its utility for representing time-series daily CO concentrations across China.

The daily variation obtained in this study is adequate to indicate the change in regulation of CO concentrations noted in the previous study (Xu et al., 2014) (Fig. S2). The ground-level CO observations show that the average CO concentrations for China was (1.07 ± 0.74) mg m^{-3} during 2013-2016. The ground-level CO observations also approximate a lognormal distribution, with a median of 0.90 mg m^{-3} and IQR of 0.69 mg m^{-3}. High CO concentrations (daily average > 4.0 mg m^{-3}) were observed in 704 monitoring sites, with (7.6 ± 0.8) days per year (CREAS and CNEMC, 2012). The CO concentrations show a strong seasonality, ranging from (0.81 ± 0.17) mg m^{-3} in summer to (1.39 ± 0.38) mg m^{-3} in winter (Table S3). The national annual average of CO concentrations decreased by 6.4% from year 2013 to 2016. Note that the scale of monitoring network was not constant, and the number of monitoring sites grew from 743 to 1603 during these four years. However, the monitoring stations were still sparse in the western China throughout the monitoring period (Fig. 1). The spatially imbalanced monitoring tends to introduce bias and uncertainty to the spatiotemporal statistics of CO concentrations.

3.2 MOPITT-CO evaluation against ground-level CO observations

The spatiotemporal pattern of the MOPITT-CO was generally consistent with that of the ground-level CO observations in China, with $r = 0.43$ for the multiyear averages and $r = 0.37$ for the daily values during 2013-2016 (Table 1). The MOPITT-CO satisfactorily reflected the west-east gradient and the seasonality (i.e., low in warm seasons and high in cold seasons) of ground-level CO concentrations (Figs. 5, 8, and S5). More severe CO

Atmospheric
Chemistry
and Physics
Discussions

pollution in the eastern part of China (including North, East, Central, and South China) indicated more intensive anthropogenic emissions (Fig. S3). At both national and regional scales, the correlation coefficients between ground-level CO observations and MOPITT-CO were generally higher in winter than the other three seasons. In addition, the correlation strength of daily values exhibited considerable spatial heterogeneity, and the r values

5   ranged from 0.58 for South China to 0.17 for Southwest China (Table S4). As expected, it was difficult to capture the CO variations under highly complex geographic conditions for Southwest China, and the high uncertainty in the emission inventories undermined the representativeness of MOPITT-CO for that region. Especially for CTP, we found that the MOPITT-CO was almost completely unaffected by the variations of ground-level CO observations, with r = -0.03 in contrast to r = 0.34 for NCP (Table 1). The CO hotspots observed in CTP (e.g.,

10   Naqu and Qamdo) were not recognized by MOPITT-CO, which even falsely showed the opposite seasonality of ground-level CO observations (Table S3; Figs. 2 and 5).

The discrepancies between the MOPITT-CO and the ground-level CO observations could be mainly attributed to the low sensitivity of the satellite instrument to the ground-level CO variation and the high uncertainty associated with the a priori states for the satellite retrieval. The measurement's lack of sensitivity to ground-level CO

15   variation caused the measurement error, which represents the retrieval uncertainty due to uncertainties in the measured radiances (associated with the instrumental noises) (ASDC, 2017). The inferred emission estimates are influenced by the emission inventories and modeling errors (i.e., CAM-Chem model) which potentially result in errors in the posterior estimation (Dekker et al., 2017). The CO emission intensities for China have been reported to be largely underestimated (Streets et al., 2003; Wang et al., 2004), which might explain the fact that the

20   MOPITT-CO was approximately half of the ground-level CO observations (Fig. 8). The results of a national-scale CTM study for China showed that the ground-level CO predictions were only about one third of the actual observations (Hu et al., 2018). Especially for the CTP, the inadequate information about the CO emissions could be the main reason why MOPITT-CO overlooked potential CO hotpots, whereas some densely populated cities (such as Naqu, Lhasa, and Qamdo) had high values of ground-level CO concentrations (Chen et al., 2019). The

25   environmental conditions of the CTP are characterized by high latitude, low temperature, and low pressure. In addition, high mountains like Qinghai-Tibet Plateau are considered to be obsrtuction which can impede pollutants transporting from South and East Asia (Yin et al., 2017). Naqu is sandwiched between the Tanggula and the Nyainqen Tanglha Mountains (Fig. 1). In the CTP, biomass (e.g., yak dung) combustion is widely used for energy but has low utilization efficiency, resulting in considerable CO emissions (Wen and Tu, 2011; Cai and Zhang,

30   2006; Xiao et al., 2015). The population in both Naqu and Qamdo are over one million, leading to relatively intensive anthropogenic activities in the CTP (NBS, 2010). It is nevertheless laborious and challenging to improve the accuracy of emission inventories, which are essential for deriving the MOPITT-CO. It is therefore critical to assimilate the MOPITT-CO and the ground-level CO observations for deriving the full-coverage ground-level CO distributions (Hooghiemstra et al., 2011).

35   **3.3 Prediction performance of the refined RF-STK**

On the basis of the cross-validation, the refined RF-STK model showed reasonable performance in predicting the daily ground-level CO concentrations, with R^2 of 0.51, RMSE of 0.54 mg m^{-3}, and MFE of 0.35 (Table S5 and Fig. 3). The prediction performance was generally stable across regions and seasons (Fig. S7), and was comparable or superior to the previous modelling performance (Table S6). In the absence of nationwide statistical modelling

work, only CTM studies were found for large scales in China. The previous CTM work for China underestimated the ground-level CO concentrations by 67.2%, with MFE of 0.96. For the RF-submodel, a concise structure was achieved through the variable selection, and spurious prediction details were excluded (Fig. S8).

The RF-submodel with the full variable set generated sharp boundaries circling the desert areas in Northwest China, which became blurred in the predictions of the reduced RF-submodel (Fig. S6). In addition, the coordinate variables (i.e., latitude and longitude) were not considered as candidate variables for the RF-submodel, as artificial strips emerged in the prediction maps after including them as was illustrated in our previous study (Zhan et al., 2017). For the STK-submodel, the predictions were further fine-tuned based on the spatiotemporal patterns of the RF-submodel residuals. Consequently, the cross-validation slope increased from 0.55 to 0.64 (Table S7), suggesting an improvement in capturing the high and low concentrations.

Compared to the original RF-STK model proposed in our previous study (Zhan et al., 2018), two modifications were made in the refined RF-STK model, including sample weighting and logarithm transformation of the response variable (i.e., ground-level CO observations). Inversely weighting the training data by their associated population densities could alleviate the effects of sampling bias towards populous areas for the national air quality monitoring network. As a result, the CO monitoring data from the sparsely populated areas (e.g., the Tibetan Plateau) gained higher weights in the model training process, leading to more realistic predictions for the unmonitored grid cells in those areas. In addition, observations with higher variation would naturally gain higher weights during model training given the loss function of squared errors, for which it was suggested to transform the predictor variable to achieve homogeneous variation (De'Ath and Fabricius, 2000). The ground-level CO observations were heavy-tailed distributed, and hence logarithm transformation was conducted for variance homogenization prior to training the RF-submodel. Compared with the original RF-submodel, the refined RF-submodel showed similar performance in the cross-validation and predicted more realistic spatial distributions of ground-level CO across China (Table S7 and Fig. S6). The predicted map of the original RF-submodel showed the prevalence of higher concentrations than those of the refined RF-submodel, due to overweighting of the training data from the areas with more serve CO pollution.

3.4 Important predictors

The predictor of MOPITT-CO, as the most important predictor in the RF-submodel, exhibited relative importance of 9.4%, and the emission-related predictors together accounted for 30.0% of the total predictor importance (Fig. 4). Absence of single dominant predictor in the model is consistent with the low or moderate correlations between the predictors and ground-level CO concentrations (Fig. S9). The partial dependence plots, which delineated the marginal effects, illustrated the complicated relationships between the predictors and the ground-level CO predictions (Fig. S10). The MOPITT-CO and most of the emission-related predictors were processed with convolution, suggesting the feasibility of incorporating neighbor conditions in the model. The derivation of MOPITT-CO was based on the emission inventories and meteorological fields, and its relatively high importance in the RF-submodel was expected. Nevertheless, the high uncertainties pertaining to the derivation process prevented the MOPITT-CO from playing a dominant role in the model. Among the emission-related predictors, the spatial-convolution-processed emission of organic carbon (OC) was the most important predictor (importance: 8.5%). Moreover, OC and BC emission sources provided valuable additional information on anthropogenic

emissions, to which industrial and residential activities contributed more than 80% (Table S9). Given the high intercorrelations among the predictors associated with anthropogenic activities, only the most relevant predictors were retained in the reduced RF-submodel (Figs. S3 and S9). The predictors which were either redundant or of high uncertainty, e.g., the emission of CO, were removed by the variable selection process.

5   As the most important group, the meteorological conditions together account for 35.6% of the total predictor importance (Fig. 4). The relative importance of temperature, evaporation, wind speed, atmospheric pressure, PBLH, relative humidity, and insolation duration ranged from 2.8 to 8.6%. Usually, stagnant weather conditions occur more frequently in winter, which is characterized by shallow mixed layer, less precipitation and slow wind speed. This can capture pollutants discharged by local emissions or transported from other areas and increased

10   their surface concentration (Wang et al., 2014). Similar to other airborne pollutants, the CO concentrations are sensitive to meteorological conditions (Xu et al., 2011). Apparently negative associations of the CO concentrations with the PBLH and wind speed were delineated by the corresponding partial dependence plots (Fig. S10). High PBLH creates favorable dispersion conditions and hence leads to diluted CO concentrations. Considering the relatively long half-life of CO in the troposphere (Goldan et al., 2000), advection associated with

15   wind speed plays an important role in the pollutant dissipation. It should be noted that the partial dependence plot illustrated the overall relationship, which tended to be mixed by both spatial and temporal variations. For instance, the partial dependence plot for temperature, with a peak around 20°C, is contrary to empirical knowledge, i.e., CO concentrations are known to be highest in winter. This "abnormal" relationship results from the fact that CO-polluted areas are generally distributed in the warmer zones of China.

20   **3.5 Spatiotemporal predictions of ground-level CO by refined RF-STK**

The spatiotemporal patterns of the refined RF-STK predictions were generally similar to those of the MOPITT-CO and reflected extensive adjustments based on the ground-level CO observations (Table 1; Figs. 2, 6, and 7). The predictions of the refined RF-STK adequately maintained the information of ground-level CO observations (ground truth), with $R = 0.95$ for the daily values (Table 1). At the national level, the multiyear (i.e., 2013-2016)

25   averages of CO concentrations were predicted to be (0.99 ± 0.30) mg m^{-3}, with the highest level of (1.32 ± 0.49) mg m^{-3} for winter and the lowest level of (0.77 ± 0.22) mg m^{-3} for summer (Table 2 and Fig. 6). Spatially, the CO concentrations were predicted to be the highest in North China and the lowest in South China, with the multiyear averages of (1.19 ± 0.30) and (0.77 ± 0.18) mg m^{-3}, respectively (Table 2). It is worth noting that the refined RF-STK predictions realistically retrieved the CO hotspots in the CTP, which were overlooked by the MOPITT-CO

30   (Figs. 2, 7, and S5). The "abnormal" CO seasonality (i.e., low in winter and high in summer) for CTP characterized by the MOPITT-CO was corrected in the refined RF-STK predictions even though the data quality of ground-level CO observations for 2013 were in question (Fig. 5). The neglected CO hotspots may be subject to the dual influences of the low combustion efficiency of the domestic stoves and the combustion of a large amount of biomass energy (Chen et al., 2015). For example, burning yak dung accounted for more than 50% of the energy

35   consumption for Nagqu of CTP (Yang and Zheng, 2015).

During 2013-2016, the predicted national average C_{PW} of ground-level CO for China decreased from (1.02 ± 0.34) to (0.95 ± 0.30) mg m^{-3} at a rate of (-0.021 ± 0.004) mg m^{-3} per year ($P<0.01$; Table 2 and Fig. 8). The relative decrease rate of 4.4% was similar to the 3.8% drop of coal consumption for China during 2013-2016 (CSY, 2010),

Atmospheric
Chemistry
and Physics
Discussions

indicating the importance of coal consumption to CO pollution. The relatively decreasing rate of CO was similar to that of NO_2, i.e., 3% (Zhan et al., 2018), but much slower than the rate of $PM_{2.5}$ (Ma et al., 2016). The regional average C_{PW} significantly decreased for all regions ($P<0.05$) except Southwest China ($P = 0.16$). The decreasing trend was most prominent for North China where CO pollution was the most severe, and the decreasing rate was

5   (-0.028 ± 0.008) mg m^{-3} per year. In comparison to the refined RF-STK predictions (which were similar to ground-level CO observations due to the good model fitness), the MOPITT-CO tended to underestimate the decreasing trends of ground-level CO concentrations to a certain extent (Fig. 8). The absolute decreasing rate of M_{PW} was approximately 60% lower than that of the refined RF-STK predictions. The M_{PW} showed no significant trends for East, Northeast, Northwest, South, and Southwest China ($P>0.05$). Accurate evaluation of the temporal trends is

10 essential for air quality management. The refined RF-STK predictions that assimilates the MOPITT-CO with ground-level CO observations provide more solid information for decision making.

4 Conclusions

The spatiotemporal distributions of ground-level CO concentrations for China during 2013-2016 are derived by using the refined RF-STK model to fuse the data from the remote sensing and the ground-level observations. The

15 RF-STK model shows feasible performance in predicting the daily CO concentrations on the 0.1o grid. As most of the monitoring sites are in urban areas, the refined RF-STK model inversely weights the ground-level observations by the surrounding population densities. Given the fact of monitoring sites clustered in cities, it is critical to take into account the effects of sampling bias on modeling the spatiotemporal distributions of atmospheric pollutants. While the general patterns are well depicted by the MOPITT-CO, the fine-scale

20 distributions are sharpened and corrected with the observations from the air quality monitoring network. By using the data-fusion approach, we obtain the comprehensive dataset of ground-level CO concentrations for China.

On the basis of the spatiotemporal predictions, the population-weighted average of ground-level CO concentrations was 0.99 ± 0.30 mg m-3 for China during 2013-2016, with an annual decreasing rate of -0.021 ± 0.004 mg m-3 per year. The CO concentrations were predicted to be the highest in North China (1.19 ± 0.30 mg

25 m-3) and the lowest in South China (0.77 ± 0.18 mg m-3). The seasonal averages of the whole China ranged from 0.77 ± 0.22 in summer to 1.32 ± 0.49 mg m-3 in winter, attributing to the seasonality of weather conditions and emission intensities as indicated by the variable importance of the RF-STK model. In summary, the present study provides important information for improving the MOPITT-CO derivation procedure, such as refining the prior status assigned to the overlooked hotspots in the Central Tibetan Plateau. The estimates of ground-level CO

30 distributions are valuable for air quality management and human exposure assessment in China, given the utility of CO for indicating air quality.

Code availability. The code for random forest is available from scikit-learn (https://scikit-learn.org/stable/). The code for spatiotemporal kriging is available from the Comprehensive R Archive Network (https://cran.r-project.org/web/packages/gstat/index.html).

35 *Data availability*. The hourly CO concentration data are from the Ministry of Ecology and Environment of the People's Republic of China (http://datacenter.mep.gov.cn/). The MOPITT data are from the Atmospheric

Atmospheric
Chemistry
and Physics
Discussions

Science Data Center (https://eosweb.larc.nasa.gov/). The estimated ground-level CO concentrations are available upon request.

Author contributions. DL and YZ performed research and wrote manuscript. BD, YL, XD, and HZ analyzed data. FY and MLG provided extensive comments on manuscript.

5 *Competing interests*. The authors declare that they have no conflict of interests.

References

Arellano, A. F., and Hess, P. G.: Sensitivity of top-down estimates of CO sources to GCTM transport, Geophys. Res. Lett., 33, 10.1029/2006gl027371, 2006.

The Level-2 Product of MOPITT Version 7: https://eosweb.larc.nasa.gov/project/mopitt/mop02j_v007, access:

5   18 Mar 2018, 2017.

Rgdal: Bindings for the 'Geospatial' Data Abstraction Library: https://CRAN.R-project.org/package=rgdal, access: 19 July 2017, 2017.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noel, S., Rozanov, V., Kelly, Goede, and A., P. H.: SCIAMACHY - Mission Objectives and Measurement Modes, J. Atmos. Sci., 56, 125--150, 10.1175/1520-

10   0469(1999)056<0127:SMOAMM>2.0.CO;2, 1999.

Cai, G., and Zhang, L.: The Present Situation and Development of Energy Utilization in Tibet (in Chinese with English abstract), Energy. China., 28, 38-42, 2006.

Chen, P., Kang, S., Bai, J., Sillanpää, M., and Li, C.: Yak dung combustion aerosols in the Tibetan Plateau: Chemical characteristics and influence on the local atmospheric environment, Atmospheric Research, 156, 58-

15   66, 10.1016/j.atmosres.2015.01.001, 2015.

Chen, P., Yang, J., Pu, T., Li, C., Guo, J., Tripathee, L., and Kang, S.: Spatial and Temporal Variations of Gaseous and Particulate Pollutants in Six Sites in Tibet, China, during 2016–2017, Aerosol and Air Quality Research, 19, 516-527, 10.4209/aaqr.2018.10.0360, 2019.

Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M.,

20   Razavi, A., and Turquety, S.: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder, Atmos. Chem. Phys., 9, 724-725 passim, 2009.

Cleveland, R. B.: STL : A Seasonal-Trend Decomposition Procedure Based on Loess, J. Off. Stat., 6, 3-33, 1990.

CREAS, and CNEMC: Ambient Air Quality Standard (in Chinese), No. 45, Pingli West Avenue, Xicheng District, Beijing, 2012.

25   China Statistical Yearbook: http://www.stats.gov.cn/tjsj/ndsj/, access: 1 June 2018, 2010.

De'Ath, G., and Fabricius, K. E.: Classification and Regression Trees: A Powerful Yet Simple Technique for Ecological Data Analysis, Ecology., 81, 3178-3192, 2000.

Deeter, M. N.: Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument, J. Geophys. Res., 108, 10.1029/2002jd003186, 2003.

30   Deeter, M. N., Edwards, D. P., Gille, J. C., Emmons, L. K., Francis, G., Ho, S. P., Mao, D., Masters, D., Worden, H., Drummond, J. R., and Novelli, P. C.: The MOPITT version 4 CO product: Algorithm enhancements, validation, and long-term stability, J. Geophys. Res., 115, 10.1029/2009jd013005, 2010.

Deeter, M. N., Martínez-Alonso, S., Edwards, D. P., Emmons, L. K., Gille, J. C., Worden, H. M., Sweeney, C., Pittman, J. V., Daube, B. C., and Wofsy, S. C.: The MOPITT Version 6 product: algorithm enhancements and

35   validation, Atmos. Meas. Tech., 7, 3623-3632, 10.5194/amt-7-3623-2014, 2014.

Deeter, M. N., Edwards, D. P., Francis, G. L., Gille, J. C., Martínez-Alonso, S., Worden, H. M., and Sweeney, C.: A climate-scale satellite record for carbon monoxide: the MOPITT Version 7 product, Atmospheric Measurement Techniques, 10, 2533-2555, 10.5194/amt-10-2533-2017, 2017.

Atmospheric
Chemistry
and Physics
Discussions

Dekker, I. N., Houweling, S., Aben, I., Röckmann, T., Krol, M., Martínez-Alonso, S., Deeter, M. N., and Worden, H. M.: Quantification of CO emissions from the city of Madrid using MOPITT satellite retrievals and WRF simulations, Atmos. Chem. Phys., 17, 14675-14694, 10.5194/acp-17-14675-2017, 2017.

Drummond, J. R., Zou, J., Nichitiu, F., Kar, J., Deschambaut, R., and Hackett, J.: A review of 9-year
5   performance and operation of the MOPITT instrument, Adv. Space Res., 45, 760-774, 10.1016/j.asr.2009.11.019, 2010.

Duncan, B. N., Logan, J. A., Bey, I., Megretskaia, I. A., Yantosca, R. M., Novelli, P. C., Jones, N. B., and Rinsland, C. P.: Global budget of CO, 1988-1997: Source estimates and validation with a global model, J. Geophys. Res-Atmos., 112, 10.1029/2007jd008459, 2007a.

10   Duncan, B. N., Strahan, S. E., Yoshida, Y., Steenrod, S. D., and Livesey, N.: Model study of the cross-tropopause transport of biomass burning pollution, Atmos. Chem. Phys., 7, 3713-3736, 10.5194/acp-7-3713-2007, 2007b.

Edwards, D. P.: Observations of carbon monoxide and aerosols from the Terra satellite: Northern Hemisphere variability, J. Geophys. Res., 109, 10.1029/2004jd004727, 2004.

15   Emmons, L. K., Edwards, D. P., Deeter, M. N., Gille, J. C., Campos, T., Nédélec, P., Novelli, P., and Sachse, G.: Measurements of Pollution In The Troposphere (MOPITT) validation through 2006, Atmos. Chem. Phys. Discuss., 8, 18091-18109, 10.5194/acpd-8-18091-2008, 2008.

Taiwan air quality monitoring networks: http://taqm.epa.gov.tw, access: 20 Feb 2017, 2017.

Hong Kong air quality monitoring data: http://epic.epd.gov.hk/EPICDI/air/station/, access: 18 Mar 2017, 2017.

20   Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Carouge, C., Clerbaux, C., Coheur, P. F., George, M., Hurtmans, D., and Szopa, S.: On the capability of IASI measurements to inform about CO surface emissions, Atmos. Chem. Phys., 9, 8735-8743, 10.5194/acp-9-8735-2009, 2009.

Friedman, J. H.: Greedy Function Approximation: A Gradient Boosting Machine, Ann. Stat., 29, 1189-1232, 2001.

25   Goldan, P. D., Parrish, D. D., Kuster, W. C., Trainer, M., McKeen, S. A., Holloway, J., Jobson, B. T., Sueper, D. T., and Fehsenfeld, F. C.: Airborne measurements of isoprene, CO, and anthropogenic hydrocarbons and their implications, Journal of Geophysical Research: Atmospheres, 105, 9091-9105, 10.1029/1999jd900429, 2000.

Goodfellow, I., Bengio, Y., and Courville, A.: Deep Learning, The MIT Press, Cambridge, MA 02142, 2016.

Gräler, B., Pebesma, E., and Heuvelink, G. B. M.: Spatio-Temporal Interpolation using gstat, R. J., 204-218,
30   10.1007/978-3-319-17885-1_1647, 2016.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning Data Mining, Inference, and Prediction, Second Edition, World Book Inc, Beijing, 192-192 pp., 2009.

Holloway, T., Levy, H., and Kasibhatla, P.: Global distribution of carbon monoxide, J. Geophys. Res-Atmos., 105, 12123-12147, 10.1029/1999jd901173, 2000.

35   Hooghiemstra, P. B., Krol, M. C., Meirink, J. F., Bergamaschi, P., van der Werf, G. R., Novelli, P. C., Aben, I., and Röckmann, T.: Optimizing global CO emission estimates using a four-dimensional variational data assimilation system and surface network observations, Atmos. Chem. Phys., 11, 4705-4723, 10.5194/acp-11-4705-2011, 2011.

Hu, J., Chen, J., Ying, Q., and Zhang, H.: One-year simulation of ozone and particulate matter in China using
40   WRF/CMAQ modeling system, Atmos. Chem. Phys., 16, 10333-10350, 10.5194/acp-16-10333-2016, 2016.

Hu, Z., Yu, S., Kan, Z., Cao, Z., and Lou, D.: Effect of altitude on Emission of Diesel engine starting pollutants (in Chinese with English abstract), China. Environ Sc., 38, 91-96, 2018.

Kazmier, L.: Schaum's Outline of Business Statistics (eBook), MCGRAW-HILL, New York, 2003.

Li, L., and Liu, Y.: Space-borne and ground observations of the characteristics of CO pollution in Beijing, 2000–
5   2010, Atmospheric Environment, 45, 2367-2372, 10.1016/j.atmosenv.2011.02.026, 2011.

Li, M., Zhang, Q., Kurokawa, J.-i., Woo, J.-H., He, K., Lu, Z., Ohara, T., Song, Y., Streets, D. G., Carmichael, G. R., Cheng, Y., Hong, C., Huo, H., Jiang, X., Kang, S., Liu, F., Su, H., and Zheng, B.: MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP, Atmos. Chem. Phys., 17, 935-963, 10.5194/acp-17-935-2017, 2017.

10   Ma, Z., Hu, X., Sayer, A. M., Levy, R., Zhang, Q., Xue, Y., Tong, S., Bi, J., Huang, L., and Liu, Y.: Satellite-Based Spatiotemporal Trends in PM2.5 Concentrations: China, 2004-2013, Environ. Health. Perspect., 124, 184-192, 10.1289/ehp.1409481, 2016.

McMillan, W. W.: Daily global maps of carbon monoxide from NASA's Atmospheric Infrared Sounder, Geophysical Research Letters, 32, 10.1029/2004gl021821, 2005.

15   MEPC Air quality daily report for China: http://datacenter.mep.gov.cn/, access: 30 Jan 2017, 2017.

The China's sixth census: http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm, access: 19 July 2017, 2010.

Classes and Methods for Spatial Data in R: the sp Package: https://cran.r-project.org/doc/Rnews/. access: 17 May 2017, 2005.

Pedregosa, F., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R.,
20   Dubourg, V., and Vanderplas, J.: Scikit-learn: Machine Learning in Python, J. Mach. Learn, Res., 12, 2825-2830, 2012.

Peng, L., Zhao, C., Lin, Y., Zheng, X., Tie, X., and Chan, L. Y.: Analysis of carbon monoxide budget in North China, Chemosphere., 66, 1383-1389, 10.1016/j.chemosphere.2006.09.055, 2007.

Pommier, M., Law, K. S., Clerbaux, C., Turquety, S., Hurtmans, D., Hadji-Lazaro, J., Coheur, P. F., Schlager, H.,
25   Ancellet, G., Paris, J. D., Nédélec, P., Diskin, G. S., Podolske, J. R., Holloway, J. S., and Bernath, P.: IASI carbon monoxide validation over the Arctic during POLARCAT spring and summer campaigns, Atmos. Chem. Phys., 10, 10655-10678, 10.5194/acp-10-10655-2010, 2010.

Pommier, M., McLinden, C. A., and Deeter, M.: Relative changes in CO emissions over megacities based on observations from space, Geophysical Research Letters, 40, 3766-3771, 10.1002/grl.50704, 2013.

30   Reeves, C. E., Penkett, S. A., Bauguitte, S., Law, K. S., Evans, M. J., Bandy, B. J., Monks, P. S., Edwards, G. D., Phillips, G., Barjat, H., Kent, J., Dewey, K., Schmitgen, S., and Kley, D.: Potential for photochemical ozone formation in the troposphere over the North Atlantic as derived from aircraft observations during ACSOE, Journal of Geophysical Research: Atmospheres, 107, ACH 14-11-ACH 14-14, 10.1029/2002jd002415, 2002.

Reid, C. E., Jerrett, M., Petersen, M. L., Pfister, G. G., Morefield, P. E., Tager, I. B., Raffuse, S. M., and Balmes,
35   J. R.: Spatiotemporal prediction of fine particulate matter during the 2008 northern California wildfires using machine learning, Environmental Science and Technology, 49, 3887-3896, 10.1021/es505846r, 2015.

Streets, D. G., Yarber, K. F., Woo, J. H., and Carmichael, G. R.: Biomass burning in Asia: Annual and seasonal estimates and atmospheric emissions, Global Biogeochemical Cycles, 17, 1099, 10.1029/2003gb002040, 2003.

Streets, D. G., Canty, T., Carmichael, G. R., de Foy, B., Dickerson, R. R., Duncan, B. N., Edwards, D. P.,
40   Haynes, J. A., Henze, D. K., Houyoux, M. R., Jacob, D. J., Krotkov, N. A., Lamsal, L. N., Liu, Y., Lu, Z.,

Martin, R. V., Pfister, G. G., Pinder, R. W., Salawitch, R. J., and Wecht, K. J.: Emissions estimation from satellite retrievals: A review of current capability, Atmos. Environ., 77, 1011-1042, 10.1016/j.atmosenv.2013.05.051, 2013.

R: A Language and Environment for Statistical Computing: https://www.R-project.org/, access: 11 July 2018,
5   2018.

von Hoyningen-Huene, W., Kokhanovsky, A. A., Wuttke, M. W., Buchwitz, M., Noël, S., Gerilowski, K., Burrows, J. P., Latter, B., Siddans, R., and Kerridge, B. J.: Validation of SCIAMACHY top-of-atmosphere reflectance for aerosol remote sensing using MERIS L1 data, Atmospheric Chemistry and Physics Discussions, 6, 673-699, 10.5194/acpd-6-673-2006, 2006.

10   Wang, T., Cheung, T. F., Li, Y. S., Yu, X. M., and Blake, D. R.: Emission characteristics of CO, NOx, SO2 and indications of biomass burning observed at a rural site in eastern China, J. Geophys. Res.: Atmos., 107, 10.1029/2001JD000724, 2002, 2002.

Wang, Y., Ying, Q., Hu, J., and Zhang, H.: Spatial and temporal variations of six criteria air pollutants in 31 provincial capital cities in China during 2013-2014, Environment international, 73, 413-422,
15   10.1016/j.envint.2014.08.016, 2014.

Wang, Y. X., Mcelroy, M. B., Wang, T., and Palmer, P. I.: Asian emissions of CO and NOx: Constraints from aircraft and Chinese station data, J. Geophys. Res., 109, 10.1029/2004jd005250, 2004.

Warner, J., Comer, M. M., Barnet, C. D., McMillan, W. W., Wolf, W., Maddy, E., and Sachse, G.: A comparison of satellite tropospheric carbon monoxide measurements from AIRS and MOPITT during INTEX-A, Journal of
20   Geophysical Research, 112, 10.1029/2006jd007925, 2007.

Wen, X., and Tu, Y.: Utilization and Prospect of Biomass Energy in Tibet (in Chinese), Tibet. Sc. Technol., 26-28, 2011.

White, J. C., Wagner, W., and Beale, C. N.: Global climate change linkages: acid rain, air quality and stratospheric ozone, Atmos. Environ. Part A, 24, 2898-2899, 1990.

25   Worden, H. M., Deeter, M. N., Frankenberg, C., George, M., Nichitiu, F., Worden, J., Aben, I., Bowman, K. W., Clerbaux, C., Coheur, P. F., de Laat, A. T. J., Detweiler, R., Drummond, J. R., Edwards, D. P., Gille, J. C., Hurtmans, D., Luo, M., Martínez-Alonso, S., Massie, S., Pfister, G., and Warner, J. X.: Decadal record of satellite carbon monoxide observations, Atmos. Chem. Phys., 13, 837-850, 10.5194/acp-13-837-2013, 2013.

Xiao, Q., Saikawa, E., Yokelson, R. J., Chen, P., Li, C., and Kang, S.: Indoor air pollution from burning yak
30   dung as a household fuel in Tibet, Atmos. Environ., 102, 406-412, 10.1016/j.atmosenv.2014.11.060, 2015.

Xu, L., Chen, F., Chen, F., Chen, W., Yu, H., Huang, X., Zeng, Y., Li, X., Hong, S., Feng, Y., and Zhong, X.: Spatial and temporal variation of nearground co concentration in the Eight economic regions in china in may and july (in Chinese with English abstract), Acta. Sci. Circumstantiae., 1934-1941, 10.13671 / j.hjkxxb.2014.0642, 2014.

35   Xu, W. Y., Zhao, C. S., Ran, L., Deng, Z. Z., Liu, P. F., Ma, N., Lin, W. L., Xu, X. B., Yan, P., He, X., Yu, J., Liang, W. D., and Chen, L. L.: Characteristics of pollutants and their correlation to meteorological conditions at a suburban site in the North China Plain, Atmos. Chem. Phys., 11, 4353-4369, 10.5194/acp-11-4353-2011, 2011.

Yang, T., and Zheng, Y.: Study on energy poverty of herdsmen in Naqu region, Tibet (in Chinese), China. Tibet., 127-133, 2015.

Yin, X., Kang, S., de Foy, B., Cong, Z., Luo, J., Zhang, L., Ma, Y., Zhang, G., Rupakheti, D., and Zhang, Q.: Surface ozone at Nam Co in the inland Tibetan Plateau: variation, synthesis comparison and regional representativeness, Atmos. Chem. Phys., 17, 11293-11311, 10.5194/acp-17-11293-2017, 2017.

Zhan, Y., Luo, Y., Deng, X., Chen, H., Grieneisen, M. L., Shen, X., Zhu, L., and Zhang, M.: Spatiotemporal prediction of continuous daily PM 2.5 concentrations across China using a spatially explicit machine learning algorithm, Atmospheric Environment, 155, 129-139, 10.1016/j.atmosenv.2017.02.023, 2017.

Zhan, Y., Luo, Y., Deng, X., Zhang, K., Zhang, M., Grieneisen, M. L., and Di, B.: Satellite-Based Estimates of Daily NO2 Exposure in China Using Hybrid Random Forest and Spatiotemporal Kriging Model, Environmental Science and Technology, 52, 4180-4189, 10.1021/acs.est.7b05669, 2018.

Zhang, M., Uno, I., Carmichael, G. R., Akimoto, H., Wang, Z., Tang, Y., Woo, J.-H., Streets, D. G., Sachse, G. W., Avery, M. A., Weber, R. J., and Talbot, R. W.: Large-scale structure of trace gas and aerosol distributions over the western Pacific Ocean during the Transport and Chemical Evolution Over the Pacific (TRACE-P) experiment, Journal of Geophysical Research: Atmospheres, 108, 10.1029/2002jd002946, 2003.

Zhao, C., Tie, X., Wang, G., Qin, Y., and Wang, P.: Analysis of Air Quality in Eastern China and its Interaction with Other Regions of the World, J. Atmos. Chem., 55, 189-204, 10.1007/s10874-006-9022-1, 2006.

Zhao, S., Yu, Y., Yin, D., He, J., Liu, N., Qu, J., and Xiao, J.: Annual and diurnal variations of gaseous and particulate pollutants in 31 provincial capital cities based on in situ air quality monitoring data from China National Environmental Monitoring Center, Environ. Int., 86, 92-106, 10.1016/j.envint.2015.11.003, 2016.

Atmospheric
Chemistry
and Physics
Discussions

Table 1. Correlations among the ground observations, MOPITT-CO, and the refined RF-STK predictions (Pearson correlation coefficients).

Region/Dataset	Pair[a]	Daily	Monthly	Seasonal	Annual	Spatial[b]
Nation	O-M	0.37	0.40	0.44	0.44	0.43
	O-P	0.95	0.97	0.97	0.97	0.98
	P-M	0.09	0.1	0.1	0.08	0.08
Central Tibetan Plateau (CTP)[c]	O-M	-0.03	-0.04	0.10	-0.17	-0.12
	O-P	0.91	0.92	0.93	0.96	1
	P-M	-0.04	-0.04	-0.05	0.004	0.6
North China Plain (NCP)[c]	O-M	0.34	0.36	0.40	0.30	0.20
	O-P	0.95	0.97	0.98	0.97	0.98
	P-M	0.36	0.40	0.46	0.52	0.57
X1[d]	O-M	0.37	0.40	0.44	0.44	0.43
X2[d]	O-M	0.36	0.39	0.42	0.42	0.40
X1_tc[d]	O-M	0.36	0.44	0.45	0.43	0.41
X1_tc_sc[d]	O-M	0.38	0.47	0.48	0.44	0.42

[a] O: ground-level CO observations; M: MOPITT-CO; and P: refined RF-STK predictions;
[b] Multiyear averages during 2013-2016.
[c] Please refer to Fig. 1 for the locations of CTP and NCP.
[d] X1: Daytime MOPITT-CO; X2: Nighttime MOPITT-CO; X1_tc: X1 processed with temporal convolution; X1_tc_sc: X1_tc processed with spatial convolution; X1 and X2 are highly correlated ($R = 0.99$).

5

Table 2. Population-weighted averages of predicted ground-level CO concentrations during 2013-2016 for China ($\mu\pm\sigma$; mg m^{-3}).[a]

Region	2013	2014	2015	2016	Spring	Summer	Fall	Winter	Annual
Central China	1.02 ± 0.32	1.01 ± 0.23	0.98 ± 0.28	0.96 ± 0.27	0.92 ±0.22	0.78 ± 0.20	0.99 ± 0.27	1.29 ± 0.41	0.99 ± 0.26
East China	0.92 ± 0.16	0.89 ± 0.15	0.88 ± 0.16	0.84 ± 0.14	0.86 ±0.14	0.72 ± 0.12	0.85 ± 0.14	1.12 ± 0.20	0.88 ± 0.14
North China	1.23 ± 0.34	1.20 ± 0.32	1.19 ± 0.30	1.14 ± 0.30	1.02 ±0.26	0.91 ± 0.24	1.17 ± 0.31	1.66 ± 0.48	1.19 ± 0.30
Northeast China	0.89 ± 0.24	0.83 ± 0.31	0.81 ± 0.27	0.79 ± 0.24	0.76 ±0.24	0.65 ± 0.21	0.81 ± 0.26	1.10 ± 0.35	0.83 ± 0.25
Northwest China	1.05 ± 0.38	1.04 ± 0.32	1.01 ± 0.32	1.01 ± 0.31	0.90 ±0.29	0.72 ± 0.18	1.00 ± 0.29	1.49 ± 0.61	1.03 ± 0.32
South China	0.78 ± 0.21	0.81 ± 0.21	0.75 ± 0.18	0.73 ± 0.16	0.79 ±0.18	0.64 ± 0.16	0.74 ± 0.18	0.90 ± 0.21	0.77 ± 0.18
Southwest China	0.79 ± 0.19	0.84 ± 0.20	0.79 ± 0.19	0.79 ± 0.19	0.76 ±0.17	0.66 ± 0.16	0.79 ± 0.18	1.00 ± 0.24	0.80 ± 0.18
Nation	1.02 ± 0.34	1.01 ± 0.30	0.98 ± 0.31	0.95 ± 0.30	0.91 ±0.25	0.77 ± 0.22	0.97 ± 0.30	1.32 ± 0.49	0.99 ± 0.30
CTP	0.86 ± 0.23	0.85 ± 0.23	0.86 ± 0.26	0.86 ± 0.25	0.85 ±0.23	0.77 ± 0.20	0.87 ± 0.25	0.94 ± 0.25	0.86 ± 0.21
NCP	1.31 ± 0.32	1.25 ± 0.26	1.26 ± 0.26	1.22 ± 0.26	1.08 ±0.24	0.94 ± 0.20	1.25 ± 0.25	1.78 ± 0.42	1.26 ± 0.26

[a] σ stands for the spatial variation.
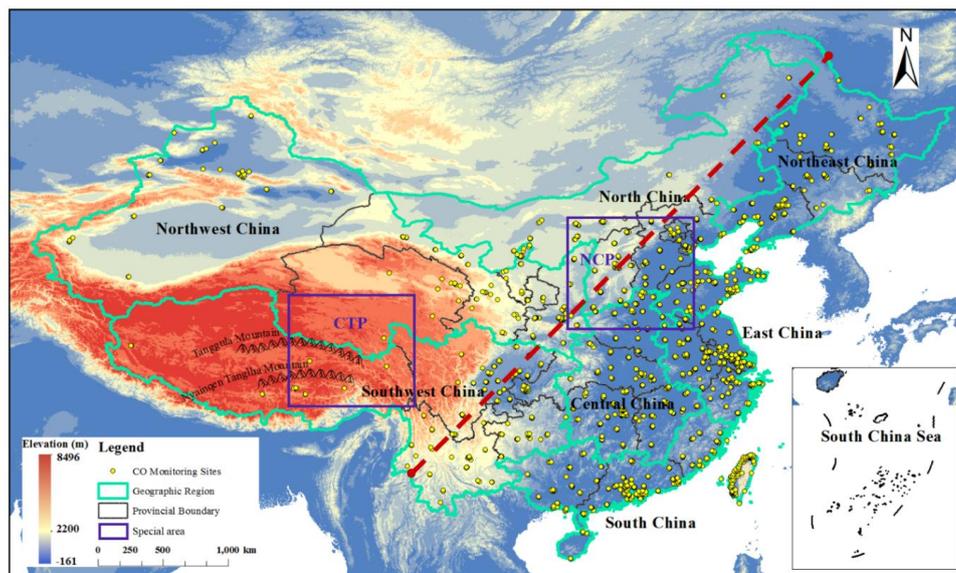
Figure 1: Ground-level CO monitoring network for China in 2013-2016 with 1656 sites in total. The Central Tibetan Plateau (CTP) and the North China Plain (NCP) are labelled on the map. The red dashed line represents Heihe-Tengchong Line.
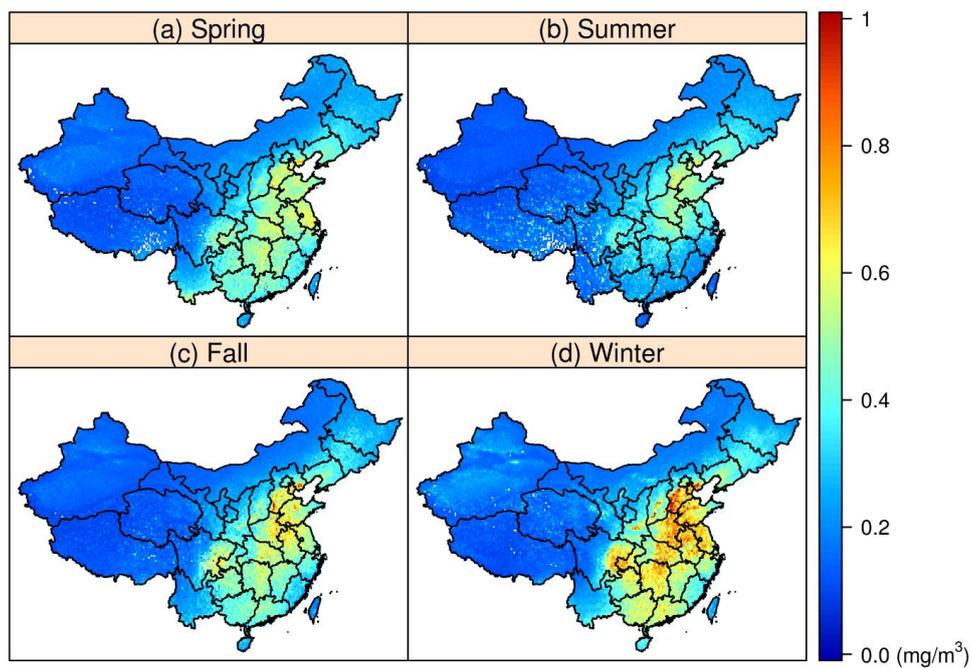
Figure 2: Average ground-level CO concentrations retrieved by the MOPITT for (a) spring, (b) summer, (c) fall, and (d) winter during 2013-2016 (unit conversion: 1.0 ppb = 873.36 mg m^{-3}).

Atmospheric
Chemistry
and Physics
Discussions
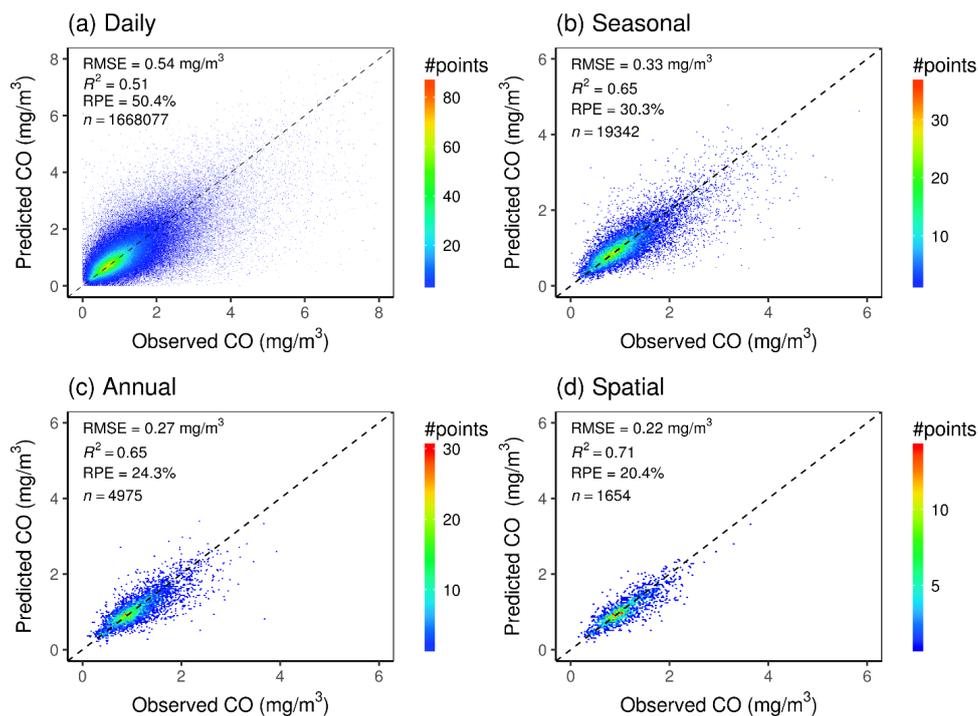
Figure 3: Performance of the refined RF-STK model in predicting (a) daily, (b) seasonal, (c) annual, and (d) spatial (i.e., multiyear average) ground-level CO concentrations across China during 2013-2016. The dashed lines represent the 1:1 relationship.

Atmospheric
Chemistry
and Physics
Discussions
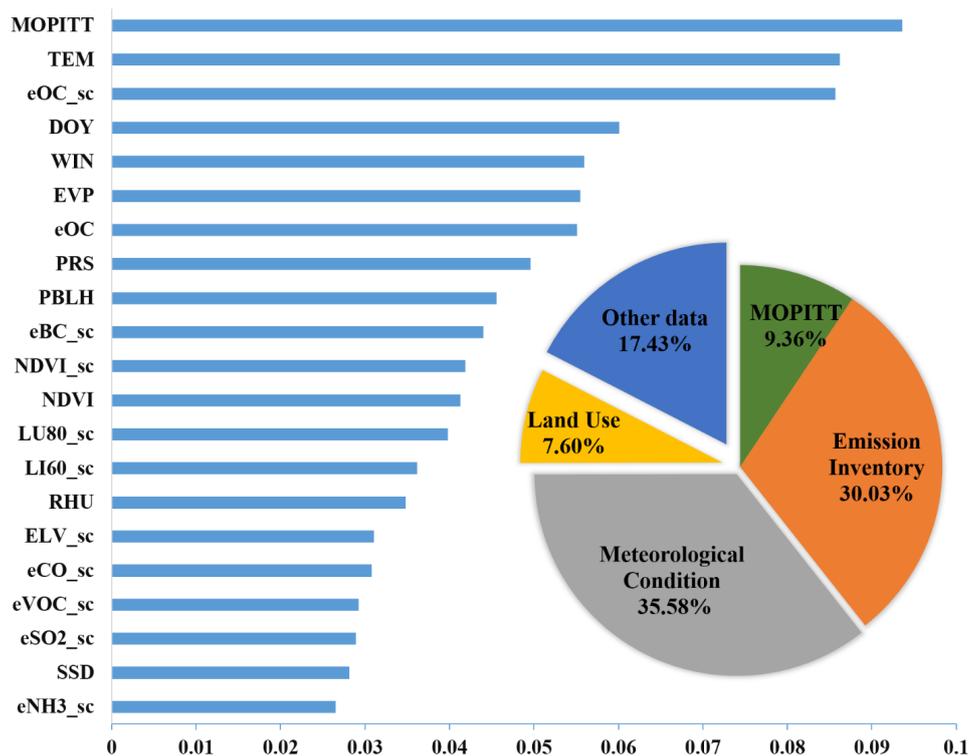
Figure 4: Relative importance of the predictor variables in the refined RF-STK model. Please refer to Table S8 for the detailed descriptions of the variables.
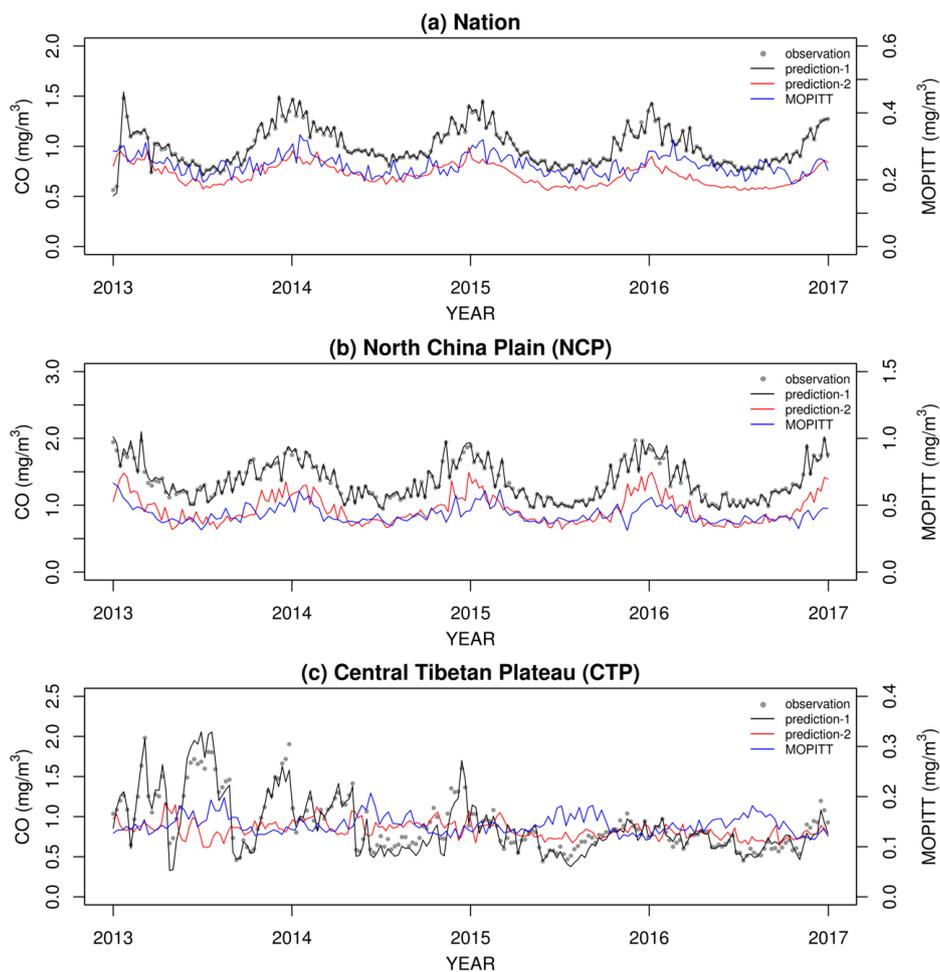
Figure 5: Trend lines of the weekly mean MOPITT-CO versus collocated ground-level CO observations for 2013-2016 for (a) the whole nation, (b) North China Plain and (c) Central Tibetan Plateau. Grey points represent the original in situ values, black solid lines represent the partial predictions which are the same with the sample sizes of situ values, red solid lines represent the whole predictions, and blue solid lines represent converted MOPITT-CO values.
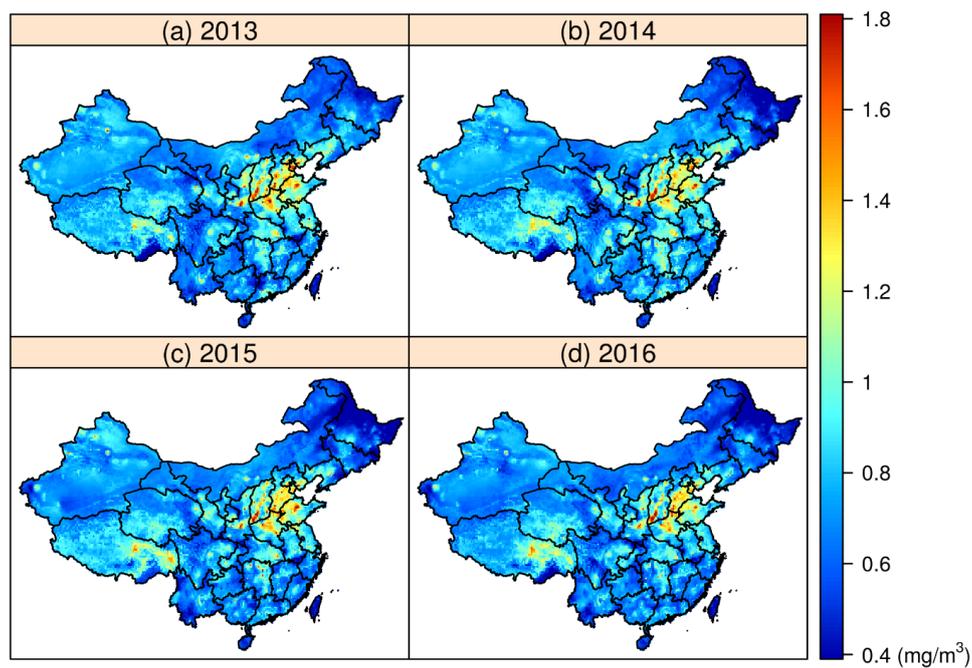
Figure 6: Multiyear average ground-level CO concentrations predicted by the refined RF-STK model for (a) 2013, (b) 2014, (c) 2015, and (d) 2016.
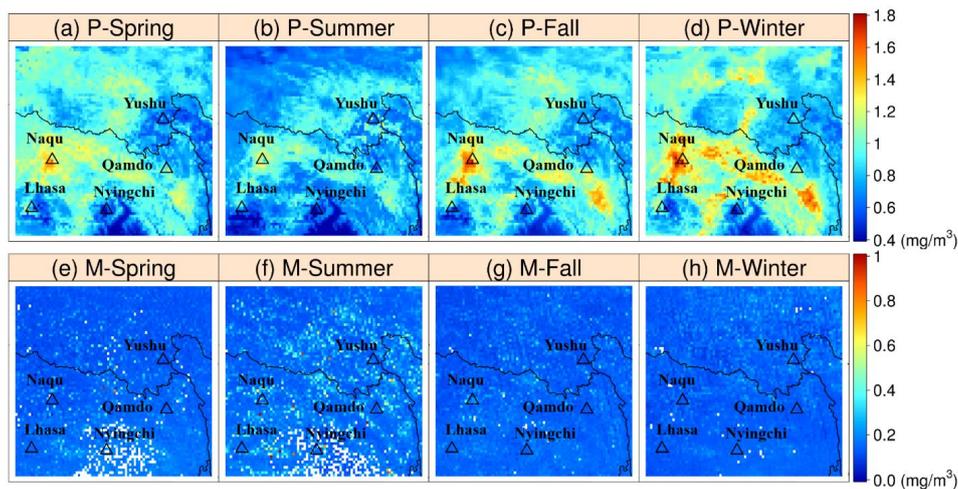
Figure 7: Seasonal predicted ground-level CO concentrations for (a) spring, (b) summer, (c) fall, and (d) winter and seasonal mixing ratio volume (mg m^{-3}) of surface CO retrieved from the MOPITT for (e) spring, (f) summer, (g) fall, and (h) winter during 2013-2016 in the Central Tibetan Plateau (CTP), respectively. Black triangles are the main cities within this area. Please note that there are two different color schemes of this figure. M: MOPITT-CO; and P: refined RF-STK prediction.
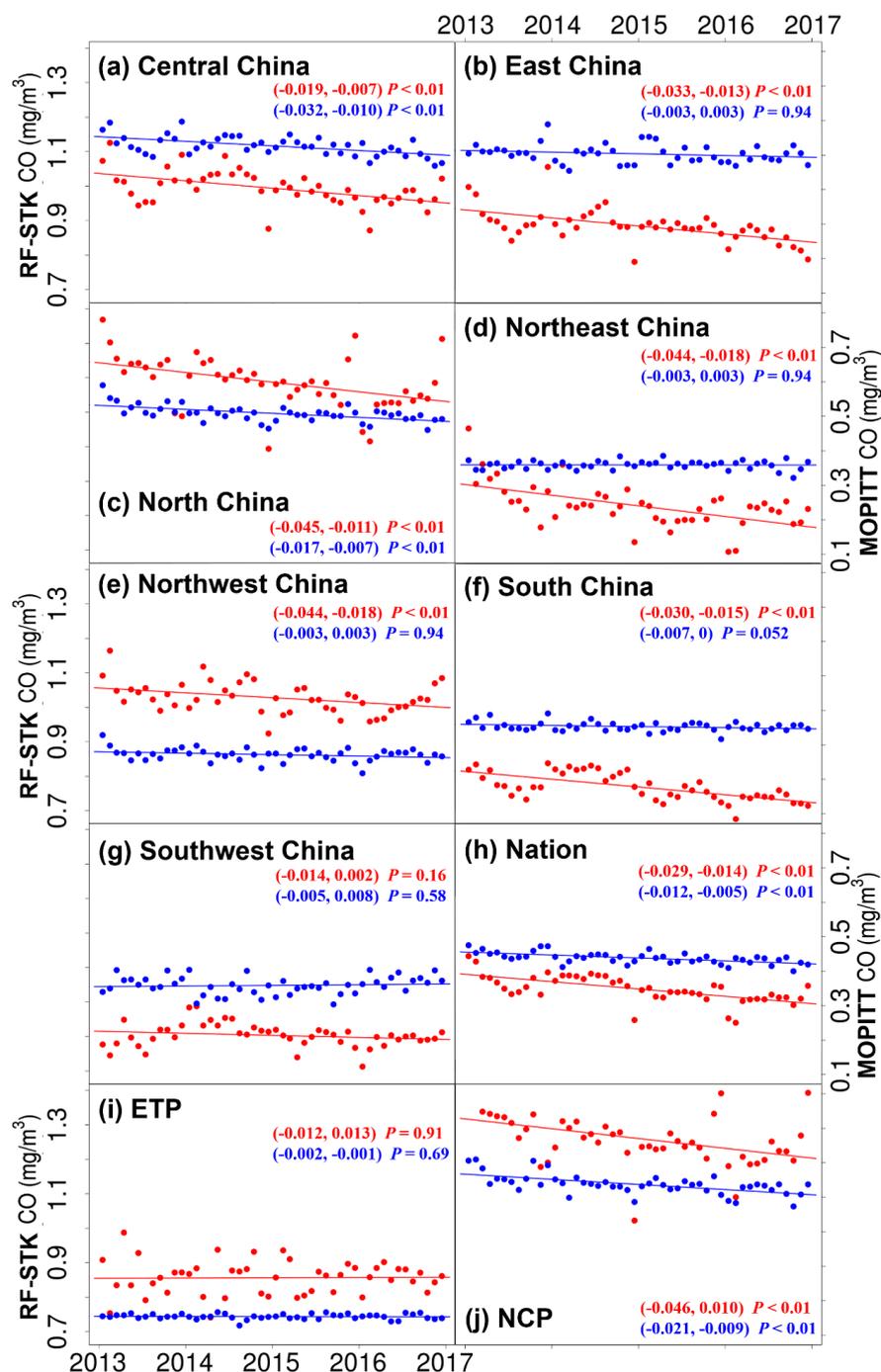
Figure 8: Population-weighted seasonal averages of predicted CO concentrations for the major regions and the whole nation of China during 2013-2016. Blue solid lines represent the refined RF-STK predictions (refer to the left bar), and red solid lines represent converted MOPITT-CO values (refer to the right bar). Since the regions studied are all on the surface and the observations are all points, no observations were added to the figures.