Ozone formation sensitivity study using machine learning coupled with the reactivity of VOC species

Junlei Zhan¹, Yongchun Liu¹*, Wei Ma¹, Xin Zhang², Xuezhong Wang², Fang Bi², Yujie Zhang², Zhenhai Wu², Hong Li²*

1. Aerosol and Haze Laboratory, Advanced Innovation Center for Soft Matter Science and Engineering, Beijing University of Chemical Technology, Beijing 100029, China
2. State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing 100012, China

Correspondence: liuyc@buct.edu.cn; lihong@craes.org.cn
Abstract

The formation of ground-level ozone (O₃) is dependent on both atmospheric chemical processes and meteorological factors. In this study, a random forest (RF) model coupled with the reactivity of volatile organic compound (VOC) species was used to investigate the O₃ formation sensitivity in Beijing, China, from 2014 to 2016, and evaluate the relative importance (RI) of chemical and meteorological factors to O₃ formation. The results showed that the O₃ prediction performance using concentrations of measured/initial VOC species ($R^2 = 0.82/0.81$) was better than that using total VOCs (TVOCs) concentrations ($R^2 = 0.77$). Meanwhile, the RIs of initial VOC species correlated well with their O₃ formation potentials (OFPs), which indicate that the model results can be partially explained by the maximum incremental reactivity (MIR) method. O₃ formation presented a negative response to nitrogen oxides (NOₓ) and relative humidity (RH), and a positive response to temperature (T), solar radiation (SR) and VOCs. The O₃ isopleth calculated by the RF model were generally comparable with those calculated by the box model. O₃ formation shifted from a VOC-limited regime to a transition regime from 2014 to 2016. This study demonstrates that the RF model coupled with the initial concentrations of VOC species could provide an accurate, flexible, and computationally efficient approach for O₃ sensitivity analysis.
1. Introduction

Ground-level ozone (O₃) pollution, which can cause adverse human health effects such as cardiovascular and respiratory diseases, has received increasing attention in recent decades (Cohen et al., 2017). Oxidation of volatile organic compounds (VOCs) will produce peroxyl radicals (RO₂) and hydroperoxyl radicals (HO₂). The RO₂/HO₂ can accelerate the conversion from NO to NO₂, subsequently, formation of O₃ by photolysis of NO₂ in the presence of O₂ (Wang et al., 2017a). The production and loss of RO₂ and HO₂ are highly dependent on the concentration ratio of VOCs and NOₓ in the atmosphere. Hence, atmospheric O₃ concentrations or production rates show a nonlinear relationship with VOCs and NOₓ. Moreover, the O₃-VOC-NOₓ sensitivity is readily influenced by VOC species (Tan et al., 2018), meteorological parameters (Liu et al., 2020a; Liu et al., 2020), and even atmospheric particulate matter (Li et al., 2019), thus, exhibiting high temporal and spatial variability. Therefore, it is urgent to develop an accurate and highly efficient method for timely assessing the sensitivity regime of O₃ production and evaluating the effectiveness of a potential measure on O₃ pollution control. The sensitivity of O₃ formation can usually be analysed using observed indicators, such as ozone production efficiency (OPE, ΔO₃/ΔNOz) (Wang et al., 2010; Lin et al., 2011), HCHO/NOy (Martin et al., 2004), and H₂O₂/NOz (or H₂O₂/HNO₃) (Sillman 1995; Hammer et al., 2002; Wang et al., 2017a), observation-based model (OBM) (Vélez-Pereira et al., 2021) and chemical transport models including community multiscale air quality (CMAQ) (Djalalova et al., 2015) and Weather...
Research and Forecasting with Chemistry (WRF-Chem) model (Wang et al., 2020a). The observed indicators can be utilized to quickly diagnose the sensitivity regime of O$_3$ production. However, the accuracy is sensitive to the precision of tracer measurements. OBMs combine *in-situ* field observations, remote sensing measurements and chemical box models, which are built on widely-used chemistry mechanisms (e.g., MCM, Carbon Bond, RACM or SAPRC), and applied to the observed atmospheric conditions to simulate the *in-situ* O$_3$ production rate (Mo et al., 2018). The sensitivity of O$_3$ production to various O$_3$ precursors, including NO$_x$ and VOCs can be diagnosed based on the empirical kinetic modeling approach (EKMA) or quantitatively assessed with the relative incremental reactivity (RIR). Chemical transport models, which are driven by meteorological dynamics and incorporated with the emissions of pollutants and the complex atmospheric chemical mechanism, provide a powerful tool for simulating various atmospheric processes, including spatial distribution, regional transport vs. local formation, source apportionment and production rates of pollutants and so on (Sayeed et al., 2021). At present, OBMs are widely used to investigate O$_3$ formation sensitivity in China. Previous studies indicated that O$_3$ formation in urban areas of China is located in a VOC-limited or a transition regime and varies with time and location (Ou et al., 2016; Wang et al., 2017a; Zhan et al., 2021). Although both OBMs and chemical transport models can assess the sensitivity of O$_3$ production and predict the O$_3$ pollution level in a scenario of control measures, the calculation accuracy is affected by the uncertainty of input parameters.
Thus, they are mostly applied to sampling cases with a short time span (days or weeks) (Xue et al., 2014; Ou et al., 2016). Compared to traditional methods, machine learning (ML) is able to capture the main factors affecting atmospheric O$_3$ formation in a timely manner with great flexibility (without the constraints of time and space) and high computational efficiency (Wang et al., 2020c; Grange et al., 2021; Yang et al., 2021a). Although attentions should be paid to the robustness of machine learning because it depends on the input dataset (observations or outputs of chemical transport models), previous studies have demonstrated that cross-validation and data-normalization can well reduce the dependence of the model on input data and improve the robustness of the model (Wang et al., 2016; Wang et al., 2017b; Liu et al., 2021; Ma et al., 2021a). Thus, it is a promising alternative to account for the effects of meteorology on air pollutants and has been intensively used in atmospheric studies (Liu et al., 2020a; Hou et al., 2022).

Recently, ML based on convolutional neural network (CNN), random forest (RF) and artificial neural network (ANN) models have been applied in simulating atmospheric O$_3$ and shown good performance in O$_3$ prediction (Ma et al., 2020; Xing et al., 2020). For example, Ma et al. (2021a) simulated O$_3$ concentrations in the Beijing-Tianjin-Hebei (BTH) region from 2010-2017 using an RF model that considered meteorological variables and output variables from chemical transport models, and the correlation coefficient ($R^2$) between the observed and modelled O$_3$ concentrations was greater than 0.8. Liu et al. (2021) also reported a high accuracy (80.4%) for classifying
pollution levels of O$_3$ and fine particulate matter with aerodynamic diameter less than 2.5 µm (PM$_{2.5}$) at 1464 monitoring sites in China using an RF model. Thus, the RF model has shown good performance in terms of prediction accuracy and computational efficiency (Wang et al., 2016; Wang et al., 2017b).

Although ML is widely used to understand air pollution, many ML studies have used total VOCs (TVOCs) to simulate O$_3$ formation and rarely considered the effect of VOC species on O$_3$ formation sensitivity (Feng et al., 2019; Liu et al., 2021; Ma et al., 2021a). Thus, they were unable to identify the chemical reactivity of a single species to O$_3$ formation, which may lead to underestimations or even misunderstandings of the role of VOCs in O$_3$ formation because the same concentration of TVOCs with different compositions may lead to different OPEs. In addition, VOCs react with OH radicals during atmospheric transport, which is the most important sink of VOCs (Carlo et al., 2004; Liu et al., 2020b). Makar et al. (1999) reported that the isoprene emissions were underestimated by up to 40% if the OH oxidation is not considered. Other studies indicated that the initial concentrations of VOCs, which account for the photochemical loss of VOCs during transport, were more representative of pollution levels in the sampling area than the observed VOCs (Yuan et al., 2013; Zhan et al., 2021). However, whether the ML model can identify the connection between the reactivity of VOC species and O$_3$ formation sensitivity has not been clarified.

It should be noted that physical interpretability of the results is an important question when ML models are applied in atmospheric studies (Hou et al., 2022).
However, explanations of ML results (e.g., RI) are somewhat vague because ML is a “black-box” model from the point view of chemical mechanism (Hou et al., 2022; Taoufik et al., 2022). In this study, we used the RF model to evaluate the prediction performance of atmospheric O\textsubscript{3} using the TVOCs, measured VOC species and photochemical initial concentration (PIC) of VOC species, which is calculated based on the photochemical-age approach (Shao et al., 2011). We compared the relative importance (RI) of the precursors (VOC species, NO\textsubscript{x}, PM\textsubscript{2.5}, CO) and the meteorological parameters (temperature, solar radiation, relative humidity, wind speed and direction) on O\textsubscript{3} formation in the summer of Beijing from 2014 to 2016. We also discussed the possibility of connecting the RIs of VOCs with their OFPs and the changes in O\textsubscript{3}-VOC-NO\textsubscript{x} sensitivity based on the RF model from 2014 to 2016. Our study indicates that the RF model combined with initial concentrations of VOC species can simulate O\textsubscript{3} concentrations well and provides a flexible and efficient tool for O\textsubscript{3} modelling in a near real-time way.

2. Methods

2.1 Sampling site and data

The sampling site (40.04°N, 116.42°E) is located at the campus of Chinese Research Academy of Environmental Sciences and was described in our previous work (Zhang et al., 2021). Briefly, the station is located two kilometers from the north 4th ring road and surrounded by a mixed residential and commercial area. The concentrations of VOCs, NO\textsubscript{x}, CO, O\textsubscript{3} and PM\textsubscript{2.5} were measured at 8 m above ground level at this location. Meteorological parameters, including temperature (T), relative humidity (RH),
wind speed and direction (WS&WD), solar radiation (SR), were monitored at 15 m above ground level. VOCs were measured by an online commercial instrument (GC-866, Chromatotec, France), which consisted of two independent analysers for detecting C₂-C₆ and C₆-C₁₂ hydrocarbon components. More details about the observations can be found in the Supplemental Materials (S1). The calculation of initial VOCs and sensitivity tests can be found in the Supplemental Materials (S2).

2.2 Random forest model

The random forest (RF) is a type of ensemble decision tree that can be used for classification and regression (Breiman 2001). During the training process, the model creates a large number of different decision trees with different sample sets at each node, and then averages the results of all decision trees as its final results (Breiman 2001). To avoid over-fitting, we trained the random forest model using cross-validation for the normalized data, which can improve the robustness of the model. Briefly, we randomly divided the normalized data into 12 subsets, then alternately took one subset as testing data along with the rest as training data. By doing this, every data point has an equal chance being trained and tested. The length of the input data from 2014 to 2016 were 1190, 1062 and 872 rows, respectively, in which different types of VOCs, NOx, CO, PM₂.₅ and meteorological parameters (including temperature, relative humidity, solar radiation, wind speed and direction) were used as input variables and O₃ as output variables. The mean values (±standard deviation) of input/output parameters are shown in Table S1. Approximately one-third of the samples are excluded from the sample,
when the decision tree is built and used to calculate the out-of-bag data error. Hence, RF can evaluate the RI of variables via the changes in out-of-bag (OOB) data error (Svetnik et al., 2003),

\[ \text{RI}_i = \frac{\sum (\text{errOOB}_2 - \text{errOOB}_1)}{N} \]  

where \( N \) represents the number of decision trees, and \( \text{errOOB}_1 \) and \( \text{errOOB}_2 \) represent the out-of-bag data error of feature \( i \) before and after randomly permuting the observation, respectively. The RI used to evaluate the importance and sensitivity of feature \( i \) to O\textsubscript{3} formation in this study. More details about workflow of RF model and the hyperparameter tuning can be found in the Text S3. The optimized parameters are shown in Table S2. To verify the stability of the model, we performed a significance test on the model results. The results showed that there was no significant difference among the different tests (\( P > 0.05, R^2 > 0.98 \)).

When plotting the O\textsubscript{3} formation sensitivity curves, we made a virtual matrix of inputs by varying the concentrations of NOx and VOCs from 0.9 to 1.1 times (with a step of 0.01) of their mean values while keeping all other inputs unchanged (i.e., the mean values). Then, the new matrix was used as testing data, while all the measured data were taken as training data. Thus, the testing data should represent the mean sensitivity regime of O\textsubscript{3} in Beijing, while the training data actually covered all the sensitivity regimes of O\textsubscript{3} formation to guarantee a sufficient coverage in the NOx-limited regime for the RF model simulations. The EKMA curves were plotted using the daily maximum 8-h (MDA8) O\textsubscript{3}. More details can be found in the SI.
3. Results and discussion

3.1 Overview of air pollutants and meteorological conditions

Figure 1 shows the time series of air pollutants and meteorological parameters during the observations from 2014 to 2016. In 2014, 2015 and 2016, the wind direction was dominated by northwest winds (Figure S1), with mean wind speeds of $3.1 \pm 2.7$ m s$^{-1}$, $2.3 \pm 2.2$ m s$^{-1}$, and $1.3 \pm 1.2$ m s$^{-1}$, respectively, and the mean daytime temperature were $22.3 \pm 5.8$, $23.9 \pm 5.0$ and $24.0 \pm 4.4$ °C, respectively. The average value of SR decreased from 162.9 to 150.8 W m$^{-2}$ during the observation period. As shown in Figure 1F-G, in 2014, 2015 and 2016, the mean VOC concentrations were $20.3 \pm 10.9$, $15.8 \pm 8.3$ and $12.1 \pm 7.7$ ppbv, respectively, while the mean initial VOC concentrations were $28.1 \pm 25.7$, $27.2 \pm 32.6$ and $16.4 \pm 16.1$ ppbv, respectively. The calculation of initial VOCs and sensitivity tests can be found in the Supplemental Materials (S2). Both the measured VOCs and initial VOCs showed a decline along with a decrease in PM$_{2.5}$ concentration from $67.2 \pm 53.5$ to $61.1 \pm 48.6$ µg m$^{-3}$ due to the Air Pollution Prevention and Control Action Plan in China (Zhao et al., 2021). However, O$_3$ concentrations showed a slight downward trend from $44.3 \pm 32.4$ to $42.7 \pm 27.9$ ppbv from 2014 to 2015 and then reach to $44.0 \pm 29.6$ ppbv in 2016. A slight upward trend was observed for NO$_x$ concentrations (Figure S2). As shown in Figure 1F-G, the concentrations of four types (alkanes, alkenes, alkynes, and aromatics) of VOCs showed significant differences from 2014 to 2016 due to the variations in emission sources (Zhang et al., 2021). In addition to VOC species, the variations in other parameters, such as
meteorological conditions and PM$_{2.5}$, should have a complex influence on O$_3$-VOC-NO$_x$ sensitivity (Li et al., 2019; Ma et al., 2021b).

**Figure 1.** Time series of air pollutants and meteorological parameters during observations in Beijing. (In A, the red arrows represent the O$_3$ concentration exceed 74.6 ppbv according to the national ambient air quality standard.)

3.2 Prediction performance of the model.

To build a robust model, we evaluated the prediction performance of the RF model for the ambient O$_3$ simulation. Figure 2 shows the O$_3$ prediction performance in 2015 when chemical species (including VOCs, NO$_x$, PM$_{2.5}$, CO) and meteorological factors (i.e., WS, WD, SR, T and RH) were used as inputs in the RF model. The prediction performance of RF model for 2014 and 2016 is shown in Figures S3 and S4 respectively.
The details of the modelling and input parameters are shown in Table S2. Figure 2A-C shows the time series of the measured and modelled O$_3$ concentrations, which were simulated using the TVOCs, measured VOC species and initial VOC species as part input variables along with the same set of other parameters. The correlation coefficients ($R^2$) of the training data were 0.77, 0.82 and 0.81 for the TVOCs, measured VOC species and initial VOC species, respectively. The corresponding root mean squared errors (RMSEs) for the predicted O$_3$ concentrations were 17.4, 12.6 and 13.9. Figure 2D-F shows the prediction performance of the testing dataset under these three circumstances. When the TVOCs were split into measured or initial VOC species, the $R^2$ increased obviously as the number of data features increased. Therefore, the VOC composition has a significant influence on O$_3$ prediction using the RF model. In previous studies using TVOCs, the influence of VOC composition was neglected (Liu et al., 2021; Ma et al., 2021a). Our results indicate that the RF model can accurately predict O$_3$ concentrations when the concentrations of measured/initial VOC species are considered.
Figure 2. Comparison of the predicted and measured O₃ concentrations in Beijing in the summer of 2015. (A and D: TVOC concentrations; B and E: measured concentrations of VOC species; C and F: initial concentrations of VOC species)

It should be pointed out that if the training dataset does not have sufficient coverage in the NOx-limited regime, then the trained algorithm essentially attempts to extrapolate in that regime, which is prone to overtraining. To avoid such overtraining, a 12-fold cross-validation by randomly dividing the observation data in each day into 12 subsets and alternately taking one subset as testing data and the rest as training data ensures that each data point has an equal chance of being trained and tested. The curves of the predicted O₃ concentrations in Figure 2 were spliced using the testing datasets in all runs. Thus, our results actually covered all the sensitivity regimes of O₃ formation.

This means that the model is robust

3.3 Relative importance of major factors
Figure 3A shows the RIs of different ambient factors, including chemical and meteorological variables on O$_3$ formation. The difference in the RIs is also compared using the TVOCs and the VOC species as inputs. Chemical factors (including VOC species, NO$_x$, PM$_{2.5}$ and CO) accounted for 79.1% of the contribution to O$_3$ production in the summer of 2016. Meanwhile, VOC species accounted for approximately 63.4% of O$_3$ production while the RIs using TVOC concentrations accounted for only 2.1%. Ma et al. (2021b) analysed the contribution of meteorological conditions and chemical factors to O$_3$ formation on the North China Plain (NCP) using the CMAQ model in combination with process analysis and found that chemical factors dominate O$_3$ formation in summer. Using probability theory, Ueno et al. (2019) also found that VOCs/NO$_x$ dominate O$_3$ production compared to meteorological variables. Thus, our results are similar to those of previous studies based on chemical models (Ueno et al., 2019), which demonstrates that the RF model can reflect the contribution of VOC species to O$_3$ production even if the observed VOC species are used.
Figure 3. Percentage of RI for O\textsubscript{3} precursors and meteorological parameters (A) and the top 10 factors with high values of RI in 2014-2016 (B-D: using initial concentrations of VOC species).

Here, we compared the RIs of VOCs calculated using the initial VOC species and the observed VOC species with the O\textsubscript{3} formation potentials (OFPs). The OFPs were calculated by the maximum incremental reactivity (MIR) method (Carter 2010). As shown in Figure S5, the RIs showed good correlations with the OFP. Interestingly, the initial concentrations of VOC species improved the correlation coefficients between the RIs and OFPs. Furthermore, we calculated the RIs and OFPs of different species using the observed data during the campaign study in Daxing District in the summer of 2019 (Zhan et al., 2021), and a stronger correlation was observed between the RIs of the initial VOC species and the OFPs (Figure S6). These results indicate that the RIs of the initial VOCs species in the ML model should partially reflect the chemical reactivity of
VOCs to produce O₃ in the atmosphere. Although the RIs calculated using the initial VOC species slightly changed compared to those calculated using the observed VOCs (Table S3), VOCs still dominated O₃ formation (Figure 3A). For example, the initial VOCs dominated O₃ production in 2014, 2015, and 2016, with RI values of 64.0, 59.0 and 63.3% respectively. Li et al. (2020a) used a multiple linear regression (MLR) model to study the contribution of anthropogenic and meteorological factors to O₃ formation in China from 2013-2019 and found that meteorological factors accounted for 36.8% and anthropogenic factors accounted for 63.2%, which is similar to our results. Figure 3B-D shows the top 10 factors having a strongly influence on O₃ production. Interestingly, NOₓ and RH showed negative responses to O₃ formation, while other variables, including T, SR, CO and all of the VOCs, showed positive responses. Thus, a decrease in NOₓ or RH will lead to an increase in O₃ concentration while a decrease in T, SR, CO and VOCs will lead to a decrease in O₃ concentration. Although O₃ formation is highly related to the photolysis of NO₂, a previous study demonstrated that it is VOC-limited in summer in Beijing (Zhan et al., 2021). This finding is consistent with the observed negative response of O₃ to NOₓ in this work. High RH usually coincides with low surface O₃ concentrations in field observations, which can be ascribed to the inhibition of O₃ formation by the transfer of NO₂/ONO₂-containing products into the particle phase and the promotion of dry deposition of O₃ on the surface (Kavassalis et al., 2017; Yu 2019). In addition, it has been shown that RH is negatively related to the
rate constant of HONO formation (Hu et al., 2011). Thus, RH might also affect the O₃ formation by influencing atmospheric OH radicals from photolysis of HONO. It should be noted that the negative response of ozone to RH might also be resulted from the dependence of RH on other parameters/conditions, such as SR. However, RH and SR showed a bad correlation ($r < 0.1$). We further tested the dependence of the RI on RH and SR with or without the counterpart as input. The stable RI values (Table S4) mean that RH and SR are independent from each other. These previous works can well explain the observed negative response of O₃ to RH in Figure 3B-D. Previous studies have observed a positive correlation between the O₃ concentration and T or SR (Steiner et al., 2010; Paraschiv et al., 2020; Li et al., 2021). Temperature can directly affect the chemical reaction rate of O₃ formation (Fu et al., 2015), and SR can promote the photolysis of NO₂ (Hu et al., 2017; Wang et al., 2020b), thus accelerating O₃ formation. As mentioned above, O₃ formation is VOC-limited in Beijing; thus, a positive response of O₃ concentration to VOCs is observed in Figure 3B. Interestingly, the RIs of isoprene showed an increasing trend from 2014 to 2016 because of the obvious reduction in anthropogenic VOCs (Figure S7) (Zhang et al., 2021). In the context of global warming, studies should focus on the factors that affect O₃ formation, including biogenic emissions, T and SR. Thus, additional efforts will be required to reduce anthropogenic pollutants in the future.

### 3.4 Ozone formation sensitivity

To further analyse the sensitivity of O₃ to VOCs and NOₓ from 2014 to 2016, we
plotted sensitivity curves for O$_3$ generation using the RF model, and the results are shown in Figure 4A-C. Moreover, EKMA curves in 2015 were also obtained using the OBM (Figure 4D). As shown in Figure 4A-C, O$_3$ formation was sensitive to VOCs in the summer of Beijing during our observations, which is consistent with previous studies that used box models (Li et al., 2020b) and chemical transport models (Shao et al., 2021). This result is also consistent with the RIs of VOCs or NO$_x$ to O$_3$ formation (Figure 3B-D). Interestingly, the O$_3$ formation sensitivity to VOCs decreases or gradually shifts from the observed point to the transition regime from 2014 to 2016 (Figure 4A-C), which is similar to that reported by Zhang et al. (2021). These phenomena can be ascribed to the increased relative importance of meteorological factors, such as T, SR, and RH, for O$_3$ formation and the variation in anthropogenic VOC emissions (Steiner et al., 2010; Ma et al., 2021b).
Figure 4. Ozone formation sensitivity curves from 2014-2016. (A, B, C: calculated by the RF model for 2014, 2015, and 2016, respectively. D: calculated by the OBM for 2015.)

We compared the relative error of simulated MDA8 $O_3$ calculated using the RF and OBM model in 2015, as shown in Figure S8. The mean relative error of simulated MDA8 $O_3$ between RF model and Box model was 15.6%. Hence, a combination of the RF model and initial VOC species can accurately depict the sensitivity regime of $O_3$ formation, while the calculated RIs correlate well with the OFPs.

4. Conclusions

In summary, this work investigated $O_3$ formation sensitivity in the summer from 2014-2016 in Beijing using the RF model coupled with the reactivity of VOC species.
The results show that the prediction performance of O$_3$ by the RF model was significantly improved when measured/initial VOC species were considered compared to TVOCs. Furthermore, after the photochemical loss of VOC species during transport was corrected, the RIs of the VOC species were well correlated with the OFPs of VOC species calculated using the MIR method, thus indicating that the RIs in the ML model reflect the chemical reactivity of VOCs. Meanwhile, both NO$_x$ and highly reactive species (such as isoprene, propene, benzene) played an important role in O$_3$ formation. An increased contribution of temperature to O$_3$ production was observed, which implied the importance of temperature to O$_3$ pollution in the context of global warming conditions. Both the RF model and the box model results showed that O$_3$ formation was sensitive to VOCs in Beijing, although the sensitivity regime shifted from VOC-limited regime to a transition regime from 2014 to 2016. Due to the high computational efficiency of ML, the O$_3$ formation sensitivity plotted by the RF model coupled with the reactivity of VOC species can provide an accurate, flexible and efficient approach for analysing O$_3$ sensitivity in a near real-time way.

**Code and data availability**

The datasets of VOCs and meteorology are available and will be provided by the corresponding authors Yongchun Liu (liuyc@buct.edu.cn) and Hong Li (lihong@craes.org.cn) upon request. The code can be seen in GitHub (https://github.com/z-12/amt-2021-367.git). The solar radiation data are publicly
Supplement

Supplementary information is available for this paper.

Author contributions
Junlei Zhan designed the idea and wrote this manuscript; Yongchun Liu and Hong Li provided useful advice and revised the manuscript; Wei Ma performed box model simulations; and Xin Zhang, Xuezhong Wang, Fang Bi, Yujie Zhang and Zhenhai Wu conducted the campaign and compiled the data. All authors contributed to the discussion of the results and writing of the manuscript.

Competing interest
The authors declare that they have no conflict of interest.

Acknowledgments
This research was financially supported by the Ministry of Science and Technology of the People’s Republic of China (2019YFC0214701), the National Natural Science Foundation of China (41877306 and 92044301) and the programs from Beijing Municipal Science & Technology Commission (No. Z181100005418015). We thank Yizhen Chen for providing the meteorological parameter data for campaign studies.
References


Carter, W. Updated maximum incremental reactivity scale and hydrocarbon bin reactivities for regulatory applications. California Air Resources Board Contract, 1, 07-339, 2010.


Kavassalis, S.C., Murphy, J.G. Understanding ozone-meteorology correlations: A role for dry


Taoufik, N., Boumya, W., Achak, M., Chennouk, H., Dewil, R., Barka, N. The state of art on the prediction of efficiency and modeling of the processes of pollutants removal based on...


