Assessment of dicarbonyl contributions to secondary organic aerosols over China using RAMS-CMAQ

Jialin Li¹, Meigen Zhang¹,²,³, Guiqian Tang¹, Yele Sun¹,²,³, Fangkun Wu¹, and Yongfu Xu¹,³

¹State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China
²Center for Excellence in Regional Atmospheric Environment, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China
³University of Chinese Academy of Sciences, Beijing 100049, China

Correspondence: Meigen Zhang (mgzhang@mail.iap.ac.cn)

Received: 17 October 2018 – Discussion started: 12 November 2018
Revised: 20 March 2019 – Accepted: 7 April 2019 – Published: 16 May 2019

Abstract. The concentration of secondary organic aerosol (SOA) is underestimated in current model studies. Recent research suggests that the reactive uptake of dicarbonyls contributes to the production of SOA, although few models have included this pathway. Glyoxal, an important representative component of dicarbonyls in models, is significantly underestimated. We therefore incorporated the reactive uptake of dicarbonyls into the regional air quality modeling system RAMS-CMAQ (the Regional Atmospheric Modeling System-Community Multiscale Air Quality) to evaluate the contribution of dicarbonyls to SOA, and we then assess the impact of the underestimation of glyoxal on the production of SOA in China during two time periods: 3 June to 11 July 2014 (episode 1) and 14 October to 14 November 2014 (episode 2). When the reactive uptake process was added, the modeled mean concentration of SOA in episode 1 increased by 3.65 µg m⁻³, which explained 34.8 % of the unaccounted-for source of SOA. Meanwhile the increase in the concentration of SOA in episode 2 was 1.82 µg m⁻³ as a result of the lower liquid water content and the lower amount of dicarbonyls produced from biogenic precursors in the fall. On this basis, when the glyoxal simulation was improved, the modeled mean dicarbonyl-derived SOA (AAQ) increased by more than a factor of 2 in both episodes relative to case 1. AAQ in episode 1 contributed, on average, 60.6 % of the total concentration of SOA and the increase in this contribution represented 69.1 % of the unaccounted-for concentration of SOA, whereas the mean AAQ in episode 2 accounted for 64.5 % of total concentration of SOA. Based on the results, the mean AAQ over China was generally higher in the east than in the west during the two episodes. The highest value (10–15 µg m⁻³) of episode 1 appeared in the areas around the lower reaches of the Yellow River, whereas the highest value of 5–10 µg m⁻³ in episode 2 was concentrated over regions from south of the lower reaches of the Yellow River to the south of Guangzhou Province as well as the Sichuan Basin. The contribution of AAQ to the concentration of SOA in episode 1 varied from 10 % to 90 % throughout China, with the highest contributions (70 %–90 %) in the coastal regions and offshore along the East China Sea to the South China Sea and in the southwestern regions. The fraction of AAQ to SOA in episode 2 was in the range of 10 %–80 % over China, with the fraction up to 80 % in a small portion of northeastern China.

1 Introduction

The fine particle fraction of aerosols (PM₂.₅, i.e., particulate matter with an aerodynamic diameter ≤ 2.5 µm) not only absorbs and scatters sunlight in the atmosphere, resulting in a reduction in visibility and effects on the global climate (IPCC, 2007), but also has adverse effects on human health (Harrison and Yin, 2000; Pöschl, 2005). Organic aerosol (OA) is a major component of fine particulate matter globally (Murphy et al., 2006; Zhang et al., 2007), typically making up 20 %–90 % of the fine particle fraction (Roberts et al., 2001; Kanakidou et al., 2005; Zhang et al., 2007), suggesting
that it has a significant effect on the characteristics and properties of fine particulate matter. Organic aerosols are formed from both primary (direct emissions) and secondary sources. Secondary organic aerosols (SOAs) are formed from precursors of volatile organic compounds (VOCs), which can be either natural or anthropogenic. However, the underestimation of organic aerosol has become a major issue in almost all current atmospheric models due to the incomplete representation of SOA (Heald et al., 2005; Morris et al., 2005; 2006; Goldstein and Galbally, 2007; Yu et al., 2008; Fu et al., 2009; Farina et al., 2010; Jiang et al., 2012; Zhang and Ying, 2012; Jo et al., 2013; Lin et al., 2016).

It has been reported that the concentration of SOA in the models is underestimated by 1 to 2 orders of magnitude (de Gouw et al., 2005; Volkamer et al., 2006). These results have motivated researchers to investigate why these models are predicting SOA concentrations so poorly. Traditionally, improvements in models have mainly concentrated on the gas-phase and derived heterogeneous formation processes, such as the formation of SOA from aromatic compounds under low- and high-NO$_x$ conditions (Ng et al., 2007), the production of SOA from the oxidation of isoprene and its subsequent growth through acid catalysis (Kroll et al., 2005, 2006; Surratt et al., 2007), the use of the volatile bias set method to depict the formation of SOA instead of the two-product model (Lane et al., 2008; Murphy and Pandis, 2009; Han et al., 2016; Lin et al., 2016), and the formation of SOA through the uptake of isoprene epoxides on the surface of aerosols (Ying et al., 2015; Hu et al., 2017). The gap has been closed to some degree through the SOA modeling efforts, but there are still large uncertainties (Tsigaridis et al., 2014).

Many studies have suggested that glyoxal and methylglyoxal can form SOAs through chemical reactions in cloud or fog water (e.g., Warneck, 2003; Ervens et al., 2004; Lim and Ziemann, 2005; Carlton et al., 2006; Loeffler et al., 2006), or by reactive uptake on the surface of cloud droplets and aqueous aerosols (e.g., Liggio et al., 2005; Corrigan et al., 2008; Galloway et al., 2009; Ervens and Volkamer, 2010; Lin et al., 2010), which is probably a significant source of SOA (Ervens et al., 2014; Curry et al., 2018). A few studies (e.g., Carlton et al., 2008, 2010; Fu et al., 2009; Lin et al., 2012, 2014; Li et al., 2013; Woo and McNeill, 2015) have incorporated aqueous SOA formation pathways into atmospheric models. Several of these studies have shown that chemical reactions only in cloud or fog water make a negligible contribution to near-surface SOA relative to reactive uptake on the surface of cloud droplets and aqueous aerosols, and that the aqueous SOA formation cannot completely explain the gaps between the observations and simulations. There are still considerable uncertainties in our knowledge of the formation of SOA.

A series of studies (Fu et al., 2008; Myriokefalitakis et al., 2008; Liu et al., 2012; Li et al., 2018) has shown that there is a substantial underestimation in the modeled vertical column densities (VCDs) of glyoxal, but few studies have considered the impact of this underestimation on simulations of SOA concentrations. Xu et al. (2017) and Ervens et al. (2011) showed that the aqueous SOA formation depends on the liquid water content (LWC), which varies between seasons. Previous studies, such as those of Fu et al. (2009) and Li et al. (2013), have only considered the contribution from SOA derived from the reactive uptake of dicarbonyls (pathway $M$) in the summer period or over evergreen areas. In this study, we not only incorporated pathway $M$ into the RAMS-CMAQ (the Regional Atmospheric Modeling System-Community Multiscale Air Quality) modeling system to evaluate the corresponding contribution of dicarbonyls to SOA, but also improved the simulation of glyoxal concentrations by investigating the reasons for its underestimation and assessing its impacts on the concentration of SOA during the two episodes in the summer and fall of China.

2 Model and data

2.1 Base model description

J. Li et al. (2017) used the RAMS-CMAQ modeling system to investigate the effects of underestimated aromatic VOC emissions and yields of SOAs from different gas precursors on the concentration of SOAs. Similarly, we used Version 4.7.1 of the CMAQ, which is coupled with the gas-phase photochemical mechanism SAPRC99 (1999 Statewide Air Pollutant Research Center; Carter, 2000) and Version 5 of the aerosol module (AERO5; Byun and Schere, 2006; Foley et al., 2010). There are three major formation pathways for SOAs in this version. Based on the two-product approach, the first pathway is the equilibrium partition of semivolatile products formed from the oxidation of seven VOC precursors: long-chain alkanes, benzene (BNZ), high-yield aromatics (mainly toluene, ARO1), low-yield aromatics (mainly xylene, ARO2), isoprene (ISOP), monoterpane (TRP) and sesquiterpenes (SESQ). The second pathway is the oligomerization of semivolatile SOAs formed through the first pathway (Kalberer et al., 2004), namely the aging process. The third pathway is the formation of SOA via the in-cloud oxidation of glyoxal (GLY) and methylglyoxal (MGLY) (Carlton et al., 2008), both of which represent dicarbonyls in the model. The details of these formation pathways are given in Carlton et al. (2010).

The meteorological fields used to drive CMAQ are obtained from RAMS, which has been described in detail by Cotton et al. (2003). The National Centers for Environmental Prediction reanalysis datasets are served as the initial and lateral boundary meteorological conditions input into RAMS. The boundary conditions used for the RAMS computations include the weekly average sea surface temperature and the monthly measured snow cover. The final modeled results are output through the four-dimensional data assimilation mode using nudging analysis.
The emission sources are derived from several different inventories. Anthropogenic emissions (M. Li et al., 2017) – including SO$_2$, NO$_x$, CO, black carbon, non-methane VOCs, organic carbon, NH$_3$ and other particulate matter – are obtained from the monthly emissions inventory of 2012. There are five emission sectors (the power, industrial, residential, transportation and agricultural sectors) in the inventory with a spatial resolution of 0.25° × 0.25° (see http://www.meicmodel.org, last access: 16 April 2019). The Model of Emissions of Gases and Aerosols from Nature (MEGAN) (Guenther et al., 2012) provides the biogenic emissions. The emissions from open biomass burning are derived from the Global Fire Emissions Database, Version 4 (Randerson et al., 2015), whereas the information of monthly lightning NO$_x$ is obtained from the Global Emissions Inventory Activity with a spatial resolution of 1° × 1° (Benkovitz et al., 1996). The emissions of NO$_x$ from the soil are derived from the Regional Emission inventory in ASia, Version 2.1, with a spatial resolution of 0.25° × 0.25° (Kurokawa et al., 2013). The online dust and sea salt emissions are calculated using the empirical model developed by Han et al. (2004) and the scheme of Gong (2003), respectively. The calculation of the model boundary conditions is supported by the Model of Ozone and Related Tracers, Version 4 (MOZART-4) (Emmons et al., 2010).

The model domain is divided into $105 \times 86$ grid cells with the center located at 35° N, 110° E (Fig. 1). The map projection is rotated polar-stereographic with a horizontal resolution of 64 km. The vertical simulation region is unequally spaced from the ground surface to ~23 km. There are 25 layers in the $\sigma$ coordinate system in RAMS, with nine layers located in the lowest 2 km to resolve the planetary boundary layer. The vertical dimension of CMAQ is divided into 15 layers, with the lowest seven layers the same as those in RAMS. We focused on the area outlined by the blue box in Fig. 1.

2.2 Adding the reactive uptake process

There is a standard reaction probability formulation for reactive uptake of gases by aerosols and clouds in Jacob (2000). In this formulation, the first-order rate constant $k$ for the chemical loss of a gas-phase species to the aerosols or cloud droplets through molecular diffusion and free collision is given by

$$k = \left( \frac{a}{D_g} + \frac{4}{v\gamma} \right)^{-1} A,$$

where $D_g$ is the gas-phase molecular diffusion coefficient, $a$ is the radius of the aerosol particle or cloud droplet, $v$ is the mean molecular speed, $\gamma$ is the reaction uptake coefficient when a collision occurs between a gas-phase molecule and the aqueous surface, and $A$ is the aqueous particle surface area per unit volume of air.

Figure 1. Geographical locations of the measurement stations in the model domain. 1: Beijing; 2: Tianjin; 3: Datong; 4: Jinan; 5: Qingshui; 6: Hohhot; 7: Fuzhou; 8: Guangzhou; 9: Urumqi; 10: Nanchang; 11: Guilin; 12: Guanyao; 13: Xinglong; 14: Yucheng. The different colors denote which variables are compared at each station. Purple: meteorological parameters; green: concentrations of aromatics; red: both meteorological parameters and aromatic compound concentrations.

We implement the reactive uptake of dicarbonyls by cloud droplets following the standard Eq. (1). In a similar manner to Fu et al. (2008) and Li et al. (2013), the cloud droplet surface area is calculated from the LWC in the cloudy fraction of the model grid by assuming an effective droplet radius of 6 µm for continental clouds and 10 µm for maritime clouds. $D_g$ is calculated according to Perry and Green (1999) following Dentener and Crutzen (1993). Similar to previous studies (e.g., Lin et al., 2014; Ying et al., 2015; Hu et al., 2017), the uptake of gas-phase dicarbonyls on aqueous aerosols is simply parameterized using the collision limitation $(v\gamma A/4)$ of Eq. (1). The value of $\gamma = 2.9 \times 10^{-3}$ for glyoxal from Liggio et al. (2005) is used in both cloud and aerosol processing. While the uptake coefficient adopted for methylglyoxal is from De Haan et al. (2018), the average value of $\gamma = 5.7 \times 10^{-3}$ measured on cloud-processed glycine aerosols is used to account for the uptake by cloud droplets; the average value of $\gamma = 3.0 \times 10^{-3}$ used in aerosol processing is derived from the uptake coefficients measured on deliquesced ammonium sulfate aerosols ($3.7 \times 10^{-3}$) and on deliquesced glycine aerosols ($2.3 \times 10^{-3}$, median value).

2.3 Observational data

To assess the contribution of dicarbonyl species to the concentration of SOA, we used the observed hourly concentration of SOA in Beijing during the summer and fall of 2014 measured by Xu et al. (2017) with an Aerodyne high-resolution time-of-flight aerosol mass spectrometer. Samples were taken from 3 June to 11 July 2014 (episode 1) and 14 October to 14 November 2014 (episode 2) at the Institute of Atmospheric Physics in China. More detailed information about the data has been reported by Xu et al. (2017).
The simulation periods are from 22 May to 11 July and from 1 October to 14 November, with the first 12 d as the spin-up time. To evaluate the reasonability in simulating the formation processes of aqueous SOAs and analyze the relative causes, the corresponding cloud water path (CWP) and cloud fraction data measured by the MODerate Resolution Imaging Spectroradiometer (MODIS) were obtained from the website https://ladsweb.modaps.eosdis.nasa.gov/search/order/4/MYDATML2--61 (last access: 16 April 2019), and the observed precipitation data were downloaded from the website http://chrsdata.eng.uci.edu/ (last access: 16 April 2019).

Both Liu et al. (2012) and Li et al. (2018) have suggested that the underestimation of glyoxal concentrations in simulations is related to the underpredicted emission of aromatic compounds. Therefore, the biases in the emission of aromatic compounds need to be evaluated through a comparison of the observed and simulated concentrations of aromatic compounds. The observed data were collected at 14:00 local standard time every Thursday by gas chromatography and mass spectrometry in Beijing, Xinglong and Yucheng (Fig. 1). Sun et al. (2016) and Wu et al. (2016) have presented the detailed information.

To evaluate the performance of our model, we also compared the simulated PM$_{2.5}$ concentrations and several meteorological factors with the results measured over 12 stations (Fig. 1) during the two episodes. The observed meteorological data were derived from the Meteorological Information Comprehensive Analysis and Process System (MICAPS) dataset, whereas the Chinese National Environmental Monitoring Center provided the measured concentrations of PM$_{2.5}$.

3 Results and discussions

3.1 Base model evaluation

Table 1 shows statistical results of the meteorological parameters (i.e., temperature, relative humidity, wind speed and wind direction) and PM$_{2.5}$ concentrations in the two analyzed episodes, where $N$ is the total number of samples; IOA is the index of agreement, which can synthetically reflect the combination of the modeled value and variable tendency being good or bad (Willmott, 1981); $C_{\text{mod}}$ and $C_{\text{obs}}$ are the average values of modeled and observed results, respectively; MB and FB are the mean and fractional bias, respectively; RMSE and FE are the root-mean-square and fractional error, respectively; and $R$ is the correlation coefficient between the observed and simulated results. The calculations of these statistical parameters can be found in Juda-Rezler et al. (2012). $P_{22.5^\circ}$ and $P_{45^\circ}$ represent the proportions of compared results that the absolute biases between the simulated and measured wind directions are within 22.5 and 45°, respectively (J. Li et al., 2017).

There are inevitably some biases in the simulated meteorological parameters relative to the observations due to the limited model resolution and system errors. Nevertheless, the model reproduces the magnitude and variation trend of the temperature and relative humidity fairly well (Figs. S1 and S2 in the Supplement), with IOAs for temperature of 0.91 and 0.95 in the two episodes and IOAs of 0.91 and 0.88 for the relative humidity, comparable with the results of Li et al. (2012) and Wang et al. (2014). The mean biases of temperature ($-0.83$, $-0.68^\circ\text{C}$) and relative humidity (1.53 %, $-0.05\%$) are small in the two episodes. The RMSEs of temperature and relative humidity are comparable with the results of Gao et al. (2016) and Zhang et al. (2012), implying a reasonable performance of the simulation. Although the capture of the variable tendency of wind speed (Fig. S3) by the model is rather poor, the small values for the mean bias ($-0.40$, $-0.22\text{ m s}^{-1}$) and the high IOAs (0.61, 0.69) indicate a better performance than that reported by Feng et al. (2016) and Wang et al. (2014) in simulating the wind speed. The RMSE of wind speed shows that the simulated results achieve the criteria for a good performance ([RMSE] $\leq 2\text{ m s}^{-1}$) given by Emery et al. (2001). For the wind direction (Fig. S4), $P_{22.5^\circ}$ and $P_{45^\circ}$ are 32.12 % and 54.52 % in episode 1 and 34.01 % and 57.41 % in episode 2, indicating the reasonable simulation of wind direction. Therefore the model provides a reasonable meteorological field for the subsequent simulations.

The modeled PM$_{2.5}$ concentrations are generally higher than the observed values in the two episodes over the 12 stations (Fig. S5). Except for the limited resolution of the model, the uncertainties of the emission inventory and the effects of the background transport may also contribute to the overestimation. However, the correlation coefficients of PM$_{2.5}$ are 0.50 and 0.56 in the two episodes, respectively, and the IOAs are 0.64 and 0.70 in the two episodes, respectively, indicating good capture of the variable tendency and magnitude by the model. Both the fractional error and the fractional bias for PM$_{2.5}$ in the two episodes fulfill the performance criteria ($\text{FE} \leq 75\%$, $|\text{FB}| \leq 60\%$) given by Boylan and Russell (2006), implying a good performance in simulating PM$_{2.5}$. These results show that the model gives a reasonable simulation of the chemical species.

The modeled and satellite-observed mean CWPs were compared in Fig. 2 during the two analyzed episodes to evaluate the reasonability in simulating the aqueous SOA formations. Figure 2a and b show that in episode 1, during one summer period, the highest observed CWP (400–500 g m$^{-2}$) mainly appears over the Yellow Sea and Bohai Sea and in parts of southeastern and southern China, whereas the observed second highest values (200–300 g m$^{-2}$) are concentrated in from the Qinghai–Tibet Plateau to the North China Plain along the Yellow River Basin. The modeled high CWP (100–400 g m$^{-2}$) appears in the same regions as in the observational dataset. In episode 2, during one fall period (Fig. 2c and d), the observed highest (300–500 g m$^{-2}$) and second...
highest (200–300 g m$^{-2}$) CWP values are mainly concentrated in the Hetao area and the regions including northeastern China and Inner Mongolia, respectively. The modeled values (50–150 g m$^{-2}$) are also high over these regions.

Overall, there are obvious biases in the numerical values between the observed and simulated CWP$\_2$ and CWP$\_5$. According to the comparisons between the observed and modeled cloud fraction and precipitation (shown in Figs. S6 and S7), this is a result that has a lot to do with the uncertainties in the cloud fraction estimations, which indirectly lead to uncertainties in the simulations of the concentrations of SOAs. However, the mean distribution patterns of the simulated CWP during the two episodes are similar to the observational results, indicating few impacts on the simulated distribution of SOA. Both the simulated and observed CWP$\_2$ and CWP$\_5$ in episode 1 are higher than in episode 2, also implying differences in the simulation of SOA between the two episodes.

### 3.2 Model results and analyses

As described in Sect. 2.3 and with reference to previous research, the underestimation of glyoxal concentrations may partly result from the underprediction of the emissions of aromatic compounds. It is therefore necessary to evaluate the emissions of aromatic compounds during the analyzed episodes with the base model before designing and implementing the sensitivity case studies. For this purpose, comparisons were made between the simulated and observed concentrations of aromatic compounds in a similar manner to the work of Zhang and Ying (2011).

Figure 3 presents the ratio of the observed-to-predicted ($O/P$ ratio) concentrations of aromatic compounds for the original emissions (Fig. 3a and c) and a 3-fold increase in the emissions of aromatic compounds (Fig. 3b and d) during the two episodes. Figure 3a shows that there are biases between the observed and simulated concentrations of aromatic compounds in episode 1 at the original emission rates of the aromatic compounds. The $O/P$ ratios are less variable for ARO1 than for ARO2, with both the mean and median $O/P$ values being close to 2. The range of $O/P$ values is large for ARO2 and there is an order of magnitude difference between the lowest and highest ratios. Both the mean and median $O/P$ ratios of ARO2 are about 20. These results show that there are underestimations in the emissions of ARO1 and ARO2. In view of the limitations of the amount of observed data and the nonlinear relationship between emissions and the concentrations (Li et al., 2018), the underestimation in the emission rates has been conservatively estimated to be a factor of 3. As a result, when the emission rates of aromatic compounds are increased (Fig. 3b), the ratios of ARO1 and ARO2 in episode 1 are less variable, especially ARO2. The mean and median ratios of ARO1 are both around 1, whereas the two ratios for ARO2 decrease from 20 to 6. The results for episode 2 are similar to those in episode 1. Figure 3c shows that the original $O/P$ ratio of ARO1 varies between 1 and 3.5, with a mean value close to 2.
the ratio of ARO2 ranges from 1 to 5, with a mean ratio of about 4, both of which indicate unaccounted-for emissions of aromatic compounds. Figure 3d shows a clear decrease in the biases of the observed and simulated concentrations of aromatic compounds. The $O/P$ ratios of ARO1 and ARO2 are concentrated and vary between 1 and 2, with a mean ratio close to 1. However, it is difficult to determine whether the factor of 3 is the actual underestimation in the emission of aromatic compounds as a result of its dependence on space and time. For convenience, a factor of 3 is chosen here as a uniform scale with which to assess the unaccounted-for emissions of aromatic compounds.

Three sensitivity simulation case studies are designed based on these results. Case 0 is run with the three default SOA formation pathways included in the standard model. According to Fu et al. (2008), the default aqueous SOA formation pathway is included in the pathway $M$. To avoid double-counting the loss of dicarbonyls through the in-cloud oxidation and the following formed SOA, the default aqueous SOA formation pathway is turned off when case 1 is run with the pathway $M$ incorporated. According to previous studies (e.g., Stavrakou et al., 2009, 2010; Li et al., 2016), the yields of glyoxal from isoprene are very likely to be underestimated. Li et al. (2018) has run a simple sensitivity case to evaluate the impacts of underestimated glyoxal yields from isoprene on glyoxal simulation. From their results, a better agreement was achieved between the observed and simulated glyoxal after a 5-fold increase of glyoxal yields from isoprene. More detailed analyses can be found in Li et al. (2018). Thus, to take the effects of the underestimations of glyoxal concentrations into consideration, case 2 is designed with a 3-fold increase in the emissions of aromatic compounds and a 5-fold increase in the molar yield of glyoxal from isoprene based on the results of case 1.

Figure 4 compares the hourly concentrations of the simulated and observed SOAs during the two analyzed episodes;
the corresponding statistical parameters are listed in Table 2. The concentrations of SOAs are measured from PM$_1$. The observed PM$_0$ / PM$_{2.5}$ ratio of 0.77 (Xu et al., 2015) was used to convert the observed concentrations for comparison with the simulated results for PM$_{2.5}$.

In case 0, the SOA concentrations in episode 1 (Fig. 4a) are significantly underestimated by an average factor of 5.7, with the differences being as high as a multiple of $\geq 60$. As a result of the impacts of uncertainties in the meteorological fields and emissions, the variation trend of the concentration of SOA is not well captured by the model ($R = 0.21$). Similarly, although the variation in SOA is reproduced well in episode 2 ($R = 0.83$) (Fig. 4b), the concentration of SOA is still underpredicted by an order of magnitude. The biases between the observed and simulated concentrations of SOAs decrease in case 1 when pathway $M$ is added to the model, especially in episode 1. It shows that the mean concentration of SOA in case 1 increases by 3.65 µg m$^{-3}$ during episode 1, explaining 34.8% of the unaccounted-for sources of SOAs. The decreased mean bias and RMSE and the slightly increased IOA also indicate the more reasonable descriptions of the formations of SOAs. In episode 2, the SOA formed through pathway $M$ contributes less to the total concentration of SOA. The mean SOA concentration during episode 2 reaches 2.87 µg m$^{-3}$, an increase of 1.82 µg m$^{-3}$. However, the decreased bias statistical parameters (e.g., RMSE and MB) also indicate the more realistic description of contributions of dicarbonyls to SOA in this episode. The larger contribution of dicarbonyls from pathway $M$ to the concentration of SOA in episode 1 is attributed to the higher LWC and the larger amount of dicarbonyls produced from biogenic precursors in the summer than in the fall. When the impact of the underestimation of glyoxal on the concentration of SOA is taken into consideration, the concentration of SOA clearly increases in case 1. Compared to case 0, the mean concentration of SOA in episode 1 is significantly improved by a factor of 5.4 and comparable to the observations, whereas the increase in episode 2 is a factor of 6.2. The statistical parameters (e.g., IOA and RMSE) also show the better performance of case 2, indicating a more realistic and reasonable representation of the formation of SOA. Aromatic compounds are not only the precursors of glyoxal, but are also the precursors of SOA in the gas phase. A compositional analysis is therefore required to evaluate the individual contributions from dicarbonyls in case 2.

Figure 5 shows the mean contributions from different sources of SOAs in the three sensitivity case studies during the two analyzed episodes. AAQ, dicarbonyl-derived SOAs, contributes little in case 0 during the two episodes. The mean contribution of AAQ to the total concentration of SOAs in episode 1 is 2.39%, larger than in episode 2 (0.89%). As a result of the greater amount of emissions from biogenic sources in summer, SOAs formed from biogenic precursors (AISO+ATRP+ASQT) contribute more (26.3%) in episode 1 than they do in episode 2 (13.5%), whereas the contributions from anthropogenic precursors (AALK+ATOL+AXYL+ABNZ) are comparable between the two episodes. When pathway $M$ is included in the model, the contribution from AAQ to SOA in case 1 clearly increases and reaches 57.1% in episode 1 and 63.7% in episode 2. The significant contributions from AAQ in the two episodes indicate the major contributions from dicarbonyls through pathway $M$ to the formations of SOAs in summer and fall. When the impact of the underestimation of glyoxal is further taken into consideration in case 2, the mean concentration of AAQ in both episodes are higher than double that of case 1. In episode 1, AAQ (7.29 µg m$^{-3}$) accounts for 60.6% of the total SOA, exceeding the sum of the contributions from ARO1 (ATOL) and ARO2 (AXYL) (16.0%) and indicating the dominant contribution of the aqueous SOA formation to the concentration of SOA in summer. The increase in AAQ relative to case 0 compensates for about 69.1% of the unaccounted-for sources of SOAs. In episode 2, AAQ (4.21 µg m$^{-3}$) contributes 64.5% to the total concentration of SOA, which is also higher than the sum of ATOL and AXYL (15.3%), implying the dominant contributions of dicarbonyls to the concentrations of SOAs. The different contribution of AAQ in case 2 during the two episodes can be attributed to the different LWC and different amount of dicarbonyls produced from biogenic precursors.

It is clear that the biases between the observed and simulated concentrations of SOAs decrease when the contributions of unaccounted-for dicarbonyls to the concentrations of SOAs are considered, especially in summer. However, the sources of unaccounted-for SOAs cannot be explained completely. As a result of uncertainties in the description of known SOA formation processes and missing pathways that are not included in the model, for example, there are many uncertainties in glyoxal simulations (Li et al., 2018). There are also many uncertainties in incorporating pathway $M$ into the model, such as the effective radius of cloud droplets (the empirical values used instead of the actual values), the reaction uptake coefficient (updated values over disparate surfaces are discussed in Curry et al., 2018) and the liquid water content in clouds (the most uncertain parameter). Other pathways for the formations of SOAs, such as the uptake of isoprene epoxides on the surface of aerosols (Lal et al., 2012; Lin et al., 2013), the aging mechanism of semivolatile primary organic aerosols (Shrivastava et al., 2008), and the oxidation of primary semi- and intermediate VOCs, are not considered in this model. Besides, recent studies (e.g., Galloway et al., 2009; De Haan et al., 2018) have presented that the reactive uptake of glyoxal and methylglyoxal can be reversible, especially methylglyoxal, but we do not consider that in this study.

To distinguish the contribution of dicarbonyls to the concentration of SOA over China in case 2 from that in case 0, the distributions of dicarbonyl-derived SOAs and their contributions to SOAs (AAQ/SOA) over China in cases 0 and 2 are analyzed.
Figure 4. Hourly concentrations of the observed and simulated near-surface SOA concentrations in episode 1 (a) and 2 (b). Case 0 is the base run, case 1 is run with the incorporation of the reactive uptake of dicarbonyls (excluding the default in-cloud dicarbonyl oxidations), and case 2 is run with the glyoxal underestimation taken into consideration based on case 1.

Table 2. Performance statistics of the modeled and observed SOA concentrations (µg m$^{-3}$).

<table>
<thead>
<tr>
<th>Case</th>
<th>Episode</th>
<th>N</th>
<th>$C_{\text{model}}$</th>
<th>$C_{\text{obs}}$</th>
<th>MB</th>
<th>RMSE</th>
<th>R</th>
<th>IOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>822</td>
<td>2.21</td>
<td>12.69</td>
<td>-10.48</td>
<td>13.48</td>
<td>0.21</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>737</td>
<td>1.05</td>
<td>17.69</td>
<td>-16.64</td>
<td>23.52</td>
<td>0.83</td>
<td>0.45</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>822</td>
<td>5.86</td>
<td>12.69</td>
<td>-6.83</td>
<td>10.86</td>
<td>0.25</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>737</td>
<td>2.87</td>
<td>17.69</td>
<td>-14.82</td>
<td>21.21</td>
<td>0.84</td>
<td>0.50</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>822</td>
<td>12.03</td>
<td>12.69</td>
<td>-0.66</td>
<td>10.19</td>
<td>0.21</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>737</td>
<td>6.53</td>
<td>17.69</td>
<td>-11.16</td>
<td>16.56</td>
<td>0.83</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Figure 6a–b and c–d show the mean concentration of AAQ in cases 0 and 2, respectively, during the two episodes. For the base case, in episode 1 (Fig. 6a), the concentration of AAQ over China is $\leq 0.2$ µg m$^{-3}$. The higher concentration of AAQ (0.1–0.2 µg m$^{-3}$) is in the areas between the lower reaches of the Yellow and Yangtze rivers and in the Sichuan Basin, probably due to the higher LWC and a greater number of sources of dicarbonyls over these areas, as discussed in Sect. 3.1 and by Li et al. (2018). In episode 2 (Fig. 6b) the concentration of AAQ is $\leq 0.1$ µg m$^{-3}$ over the regions due to the lower LWC and the fewer sources of dicarbonyls from biogenic precursors. It is clear that the concentration of AAQ is improved in case 2, when pathway $M$ is added and the impact of the underestimation of glyoxal is considered. Overall, the concentration of AAQ is higher in eastern China than in the west during the two episodes. In episode 1 (Fig. 6c), the concentrations of AAQ mostly vary from 2 to 15 µg m$^{-3}$ over central and eastern China, with the highest value (5–10 µg m$^{-3}$) concentrated over regions from south of the lower reaches of the Yellow River to the south of Guangzhou Province as well as the Sichuan Basin. The concentration of AAQ in western China is also $\leq 1$ µg m$^{-3}$, with the lowest value ($\leq 0.1$ µg m$^{-3}$) in the Qinghai–Tibet Plateau. Outside China, the highest concentration of AAQ (15–20) appears in the northeastern India due to more primary sources of dicarbonyls from the large scale of postharvest paddy residue burnings (Chandra and Sinha, 2016) and the barrier of precursor gases diffusions from Himalayan orogen as well as the low wind speed.

Figure 6e–f and g–h show the spatial distribution of the mean AAQ/SOA in cases 0 and 2 during the two episodes, respectively. Figure 6e shows that the AAQ fraction over China in episode 1 is $\leq 10\%$, except in Yunnan Province and some parts of the South China Sea, where AAQ/SOA reaches 10%–20%. In episode 2 (Fig. 6f), the mean AAQ/SOA is $\leq 10\%$ over all regions. When the contributions of dicarbonyls from pathway $M$ and improved glyoxal to SOA are considered, there is a large increase in AAQ/SOA over these regions in two episodes. As shown...
consideration of the underestimation of glyoxal in the two
the SOA simulations with the inclusion of pathway
M
three sensitivity groups showed different improvements in
and 14 October to 14 November 2014. Comparisons between
of SOAs during the two episodes from 3 June to 11 July 2014
contributions of dicarbonyls from the reactive uptake process
The RAMS-CMAQ modeling system was used to assess the

4 Conclusions

The RAMS-CMAQ modeling system was used to assess the contributions of dicarbonyls from the reactive uptake process and unaccounted-for sources of glyoxal to the concentrations of SOAs during the two episodes from 3 June to 11 July 2014 and 14 October to 14 November 2014. Comparisons between the observed and simulated concentrations of SOAs from three sensitivity groups showed different improvements in the SOA simulations with the inclusion of pathway M and consideration of the underestimation of glyoxal in the two episodes. Due to the high LWC and large amount of dicarbonyls produced from biogenic precursors in summer, the contributions of dicarbonyls were greater in episode 1 than that in episode 2. When pathway M was added in case 1, the modeled mean concentration of SOA in episode 1 increased by 3.65 µg m\(^{-3}\), explaining about 34.8% of the unaccounted-for SOA sources, while there was a 1.82 µg m\(^{-3}\) increase in the mean result during episode 2. When the impacts of glyoxal underestimation were further taken into consideration in case 2, the modeled mean AAQ in episode 1 was improved to 7.29 µg m\(^{-3}\) and contributed 60.6% of the total concentration of SOA. The increase due to AAQ relative to case 0 is equivalent to 69.1% of the unaccounted-for source of SOA, whereas the mean concentration of AAQ in episode 2 was 4.21 µg m\(^{-3}\) and accounted for 64.5% of total concentration of SOA. Although the contributions of dicarbonyls to SOAs are different in the two episodes, the simulated SOA results are both improved and close to the observations, indicating a more realistic description of aqueous SOA formation.

The mean AAQ in case 2 during the two episodes was clearly improved over China relative to case 0 and was generally higher in the east than in the west. In episode 1, the highest value (10–15 µg m\(^{-3}\)) was seen in the areas around the lower reaches of the Yellow River, while the highest value (5–10 µg m\(^{-3}\)) in episode 2 was concentrated over regions from south of the lower reaches of the Yellow River to the south of Guangzhou Province as well as the Sichuan Basin. As a result, the contribution of AAQ to the concentration of SOA in two episodes was also improved in case 2. In episode 1, the fraction varied from 10% to 90% throughout China, with the highest contribution (70%–90%) in the coastal regions and offshore along the East China Sea to the South China Sea. The fraction of AAQ in the west is relatively low and usually ≤50%. However, the contribution of AAQ to SOA is up to 80% in southwestern regions (e.g., Yunnan Province and parts of Sichuan Province). In episode 2 (Fig. 6h), the contribution of AAQ to SOA is in the range of 10%–80% throughout China. In central and eastern areas, the fraction of AAQ is generally in the range of 50%–70% and up to 80% in a small portion of northeastern China. The fraction of AAQ in the west is also lower and usually ≤50%. However, the contribution of AAQ to SOA is up to 60%–70% in a small part of Sinkiang Province.
Figure 6. Modeled distributions of the mean (a, b) wind field with SOA formed from dicarbonyls (AAQ) in case 0, (c, d) AAQ in case 2, (e, f) AAQ/SOA in case 0 and (g, h) AAQ/SOA in case 2, over the regions during two episodes. Case 0 is the base run; case 2 is run taking into consideration the incorporation of the reactive uptake of dicarbonyls (excluding the default in-cloud dicarbonyl oxidations) and the underestimation of glyoxal.
the evaluations of the underestimations in the aromatic compound emissions which are not well discussed due to the limited observed data can also bring the uncertainties into our results. Thus, more work on the dicarbonyl-derived SOA needs to be done in the future.

Besides, the aqueous SOA formation is not only relative to the distributions of dicarbonyl concentrations, but also depends on the liquid water content (LWC). Due to the large space and time dependence, one single station measurement of SOA concentration is not enough to evaluate the model performance over China, especially the impacts of glyoxal underestimations on dicarbonyl-derived SOA. Thus, more observed SOA data from different stations need to be collected and used for comparisons to reduce the uncertainties in the conclusions.

Data availability. The datasets of measured SOA concentrations, observed aromatic compound concentrations and corresponding modeled results used in this study can be accessed by contacting sunyele@mail.iap.ac.cn, tgz@dq.cern.ac.cn and mgzhang@mail.iap.ac.cn, respectively.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/acp-19-6481-2019-supplement.

Author contributions. In this study, JL designed the sensitivity experiments, developed the model code and performed the corresponding simulations. MZ co-designed the experiments and provided valuable advice about the model operations. YS carried out the measurements of SOA and provided the corresponding data for evaluating the modeled results. GT and FW carried out the measurements of SOA and provided the corresponding data for evaluating the simulated results. YX provided valuable advice on model result analysis.

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

Acknowledgements. This study was supported by the National Key R&D Programs of China (no. 2017YFC0209803), the National Natural Science Foundation of China (91544221), and the Beijing Municipal Science and Technology Project (ZL17110000617002). We thank Qi Ying from Texas A&M University for helping to incorporate the reactive uptake of dicarbonyl pathways into the model.

Review statement. This paper was edited by Xiaohong Liu and reviewed by three anonymous referees.

References


www.atmos-chem-phys.net/19/6481/2019/

Atmos. Chem. Phys., 19, 6481–6495, 2019
6494

J. Li et al.: Assessment of dicarbonyl contributions to secondary organic aerosols


