We thank the reviewers for their thoughtful comments that have helped us clarify and improve the manuscript. We have reproduced the reviewer comments in black text. For ease of review, our responses are given in blue text, while the text added to the manuscript is given in *blue Italics.*

Referee 1 Comments

R1.1. This paper describes modeling the secondary organic aerosol (SOA) formation in Los Angeles as measured during the CalNex study in May/Jun 2010. The paper evaluates a variety of different models/parameterizations – WRF-CHEM, WRF-CMAQ, and a box model with several different mechanisms. None of the models are novel, but applying them to CalNex is new (EPA has been doing some evaluation of the CMAQ model, e.g. a paper is up on ACPD right now http://www.atmos-chem-physdiscuss.net/15/157/2015/acpd-15-157-2015.html). Given the comprehensive nature of the CalNex dataset it is valuable to use it to evaluate SOA models. The analysis is similar to previous work the authors have published for Mexico City.

The conclusions from the work appear to be largely consistent with those of previous studies. The old, traditional SOA models that only consider a relatively small number of volatile organic compounds cannot explain the measured SOA, even if one makes extreme corrections to yields for potential vapor wall losses to chamber walls. The paper shows that mass closure on the SOA can be achieved if one includes semivolatile and intermediate volatility precursors in a model. The paper evaluates several different schemes for doing this. The paper evaluates models against SOA mass, degree of oxygenation, and fossil/modern split. The parameterization with IVOC/SVOC all appear to work reasonably well (but not perfectly), but are not well constrained with data. The simple IVOC/SVOC schemes the authors consider appear to overpredict the amount of SOA at longer time scales (~ 3 days), presumably (as mentioned by the authors) because they do not include any fragmentation. There is reasonable agreement between fossil modern carbon data and their apportionment. Finally, the authors show that a simple parameterization they developed in Mexico City for SOA formation appears to work in LA as well. These are basically the same set of conclusions as the Mexico City work, with the addition of fossil/modern carbon.

This is a well written paper. It is long and very comprehensive, but I found it not difficult to read and follow. It makes a contribution. Its weakness is that it does not break much new ground. There is value to show that they largely reach the same conclusions as they did with Mexico City. I would recommend the paper for publication after the authors address the following comments.

We thank the reviewer for his/her good summary of the paper and comments.

We note that it is not obvious that the results from LA were going to be similar to those from Mexico City, and many scientists had expressed skepticism that this would be the case. For example a prominent EPA scientist stated in a public comment on a past ACPD paper that "I haven't paid enough attention to Mexico City modeling studies in the past because [...] I did not expect the findings from intensive field campaigns in Mexico City to be generally transferable when modeling other less-polluted regions of the world." [P. Bhave, Atmos. Chem. Phys. Discuss., 10, C878–C884, 2010].

R1.2. "The work presented here quantitatively demonstrates that PAHs are relatively unimportant compared to other precursors such as methylbenzenes" This statement is too strong. The model only included a few PAH (napthalenes, methylnaphthalene). It seems clear that these PAHs do not appear to be major precursors, but large amounts of unspeciated IVOCs include a fair number of more alkylated or funcationalized PAHs. The authors claim that unspeciated IVOC are mainly alkanes, but that has not really been established. If a reasonable fraction (10s of %) were PAHs then PAHs could play an important role.

We have substantially modified the relevant discussion and, for reference, have pasted the new text below.

"Figure 5B also shows a comparison for the naphthalenes. The tracer estimates are over an order-of-magnitude higher than the model predictions when using the SOA yields from the literature (which are ~20% for the conditions of our study) and the emission ratios determined from the regression analysis of nighttime measurements shown in Figure SI-3. The model is also

run using the empirically adjusted emission ratios that better match the observed concentrations of the naphthalenes. The model for this variation is still much lower than the tracer estimate. As an additional sensitivity study, we also run the model with the adjusted emissions and a yield of 150% that places all the oxidized mass in the C=1 μg m^{-3} volatility bin. This last variation represents an upper limit estimate of SOA from naphthalenes, in which nearly all of the mass plus the added oxygen partitions to the particle phase, which is much higher than laboratory observations. The tracer estimate, however, is still about a factor of two higher than the model. It is known that the tracer estimate is an upper limit, because the tracer compound, phthalic acid, may not be a unique tracer, and it potentially could be emitted from primary sources (Kleindienst et al., 2012). However, there may be other alkylated or functionalized PAHs that are not explicitly accounted for in the box model, and some of them might produce the tracer.*

The best estimate from the model with the adjusted emissions results in 0.7% of the predicted SOA being formed from the measured napthalenes. Utilizing the upper limit of the model results for the PAHs, including that from the parameterization with a purposefully high yield, it is apparent that naphthalene, 1-methyl naphthalene, and 2-methyl naphthalene account for less than 4% of the SOA mass. While previous work has suggested that PAHs are important precursors for SOA in the Los Angeles Basin (Hersey et al., 2011) these earlier findings were qualitative and based on the observation of phthalic acid in samples. The work presented here, both the modeling results as well as the tracer results, quantitatively demonstrates that SOA from identified PAHs is relatively small but not negligible when compared to the total SOA concentration. An upper limit for the contribution of this group of precursors is 8(±3)% of the SOA. This percentage is calculated using the tracer method in which the SOA concentration from PAHs is higher than in the box model and a 30% uncertainty for the SV-OOA concentration. Lastly, we note that no suitable tracers for alkane oxidation have been identified yet, which prevents carrying out similar model/tracer comparisons with respect to the P-S/IVOCs, since these compounds are thought to be composed primarily, although not exclusively, of alkanes."

R1.3. The authors should report the effective yields of the SI-SOA precursors. How do they compare of the recent work of Jathar (PNAS, 2014), which derived yields from chambers? How do they compare to the yields proposed by Zhao (EST, 2014) who performed this sort of model

with more extensive set IVOC data? How do they compare to smog chamber data for traditional precursors (alkanes, aromatics, etc.)?

The questions posed by the reviewer regarding the effective yields are very interesting. We have responded to them by adding the following paragraph to Section 3.1.1.

"In addition, the effective SOA yields predicted in the box model for P-S/IVOCs can be compared against those determined in previous modeling and smog chamber studies. The effective yield is a function of photochemical aging, and thus for the purpose of this comparison we focus on the effective box model yields for 12:00 – 15:00 when there was a moderate amount of photochemical aging (5 h at an average OH concentration of 1.5 x 10^6 molecules cm^{-3}) comparable to the degree of aging typically achieved in chambers. During this period the effective yields for P-S/IVOCs were 12%, 27%, and 36% for ROB+TSI, GRI+TSI, and PYE+TSI, respectively. Zhao et al. (2014) recently carried out a modeling study of SOA formed in Pasadena that was constrained with an extensive set of IVOC data and found an overall SOA yield for IVOCs of 29%, which falls within the range of effective yields for P-S/IVOCs that are predicted by the box model for the three different parameterizations. Jathar et al. (2014) also recently estimated from chamber studies an effective SOA yield of 10 – 40% for unspeciated organic emissions from combustion sources, which is also consistent with the P-S/IVOC yields from our box model. For reference, the effective yields for the aromatic VOCs m-xylene, toluene, and benzene under high-NO_X conditions in chamber studies range from 4 – 28% depending on the precursor identity and chamber conditions (Ng et al., 2007). Similar chamber studies on 12-carbon alkanes determined effective yields ranging from 11% – 160%, where the highest yield corresponded to a cyclic alkane (Loza et al., 2014). In general, it appears that the effective yields resulting from the box model for the lower photochemical ages used here are similar to those determined from other chamber and modeling studies."

R1.4. The paper uses the results to apportion SOA to different sources. Since the SVOC and IVOC concentrations are scaled with primary organic aerosol; this apportionment is presumably the same as the primary organic aerosol. The authors should clarify this point in the manuscript. I

suspect it is highly uncertain. For example, gasoline vehicles emit relatively little POA but likely contribute significant SOA (Jathar et al. 2014).

The reviewer is correct that the initial SVOC and IVOC concentrations are scaled from primary organic aerosols. However, the apportionment of SOA from P-S/IVOCs is not the same as that for primary organic aerosol, because the amount of SOA from each source is also a function of the photochemical age. For example, there are substantial cooking emissions in the evening, but these produce little SOA over the urban scale considered here, due to the short ambient photochemical ages at that time. In addition, the SOA apportionment takes account of SOA from VOCs, which are largely emitted by gasoline vehicles, and which are not scaled to POA.

Still, the reviewer is correct that in scaling the P-S/IVOCs to POA we have assumed that the P-S/IVOCs to POA ratio as a well as the volatility distribution of P-S/IVOCs is the same for the different sources, which may not be accurate. To explain this uncertainty we have added the following paragraph to Section 3.1.4.

"The uncertainties in Figure 6 do not however account for certain assumptions that were made in order to perform the source apportionment. In particular, it was assumed that the P-S/IVOCs to POA ratio as well as the volatility distribution of P-S/IVOCs is the same for all sources, which is likely not the case. However, to our knowledge there is insufficient information in the literature to prescribe different volatility distributions for the different sources."

R1.5. Zhao (EST, 2014) recently reported measurements of more IVOCs, including an estimate of the unspeciated IVOC. It appears that the naphthalene data came from the same dataset? How do the author's estimates of unspeciated IVOC inferred from primary organic aerosol compare to the measured data of Zhao et al.? How do the predictions of the authors compare to Zhao et al.?

In response to this comment we have added a table to the supporting information (Table SI-3) that compares the estimated IVOCs from the model versus the measured data of Zhao et al. (2014). The table is also included below for the reviewer's convenience. We have also added a paragraph to Section 3.1.1 that discusses the comparison (see quoted text below).

In addition, we have performed a sensitivity study in which all the IVOC emissions from cooking as well as diesel/gasoline have been reduced by one-half. This sensitivity study is summarized in a new paragraph in Section 3.1.5 (text is also quoted below).

"It is also possible to compare the predicted IVOC concentrations in the box model versus the concentrations measured by Zhao et al. (2014). The comparison is summarized in Table SI-3 of the supporting information. In total, the initial IVOC concentrations in the box model are two times higher compared to those determined from measurements (16 $\mu g\ m^{-3}$ versus 8(±1) $\mu g\ m^{-3}$). In addition, there is a larger difference for the $C^* = 10^3$ bin (2.5 $\mu g\ m^{-3}$ versus 0.2(±0.1) $\mu g\ m^{-3}$).At the same time, the model used by Zhao et al. to predict urban SOA is lower than the measurements by 50% on the urban scale, whereas as the box model used here does not exhibit such a low bias. Given these differences we have run two sensitivity studies to explore how the model predictions depend on the IVOC emissions that are discussed in the following sections. The first sensitivity study reduces the emission of P-S/IVOCs from cooking emissions to zero (Section 3.1.2), and the second sensitivity study reduces all IVOC emissions by one-half (Section 3.1.5). Both of these variations greatly improve the agreement between the modeled and measured IVOC concentrations."

"Table SI-3. *Initial concentrations of primary IVOCs predicted by the box model (ROB parameterization) in comparison with data from Zhao et al. (2014) as a function of the saturation concentration (C*) at 298 K. Note that the corresponding results for the GRI parameterization are very similar with the concentration being 8% higher due to differences in the* ΔH_{vap}.

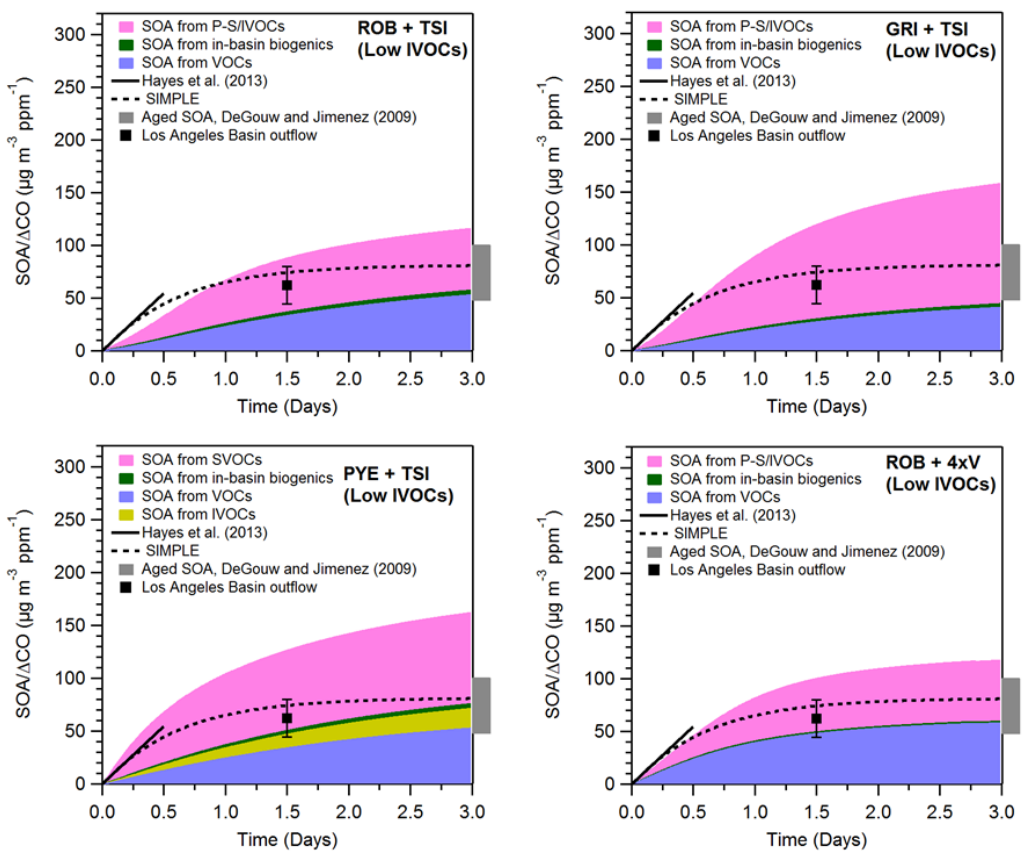
C* (μg m^{-3})	**Estimated Primary IVOCs (μg m^{-3})**	**Estimated Primary IVOCs without cooking emissions (μg m^{-3})**	**Measured Primary IVOCs (μg m^{-3})**
10^3	2.47	1.70	0.21 (± 0.07)
10^4	3.30	2.27	1.39 (± 0.29)
10^5	4.12	2.84	2.64 (± 0.64)
10^6	6.59	4.54	3.82 (± 0.99)

From Section 3.1.5:

"A third explanation is the potential overestimation of IVOC emissions in the box model. As discussed in Section 3.1.1, the initial concentration of primary IVOCs in the model is a factor of 2 higher than the values determined from field measurements by Zhao et al (2014). To investigate this possibility, a sensitivity study was run in which the initial concentrations of primary IVOCs in the volatility bins C=* 10^3, 10^4, 10^5, *and* 10^6 *were decreased by one-half. The results of this sensitivity study are shown in Figure 8. In general, ROB+TSI, GRI+TSI, PYE+TSI, and ROB+4xV all show better agreement with measurements at long photochemical ages, although all four variants still overestimate the measurements. For shorter photochemical ages (in the urban scale) ROB+TSI under-predicts the SOA concentration, whereas GRI+TSI and ROB+4xV both predict SOA/ΔCO ratios that are not significantly different from the measured values (Hayes et al., 2013), and lastly PYE+TSI overestimates the SOA concentration.*

Thus, IVOCs emissions that are too high in the box model may be responsible for some, but not all, of the overestimation of SOA concentrations at long photochemical ages.

For reference, we note that when the IVOC concentrations are halved the four variations all predict less SI-SOA for the Pasadena ground site (Figure SI-7), but the contribution of P-S/IVOCs to SOA formation remains important: 59 – 73% (ROB+TSI), 72 – 80% (GRI+TSI), 79 – 92% (PYE+TSI), 38 – 48% (ROB+4xV).Furthermore, all four variations still predict a fossil fraction of urban SOA consistent with the ^{14}C measurements at the Pasadena site: 66(±9)%, 64(±10)%, 61(±12)%, 78(±6)%, respectively. Note that in calculating these fossil fractions the IVOCs emissions from cooking and gasoline/diesel were reduced by the same amount (i.e. one-half).”



“Figure 8: *SOA concentration predicted by the ROB+TSI, GRI+TSI, PYE+TSI, and ROB+4xV parameterizations for up to 3 days of photochemical aging at a reference ·OH concentration of 1.5×10^{6} molec cm^{-3}.These predictions correspond to the sensitivity study in which the concentration of IVOCs in the volatility bins $C^{*}=10^{3}$ – 10^{6} were reduced by one-half. Also*

shown in the three panels is the same result for the SIMPLE model using the optimized parameters (see Section 3.3 for further discussion). Note that the SOA concentrations have been normalized to the background subtracted CO concentration to account for changes in emission strengths, and the processed data are identified by the symbol SOA/ΔCO. In addition, the SOA/ΔCO data determined for the Pasadena site from the measurements of Hayes et al. (2013) are shown (black line) as well as similar airborne measurements downwind of Pasadena performed by Bahreini et al. (2012) aboard the NOAA P3 (black marker). The Bahreini et al. point corresponds to an average of all LA Basin outflow data between 1 – 2 days of photochemical aging. The OA/ΔCO ratio reported by de Gouw and Jimenez (2009) is also indicated (gray box) to serve as an estimate of OOA/ΔCO in highly aged air masses."

R1.6. "Indeed, decreases in SOA concentration at high photochemical ages have been observed in flow-tube studies, although typically at photochemical ages much longer than 3 days (George and Abbatt, 2010)." More recent flow tube studies (Lambe et al. EST 2013; Tkacik et al. EST 2014) suggest that the fall off may occur at shorter photochemical ages (between 2 and 3 days).

The shorter photochemical ages in the Tkacik et al. study are not a real difference, but rather the result of that study using a higher OH concentration for the calculation of the ages. The decrease in the Tkacik et al. data occurs at 5 days, when one uses the same OH = 1.5 x 10^6 molec cm^{-3} used in this paper. Lambe et al. (ES&T 2013) does not include any such data. The reviewer may have meant to cite Lambe et al. (ES&T 2012), for which the decreases are observed between 4-8 days of equivalent exposure at OH = 1.5 x 10^6 molec cm^{-3}.

In addition the apparent losses in recent flow tube studies where the SOA is formed in the reactor may be due to lack of SOA formation (e.g. due to oxidation of semivolatile species before they have time to condense, leading to fragmentation). The studies of George and Abbatt started with pre-existing OA and would not suffer from that potential problem, and thus we consider them more trustworthy for this point. Thus, we have not modified the text in response to this comment.

R1.7. SIMPLE parameterization is interesting. It is not clear how much fitting versus predicting is being done. This needs to be clarified in the manuscript. It appears that it has been retuned for

CalNex. If so then that is not as impressive – it is just a curve fit. If it is using previous parameters then that would be good. It would be interesting to better characterize the atmospheric conditions under which it may be applicable / influence SOA formation (VOC/NOx, etc.). Presumably LA and Mexico City are pretty consistent.

We have modified the text to clarify the discussion of the SIMPLE parameterization (see below). We agree with reviewer that it would be interesting to better characterize the performance of SIMPLE under various atmospheric conditions and have noted this now in the text. We also note that we are proposing this parameterization for urban SOA sources, and not for, e.g., biogenic SOA. It is difficult to carry out the characterization described by the reviewer for our case study, however, due to the limited set of atmospheric conditions explored. For example, in Los Angeles high-NO_X conditions are generally observed. Thus, we recommend that future modeling studies also apply this parameterization to studies in other locations.

"We replaced the SOA parameterizations discussed above with the SIMPLE parameterization just described, and ran the box model for a large number of possible parameter value combinations (i.e. emission ratio of VOC/CO and ·OH rate constant). Figure 11A shows the difference between model and measurements over that parameter space. The diurnal cycle predicted by the SIMPLE parameterization with the optimum parameters for Pasadena is shown in Figure 11B. The SIMPLE model with the optimized parameters performs comparably to the more complex parameterizations used in this work. At the same time, the SIMPLE parameterization is unable to capture perfectly the location of the peak in time because it depends solely on CO and photochemical age. The CO concentration at the site peaks at 12:00 and photochemical age peaks at 13:00 (Figure 3A) while the measured SOA has a broad peak between 14:00 – 16:00. The fact that SOA does not peak at the same time as CO and photochemical age indicates the assumption in SIMPLE that VOC*/CO does not vary in time is probably not completely accurate. Still, the performance of the SIMPLE parameterization for urban SOA is sufficient for many applications and certainly far better than many models currently used.*

Interestingly, the optimal model parameters for Mexico City and Pasadena are very similar. In other words, when tuning the model separately for each city, the parameters obtained

are identical within the estimated uncertainties. This result suggests SIMPLE, with the parameters reported for Mexico City or Pasadena, can be applied to other polluted urban regions as well. In addition, the optimal parameters for Pasadena (and Mexico City) are consistent with the OA/ΔCO ratios observed for highly aged air masses by Bahreini et al. (2012) from the NOAA P3 aircraft in the LA basin outflow, as well as for other urban areas as summarized by de Gouw and Jimenez (2009) and shown in Figure 7. However, it should be noted that a range of SIMPLE parameter combinations still remains in which the different combinations perform similarly in the model/measurement comparison, and this range is indicated by the dashed box in Figure 11A. While the SIMPLE model is promising, additional work should be carried out to verify the optimal SIMPLE model parameters including analysis of data for a broad range of ages, e.g., by utilizing results from ambient air processed by oxidation flow reactors (Ortega et al., 2013). Also, the accuracy of the SIMPLE model for predicting urban SOA under a variety of atmospheric conditions should be explored (e.g. VOC/NO_X or relative amounts of gasoline versus diesel emissions.) Finally, we note that the SIMPLE model parameterizes urban SOA, and is not applicable to biogenic SOA."

R1.8. The importance cooking emissions to SOA formation is interesting. I am not aware of data to support this. I believe that the 1.5 factor of Robinson et al. (Science 2007) is based on data for combustion sources. Is there evidence from source test data that it might also apply to cooking? What evidence is there for SOA from cooking?

This is an excellent question. In essence, there is evidence for SOA from cooking, but the number of studies and thus constraints available for developing cooking SOA parameterizations is currently very limited. In particular the Prévôt group at PSI (coauthors in this paper) has a paper in preparation detailing results from one extensive study where significant SOA formation was clearly observed, and that has been presented at the AAAR Annual Meeting and other conferences.

We hope that our study will highlight the potential importance of cooking SOA and motivate further work on this topic. We have added the following paragraph to the end of Section 3.1.4 in

order to better explain the potential importance as well as the uncertainty regarding cooking SOA.

"Lastly, the percentage of SOA attributed to cooking emission in this work also requires discussion. Compared to gasoline or diesel vehicles there is relatively little data on the SOA forming potential of cooking emissions, but nevertheless there is both direct and indirect data supporting the SOA forming potential of cooking emissions. First, it is clear from numerous source apportionment studies that cooking emissions are a source of organic matter in the atmosphere (e.g. Robinson et al., 2006; Mohr et al., 2011; Hayes et al., 2013). Second, molecular speciation of cooking emissions has demonstrated that cooking activities emit a variety of volatile and semi-volatile compounds that are known SOA precursors (Schauer et al., 1999, 2002). Third, chamber studies have demonstrated SOA formation from cooking emissions. The latter results have been presented at several major conferences, but not yet been published in the peer-reviewed literature (El Haddad et al., 2012). Thus, it is reasonable to conclude that SOA models should include the SOA resulting from chemical processing of cooking emissions, but there is a lack of chamber yields that could be used to develop specific SOA parameterizations We have therefore assumed that SOA from cooking emission can be described using the same parameterizations as used for SOA from vehicular P-S/IVOCs. We also perform a sensitivity study where we assume that cooking emissions do not produce any SOA. Ultimately, the percentage of SOA from cooking emissions reported here should be considered a first-order estimate that should be updated when additional data regarding SOA from cooking emissions becomes available."

References.

Jathar, S. H.; Gordon, T. D.; Hennigan, C. J.; Pye, H. O. T.; Pouliot, G.; Adams, P. J.; Donahue, N. M.; Robinson, A. L., Unspeciated organic emissions from combustion sources and their influence on the secondary organic aerosol budget in the United States. Proceedings of the National Academy of Sciences 2014, 111, (29), 10473-10478.

Tkacik, D. S.; Lambe, A. T.; Jathar, S.; Li, X.; Presto, A. A.; Zhao, Y.; Blake, D.; Meinardi, S.; Jayne, J. T.; Croteau, P. L.; Robinson, A. L., Secondary Organic Aerosol Formation from in-Use

Motor Vehicle Emissions Using a Potential Aerosol Mass Reactor. Environ. Sci. Technol. 2014, 48, (19), 11235-11242.

Lambe, A. T.; Cappa, C. D.; Massoli, P.; Onasch, T. B.; Forestieri, S. D.; Martin, A. T.; Cummings, M. J.; Croasdale, D. R.; Brune, W. H.; Worsnop, D. R.; Davidovits, P., Relationship between Oxidation Level and Optical Properties of Secondary Organic Aerosol. Environ. Sci. Technol. 2013, 47, (12), 6349-6357.

Lambe, A. T.; Onasch, T. B.; Croasdale, D. R.; Wright, J. P.; Martin, A. T.; Franklin, J. P.; Massoli, P.; Kroll, J. H.; Canagaratna, M. R.; Brune, W. H.; Worsnop, D. R.; Davidovits, P., Transitions from Functionalization to Fragmentation Reactions of Laboratory Secondary Organic Aerosol (SOA) Generated from the OH Oxidation of Alkane Precursors. Environ. Sci. Technol. 2012, 46, (10), 5430-5437.

Zhao, Y.; Hennigan, C. J.; May, A. A.; Tkacik, D. S.; de Gouw, J. A.; Gilman, J. B.; Kuster, W. C.; Borbon, A.; Robinson, A. L., Intermediate-Volatility Organic Compounds: A Large Source of Secondary Organic Aerosol. Environ. Sci. Technol. 2014, 48, (23), 13743-13750.

We thank the reviewers for their thoughtful comments that have helped us clarify and improve the manuscript. We have reproduced the reviewer comments in black text. For ease of review, our responses are given in blue text, while the text added to the manuscript is given in *blue Italics.*

Referee 2 Comments

R2.1. Hayes et al. present results from box model simulations of in-basin SOA formation focusing on Pasadena, CA and measurements made during the CalNex 2010 study. Their box model is explicitly constrained by observations of ambient traditional VOC concentrations. They also include emissions of intermediate and semi-volatile organic compounds (I/SVOCs) that are not included in standard emissions inventories. They link their I/SVOC emissions to POA emissions and assume that all POA sources have identical volatility distributions. They also consider a first order estimate of the influence of chamber biases on SOA yields by multiplying their yields of semi-volatile VOC product species by 4. They have considered a number of different schemes for treating the oxidation of S/IVOCs and even of VOCs. They find that the various combinations all exhibit a diurnal profile shape consistent with the observations, but that certain model formulations perform better than others when the model results are compared with observations of the semi-volatile oxygenated organic aerosol (SV-OOA) factor determined from PMF analysis of the ambient organic aerosol time-series. They conclude that, overall, there is a need for some contribution of I/SVOCs to allow for simulation of SOA levels that are consistent with the average SV-OOA diurnal profile. They find that biogenics from within LA ("in-basin") contribute negligibly to the total in-basin SOA, although biogenic-SOA may contribute to the background OA burden. They find when they extrapolate their simulations based on these literature parameterizations to longer times that they overestimate the amount of SOA that should be formed.

Overall, the authors do a good job explaining what they have done, describing the results, and comparing between the different literature model formulations. I believe that this work should be publishable after they address the comments below.

We thank the referee for their thoughtful review and have provided point-by-point responses to their comments below.

R2.2. The authors apply a number of different model formulations, taken from the literature, for the simulation of SOA concentration diurnal profiles within a constrained box model. Each of these model formulations/parameterizations has particular limitations that ultimately lead to differences in the box model results. Some aspects of the model formulations are more constrained than others. For example, the VBS yields for the traditional VOCs are constrained from comparisons with chamber observations whereas the literature ageing scheme(s) applied have generally weak, or even no, constraints. The O/C parameterization for S/IVOCs are similarly underconstrained, while the O/C estimates for SOA from VOCs are guided by observations from laboratory studies. The S/IVOC emission scheme is linked to literature measurements of POA volatility, although connecting these observational constraints on the volatility distribution to actual emissions is challenged by a lack of knowledge regarding the conditions under which the emission inventory was determined. I therefore suggest that it would be useful if the authors were to note a bit more explicitly what aspects of the literature parameterizations are more/less constrained by previous observations. They already do a good job of describing the model formulations, but some minor addition regarding the nature of these formulations might facilitate greater understanding by the reader. I also suggest that, given the inherent uncertainties in the parameterizations, that the authors limit the scope of any conclusions regarding whether one particular parameterization is particularly better than another.

We generally agree with the descriptions of the reviewer in this comment. Some of the parameterizations are based on a larger amount of data than others, and some have only weak constraints. We have included all the constraints that we were aware of, but in some cases assumptions needed to be made due to the limitations of the available information. In all cases we have made what we thought was the most reasonable assumption, stated clearly what was done, and in multiple cases we have presented several model alternatives to illustrate the sensitivity of the results to different assumptions.

We have modified the abstract and the conclusions to better clarify these points. We have removed the sentence that stated that the GRI parameterization did better in reproducing the observations. In particular, the following text has been added to the abstract:

"*The results from the 3 parameterizations show large differences (e.g. a factor of 3 in SOA mass) and are not well constrained, underscoring the current uncertainties in this area.*"

and

"The relative contribution of each source is uncertain by almost a factor of 2 depending on the parameterization used."

R2.3. Some additional clarification regarding the specification of the BVOC emissions would be useful.

We have added a paragraph to the manuscript to clarify the specification of the biogenic VOC emissions. As an example, we have also included in Figure SI-2 a comparison of the diurnal cycle predicted and measured for isoprene.

"The biogenic VOCs are not expected to be emitted proportionally with CO, and therefore the approach described in the previous paragraph cannot be used to specify the biogenic VOC emissions. Rather, the emissions of biogenic VOCs were adjusted empirically to match the observed concentrations of isoprene and terpenes, after accounting for anthropogenic isoprene using Δ(isoprene)/ΔCO (Borbon et al., 2013). Only ~4% of the daily average isoprene is from anthropogenic sources. In addition, the diurnal profile of emissions was assumed to be proportional to temperature.

The model consistency with the VOC measurements, including for biogenic VOCs, is evaluated by comparing the measured and modeled diurnal cycles. Some of the cycles compared are given in Figure SI-2 as an example. It is observed that the model is generally consistent with the biogenic VOC measurements."

R2.4. P32340: The authors should provide some brief discussion regarding their use of a constant [OH] in the simulations.

We have added substantially more discussion regarding the use of constant [OH] in the simulations. Please see the updated text below, which is from Section 2.3.

"The second consideration is that the purpose of using the ratio of VOC concentrations is to determine the ·OH exposure for the air mass at the Pasadena site. (·OH exposure is the concentration integrated over time for an air parcel.) While the ·OH exposure for the site is therefore well-constrained, the actual ·OH concentration in the modeled air parcel as a function of time is not as well-constrained. Thus, the photochemical ages used here (Figure 3) are calculated using an average ·OH concentration of 1.5×10^{6} molec cm^{-3}, as described in our previous work (Hayes et al., 2013), and the model is run with the same concentration. Insofar as the model produces the same ·OH exposure as determined from measurements, which is always the case in this modeling study, the actual concentration of ·OH used in the model is not expected to substantially influence the results. In other words, while the concentration ·OH in the model is assumed to be 1.5×10^{6} molec cm^{-3}, the integral of the ·OH concentration over time is constrained by the observed VOCs ratios. As expected, in the middle of the day the photochemical age will be longer than the transport age, and the opposite will be true during periods with low ambient ·OH."

R2.5. Section 2.6: Some clarification regarding the correction for "higher OA concentrations upwind of Pasadena" would be useful.

We have added a more in-depth introduction to this section in order to clarify this correction (see below). Perhaps, a useful way to understand this correction is the following: the initial concentration of precursors are calculated using the emission ratios to CO, which are multiplied by the measured excess CO (above the background level) to determine the initial precursor concentrations. Given that CO and precursors should undergo the same amount of vertical and horizontal dilution during transport, this approach implicitly dilutes the precursors by the appropriate factor <u>before oxidation occurs</u>. In reality, the species in the air parcel are oxidized

and diluted at the same time. This difference then means that the partitioning calculated in the model upwind of Pasadena will not be perfectly accurate. (However, the partitioning calculated at Pasadena is accurate.) We have applied a correction for this effect, although this correction is actually small due to the fact that the aerosol mass is distributed in volatility space over many orders of magnitude, whereas the OA concentration generally varies over a much smaller range (2 - 20 μg m^{-3} for Pasadena).

"2.6. Correction for changes in partitioning due to emissions into a shallower boundary layer upwind of Pasadena

To account for changes in partitioning due to lower planetary boundary layer (PBL) heights, and thus, increased OA concentrations upwind of Pasadena, the concentrations of POA, V-SOA, and SI-SOA are increased upwind of Pasadena beyond the amount already simulated in the model. This correction is necessary because using CO as a conservative tracer of emissions does not account for how the shallow boundary layer over Los Angeles in the morning influences partitioning between the gas and particle phases. Specifically, during the afternoon Pasadena is a receptor site for pollution from downtown Los Angeles that was generally emitted into a shallower boundary layer during the morning. The reduced vertical dilution will lead to higher concentrations of POA as well as any urban SOA formed, which in turn leads to higher partitioning to the particle phase and less gas phase oxidation of primary and secondary S/IVOCs."

R2.6. Questions and concerns regarding results and discussion:
I suggest that the introduction of the model variant that considers the influence of vapor wall losses be moved to methods.

In accordance with the reviewer's comment we have moved this subsection to the methods section.

R2.7. P32347/L10: The authors conclude here that their simulation results point to the "importance" of S/IVOCs. "Importance" is such a cagey word. I suggest that the authors be more

precise, stating that their results suggest that S/IVOCs contribute anywhere from X-Y% of the total SV-SOA.

We agree with the review that giving a percentage would be more precise, and so we have calculated the percentage of SOA formed from P-S/IVOCs in each model and added text describing the results to the manuscript (see quoted text below).

"Specifically, the contribution to total SOA from P-S/IVOCs in the box model is 65-75% (ROB+TSI), 80-87% (GRI+TSI), 80-92% (PYE+TSI), and 44-51% (ROB+4xV). The range indicates the variation in the contribution with the time-of-day. Thus, only in the ROB+4xV model variation is the estimated contribution to SOA from VOCs generally larger than or equal to that from the P-S/IVOCs. We note however these percentages include only the urban SOA and not the background OA, which is likely also SOA."

R2.8. Although the authors do clearly distinguish between SV-OOA and LV-OOA, in general, it could be useful if they take opportunities to remind readers more often that the SOA being investigated here excludes background OA, a fraction of which is likely SOA.

We have added several reminders of this point to the text. In particular, these reminders have been added to the 1st paragraph of Section 3.1.1., the first paragraph of Section 3.1.2., and the figure caption of Figure 6.

R2.9. Fossil vs. Modern: 1. The fossil/modern carbon analysis was determined for samples collected over only 7 days, a small subset of the overall campaign. The authors should note whether this week was generally representative of the overall campaign. 2. The authors might point out more strongly that the fossil/modern split determined here is linked to the assumptions regarding the assumed I/SVOC volatility distributions and emissions of these same compounds between sources.

In accordance with the reviewer's first comment we have added the following text to Section 2.4.

"In particular, the dates that the filters were collected were 30 May as well as 3, 4, 5, 6, 13, and 14 June 2010. Thus, these filters are more representative of the second half of the campaign that was more strongly influenced by pollution from the basin, compared to the first half of the campaign were regional advection played a more important role (Ryerson et al., 2013). Given the cost of the ^{14}C analyses, these days were chosen on the basis of the larger urban influence determined from the real-time measurements, and are therefore better suited to constrain urban sources (the subject of this paper) than if the analyses had been performed on filters from randomly-chosen days. However, it is noted that the relative concentrations of the different components of the OA were similar when averaging the second half of the campaign or the entire campaign: 14% vs. 12% for HOA, 5% vs. 5% for LOA, 12% vs. 17% for COA, 28% vs. 34% for LV-OOA, 40% vs. 34% for SV-OOA. Thus, it appears reasonable to assume that the relative results from the ^{14}C analysis are representative of the entire campaign. "

With respect to the second comment, we have added the following discussion to Section 3.1.2.

"It should be noted that the fossil/modern split from the box model that is described above depends on the initial P-S/IVOCs concentrations and volatility distribution assumed in the model. These parameters are not well constrained for cooking emissions, as discussed in further detail in Section 3.1.4 below. In addition, as discussed in the previous section (3.1.1) the concentration of primary IVOCs in the box model is higher than that measured. Thus, as an extreme sensitivity study, the model variations were also run under the assumption that cooking sources did not emit any P-S/IVOCs or, in the case of the PYE+TSI variation, any SVOCs (Figure SI-5). In this sensitivity study there is improved model/measurement agreement for primary IVOCs as shown in Table SI-3."

R2.10. SOA Apportionment: P32354/L14: I find the point associated with the mention of these PMF results from filters to be somewhat unclear. Are the authors simply saying that someone else said that diesel might contribute something?

Yes, that is correct. We have modified the text to make this point clearer (see quoted text below.)

"In addition, the existence of a diesel contribution in the model is consistent with PMF analysis of FTIR spectra of OA filter samples collected in Pasadena, in which, one SOA component exhibited relative peak intensities in the C-H stretching region that suggest some contribution from diesel emissions (Guzman-Morales et al., 2014), although the percentage of SOA from diesel could not be determined in this previous work."

R2.11. Evolution for 3 days: P32356/L12: I can see the concern raised regarding potential overestimates of OA downwind from urban regions in models. However, I think it is equally important to note that this is intimately tied to the model formulation that is used. Many models use very simple parameterizations that will not overpredict (necessarily), potentially even underpredicting (as shown with the WRF-CMAQ model). I think that the model dependency of this conclusion needs to be emphasized to a greater extent.

We have updated the text in this section (3.1.5) to more clearly state that this finding only applies to the parameterizations discussed in this section. The revised text reads:

"We also note that all of the parameterizations used in this section produce SOA/ΔCO ratios substantially larger (by factors of 2 or more) than those observed globally for aged air masses (i.e. photochemical ages greater than one day at an average OH concentration of 1.5×10^{6} molec cm^{-3})."

R2.12. WRF-CMAQ box modeling: I think that the comparison here can go even further than what is already noted by the authors. The non-ageing VBS treatment of VOCs in the primary box model is essentially equivalent to the 2-product model in CMAQ. The only difference is really that there are four products instead of two. Yet in the primary box model if only the VOCs are considered the underprediction is not a factor of 25, as is seen when using CMAQ in box model form. It is more like a factor of 5 or 6. While still substantial, I think that this also indicates that there is a fundamental difference in the basic model parameters used in the CMAQ 2 product formulation vs. the VBS 4-product formulation. In other words, toluene in CMAQ does not equal toluene in VBS. The authors are encouraged to emphasize this upon revision. They may wish to

refer to (Barsanti et al., 2013) who discuss issues associated with refitting data to determine 2 product parameters.

The statement by the reviewer about the two parameterizations being "essentially equivalent" is incorrect. The two parameterizations of V-SOA are conceptually equivalent but produce substantially different results. As stated in P32357, L17-24, for this comparison we use the older V-SOA yields from Koo et al. (2003), which are significantly lower than those of the Tsimpidi et al. (2010) parameterization used in the rest of the paper. The latter parameterization was updated using more recent chamber results such as those of Ng et al. (2007). We use the Koo et al. (2003) parameterization when comparing the results of the box model to WRF-CMAQ because the parameterization in CMAQ is of the same vintage, and thus we can show that both models produce very similar results when using similar parameterizations. We have added the following text at the end of P32357 L24 to clarify this point:

"Note that the Koo et al. (2003) parameterization produces substantially lower V-SOA yields than the Tsimpidi et al. (2010) parameterization used in the rest of this work. The latter parameterization was updated using more recent chamber results (with higher yields) such as those of Ng et al. (2007). The use of the older Koo et al. (2003) parameterization for this specific comparison was motivated by the fact that the parameterization for urban SOA in the version of CMAQ used here is based on the same older data, and does not contain the higher updated yields in, for example, Tsimpidi et al. (2010)."

R2.13. Oxygen content comparison: The literature oxygenation (i.e. O/C) parameterizations used for S/IVOCs are particularly underconstrained. I suggest that the authors take care to note that, given these inherent uncertainties in the O/C model formulation, the results here should not necessarily be taken as an indication that one model performs better than another.

We certainly agree with the reviewer, and have added the following paragraph in Section 3.2 of the manuscript in order to highlight this important point.

"Lastly, we note that there are large uncertainties in the parameters used to predict O:C such as the distribution of O:C values as a function of volatility for V-SOA or the amount of oxygen mass added to the SI-SOA for each oxidation reaction. It is therefore not very meaningful to conclude from this study that one parameterization performs better than another. Rather, it is apparent that when using several different previously published SOA parameterizations (i.e. ROB+TSI, GRI+TSI, and TSI alone) it is possible to reproduce the observed O:C at the Pasadena ground site."

R2.14. The updated SOA budget: I suggest that the authors focus this discussion on the anthropogenic SOA budget, with less consideration given to the BVOC SOA budget, given that the (non-cooking related) urban SOA here is predominately anthropogenic in origin.

We agree with the reviewer that the BVOC SOA budget is highly uncertain and unrelated to the new constraints derived in this paper, and we have thus removed all references to this topic from the text.

R2.15. Conclusions: Given the differences between the different parameterizations regarding the relative contributions from S/IVOCs versus VOCs towards the urban SOA, I suggest that the authors further emphasize that the relative concentrations remain quite uncertain.

We agree with the reviewer and had already included a statement to this effect in the conclusions. However, we have added an additional sentence to the conclusions to emphasize the uncertainty in the relative contributions from each category of precursor. The updated text from the conclusions is quoted below.

"The relative importance of VOCs and P-S/IVOCs as contributors to urban SOA over different time and length scales remains unclear. Depending on the parameterization used in the box model, the amount of urban SOA from VOCs can range between 15 – 53% of the total predicted SOA for the Pasadena ground site. This range is 16 - 58% in a sensitivity study in which the IVOC emissions are reduced by one-half."

R2.16. Figures: Figure 1 is fantastic and is very helpful to the reader to understand how the model formulations work.

We are very happy to hear that the referee found Figure 1 to be helpful.

References.

Barsanti, K. C.; Carlton, A. G.; Chung, S. H. Analyzing experimental data and model parameters: implications for predictions of SOA using chemical transport models. Atmos. Chem. Phys., 13, 12073-12088, 2013.

Koo, B. Y.; Ansari, A. S.; Pandis, S. N. Integrated approaches to modeling the organic and inorganic atmospheric aerosol components. Atmos. Environ., 37, 4757-4768, 2003.

Ng, N. L.; Kroll, J. H.; Chan, A. W. H.; Chhabra, P. S.; Flagan, R. C.; Seinfeld, J. H. Secondary organic aerosol formation from m-xylene, toluene, and benzene. Atmos. Chem. Phys., 7, 3909–3922, 2007.

Tsimpidi, A. P.; Karydis, V. A.; Zavala, M.; Lei, W.; Molina, L.; Ulbrich, I.M.; Jimenez, J. L.; Pandis, S. N. Evaluation of the volatility basis-set approach for the simulation of organic aerosol formation in the Mexico City metropolitan area. Atmos. Chem. Phys., 10, 525-546, 2010.

Modeling the formation and aging of secondary organic aerosols in Los Angeles during CalNex 2010

Patrick L. Hayes[1,2,3], Annmarie G. Carlton[4], Kirk R. Baker[5], Ravan Ahmadov[1,6], Rebecca A. Washenfelder[1,6], Sergio Alvarez[7], Bernhard Rappenglück[7], Jessica B. Gilman[1,6], William C. Kuster[6], Joost A. de Gouw[1,6], Peter Zotter[8], Andre S. H. Prévôt[8], Sönke Szidat[9], Tadeusz E. Kleindienst[5], John H. Offenberg[5], Prettiny K. Ma[3], Jose L. Jimenez[1,2]

(1) Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado, Boulder, CO, USA.

(2) Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO, USA.

(3) Université de Montréal, Department of Chemistry, Montreal, QC, CANADA

(4) Department of Environmental Sciences, Rutgers University, New Brunswick, NJ, USA.

(5) US Environmental Protection Agency, Research Triangle Park, NC, USA.

(6) Earth System Research Laboratory, National Oceanic and Atmospheric Administration (NOAA), Boulder, CO, USA.

(7) Department of Earth and Atmospheric Sciences, University of Houston, TX, USA.

(8) Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, Villigen, Switzerland.

(9) Department of Chemistry and Biochemistry & Oeschger Centre for Climate Change Research, University of Bern, Switzerland.

Correspondence to:

P. L. Hayes (patrick.hayes@umontreal.ca) and J. L. Jimenez (jose.jimenez@colorado.edu)

Abstract

Four different literature parameterizations for the formation and evolution of urban secondary organic aerosol (SOA) frequently used in 3D models are evaluated using a 0-D box model representing the Los Angeles Metropolitan Region during the CalNex 2010 field campaign. We constrain the model predictions with measurements from several platforms and compare predictions with particle and gas-phase observations from the CalNex Pasadena ground site. That site provides a unique opportunity to study aerosol formation close to anthropogenic emission sources with limited recirculation. The model SOA formed only from the oxidation of VOCs (V-SOA) is insufficient to explain the observed SOA concentrations, even when using SOA parameterizations with multi-generation oxidation that produce much higher yields than have been observed in chamber experiments, or when increasing yields to their upper limit estimates accounting for recently reported losses of vapors to chamber walls. The Community Multiscale Air Quality (WRF-CMAQ) model (version 5.0.1) provides excellent predictions of secondary inorganic particle species but underestimates the observed SOA mass by a factor of 25 when an older VOC-only parameterization is used, which is consistent with many previous model-measurement comparisons for pre-2007 anthropogenic SOA modules in urban areas.

Including SOA from primary semi-volatile and intermediate volatility organic compounds (P-S/IVOCs) following the parameterizations of Robinson et al. (2007), Grieshop et al. (2009), or Pye and Seinfeld (2010) improves model/measurement agreement for mass concentration. The results from the 3 parameterizations show large differences (e.g. a factor of 3 in SOA mass) and are not well constrained, underscoring the current uncertainties in this area. ~~When comparing the three parameterizations, the Grieshop et al. parameterization more accurately reproduces both the SOA mass concentration and oxygen-to-carbon ratio inside the urban area.~~ Our results strongly suggest that other precursors besides VOCs, such as P-S/IVOCs, are needed to explain the observed SOA concentrations in Pasadena. All the recent parameterizations over-predict urban SOA formation at long photochemical ages (≈3 days) compared to observations from multiple sites, which can lead to problems in regional and especially global modeling. However, reducing IVOC emissions by one-half in the model to better match recent IVOC measurements improves SOA predictions at these long photochemical ages.

Among the explicitly modeled VOCs, the precursor compounds that contribute the greatest SOA mass are methylbenzenes. Measured Polycyclic aromatic hydrocarbons (naphthalenes) contribute 0.7% of the modeled SOA mass. The amounts of SOA mass from diesel vehicles, gasoline vehicles, and cooking emissions are estimated to be 16 – 27%, 35 – 61%, and 19 – 35%, respectively, depending on the parameterization used, which is consistent with the observed fossil fraction of urban SOA, 71(±3)%. The relative contribution of each source is uncertain by almost a factor of 2 depending on the parameterization used. In-basin biogenic VOCs are predicted to contribute only a few percent to SOA. A regional SOA background of approximately 2.1 μg m^{-3} is also present due to the long distance transport of highly aged OA, likely with a substantial contribution from regional biogenic SOA. The percentage of SOA from diesel vehicle emissions is the same, within the estimated uncertainty, as reported in previous work that analyzed the weekly cycles in OA concentrations (Bahreini et al., 2012; Hayes et al., 2013). However, the modeling work presented here suggests a strong anthropogenic source of modern carbon in SOA, due to cooking emissions, which was not accounted for in those previous studies, and which is higher on weekends.

Lastly, this work adapts a simple two-parameter model to predict SOA concentration and O/C from urban emissions. This model successfully predicts SOA concentration, and the optimal parameter combination is very similar to that found for Mexico City. This approach provides a computationally inexpensive method for predicting urban SOA in global and climate models. We estimate pollution SOA to account for 26 Tg yr^{-1} of SOA globally, or 17% of global SOA, 1/3 of which is likely to be non-fossil.

1. Introduction

Submicron aerosols impact regional to global climate (IPCC, 2013), visibility (Watson, 2002), and human health (Dockery and Pope, 1994). Quantification of the environmental and health impacts of atmospheric aerosols is difficult however, because of our incomplete understanding of aerosol physical and chemical properties. Atmospheric aerosols are typically a mixture of organic and inorganic matter, and the organic fraction is normally composed of hundreds or even thousands of compounds. Due to this complexity, accurate prediction of OA concentrations, as well as chemical properties is challenging (McKeen et al., 2007; Heald et al., 2011; Spracklen et al., 2011). This problem is especially important given that OA represents roughly half of the total tropospheric submicron aerosol mass in many environments including polluted urban regions (Murphy et al., 2006; Jimenez et al., 2009).

Given its complexity, OA is often categorized based on sources. Primary organic aerosols (POA) are emitted directly into the atmosphere from sources such as motor vehicles, food cooking, and wildfires. SOA is formed in the atmosphere by photooxidation and/or heterogeneous or cloud processing of gas-phase precursors. The gas-phase precursors for SOA potentially have many sources including vehicle emissions, the biosphere, biomass burning, and food cooking (e.g. Schauer et al., 1999; Hallquist et al., 2009; Hodzic et al., 2010b; Bahreini et al., 2012). A large portion of the submicron OA throughout the world can be classified as SOA (Zhang et al., 2007; Jimenez et al., 2009). Even in urban areas such as the Los Angeles Metropolitan Area, SOA is often found to be larger than POA, especially in the summer (Docherty et al., 2008; Hersey et al., 2011; Hayes et al., 2013).

Traditional models for SOA formation use a semi-empirical approach wherein SOA formation is described in two steps: the gas-phase oxidation of VOC precursors resulting in the formation of semi-volatile organic compounds (SVOCs), followed by partitioning of the SVOCs to the particle phase. The parameters for these models (yields, saturation concentrations, etc.) are typically derived from smog chamber experiments on individual VOCs (Hallquist et al., 2009). Since about 2005, it has been shown in multiple publications from several field studies that traditional models under-predict observed SOA in urban areas by a large amount with a difference of up to a factor of 19. (Volkamer et al., 2006; de Gouw and Jimenez, 2009; Dzepina et al., 2009; Hodzic et al., 2010a). A similarly large underestimate is typically not observed in

areas dominated by biogenic SOA (Tunved et al., 2006; Chen et al., 2009; Hodzic et al., 2009; Slowik et al., 2010). In response, new precursors and pathways for SOA formation have been identified from measurements and incorporated into SOA models. The new formation pathways include SOA formation from primary semivolatile and intermediate volatility organic compounds (P-S/IVOCs) (Robinson et al., 2007), aqueous phase production in clouds (e.g. Lim et al., 2005) and aerosols (Ervens and Volkamer, 2010; Knote et al., 2014b), as well as the oxidation of VOCs such as isoprene, benzene, and acetylene that were previously thought to produce little or no SOA (Martin-Reviejo and Wirtz, 2005; Kroll et al., 2006; Volkamer et al., 2009).

The introduction of the volatility basis set (VBS) approach represents a conceptual advance for modeling OA (Donahue et al., 2006). This approach distributes organic species into logarithmically spaced volatility bins, which are used to calculate absorptive partitioning between the gas and particle-phases. Mass is transferred between the bins as photochemical oxidation proceeds and environmental parameters (i.e. temperature, dilution) change. The VBS approach has been applied to SOA from biogenic and anthropogenic VOCs as well as to P-S/IVOCs and the SOA formed from them (Robinson et al., 2007; Tsimpidi et al., 2010).

Although these updates have led to substantial reductions in the gaps between observed and predicted OA concentrations, major inconsistencies and uncertainties remain, and it is not clear that improved agreement is achieved for the right reasons. For instance, both Dzepina et al. (2011) and Hodzic et al. (2010a) reported that the Robinson et al. (2007) parameterization for the production of SOA from P-S/IVOCs contributed substantially to successful predictions of SOA concentration in a box and a regional model for the Mexico City region, but the predicted O/C values were approximately a factor of 2 too low. A different parameterization of SOA from P-S/IVOCs published by Grieshop et al. (2009) led to overpredicted total SOA concentration, but successfully reproduced the measured O/C values.

Complicating the picture further was the additional finding in Dzepina et al. (2011) that if the VBS with multi-generational aging was applied to VOCs following Tsimpidi et al. (2010), then all the SOA mass could be successfully predicted without considering P-S/IVOCs. A similar finding was observed in Tsimpidi et al. (2010) wherein the inclusion of P-S/IVOCs and an "aging VBS" treatment of VOC oxidation worsened over-prediction in the model during the

afternoon. Thus, the relative importance of P-S/IVOCs versus VOCs in urban SOA production remains very uncertain. More generally, robust model/measurement closure – in which SOA chemistry is accurately represented – is an important step towards implementing effective particulate matter pollution controls in urban areas.

Here we compare the results of a constrained SOA box model against measurements carried out at the Pasadena ground site during the California Research at the Nexus of Air Quality and Climate Change (CalNex) campaign. The use of a box model allows multiple state-of-the-art parameterizations to be tested. Once constrained by measurements, the box model facilitates the improved source apportionment of SOA in the Los Angeles Metropolitan Area. In particular, the amount of SOA formed from different precursors is quantitatively evaluated. The importance of diesel versus gasoline emissions as sources of SOA precursors – a topic that has received much recent interest – is discussed as well (Bahreini et al., 2012; Gentner et al., 2012; Hayes et al., 2013; Ensberg et al., 2014). Results are also compared to those of the 3-D WRF-CMAQ model. The CalNex field campaign, which took place in Spring/Summer 2010, provides a unique data set for evaluating gas-phase SOA models because of, in part, the large scope of the campaign, and the generally clear-sky conditions during the campaign that limited the effects of cloud chemistry. Specifically at the Pasadena ground site, which operated from May 15 2010 to June 15 2010, there were over 70 gas and particle phase measurements including cutting-edge techniques that provide new insights into SOA sources and chemistry. For example, highly time resolved ^{14}C measurements with 3 – 4 h resolution are utilized in this work, whereas typically 12 h or lower resolution has been reported (Zotter et al., 2014). By comparing the CalNex dataset to recently proposed SOA models, the research described below aims to evaluate recently proposed SOA models and assess the importance of different SOA sources and formation pathways.

2. Modeling methods

2.1. Pasadena ground site meteorology

An overview of the CalNex study has been recently published by Ryerson et al. (2013). The location and meteorology of the Pasadena ground site has been described in detail previously (Washenfelder et al., 2011; Hayes et al., 2013). Briefly, the site was located in the Caltech campus about 18 km northeast of downtown Los Angeles (34.1406 N, 118.1225 W).

Pasadena lies within the South Coast Air Basin (SoCAB) and the Los Angeles metropolitan area. The prevailing wind direction during daytime in Pasadena was from the southwest, which brought air masses from the Santa Monica and San Pedro bays through Los Angeles to Pasadena. Thus, Pasadena during the daytime is predominately a receptor site for pollution emitted in the western Los Angeles metropolitan area that is then advected over a period of several hours (about 3 – 5 h). While more local emissions and background concentrations of atmospheric species must influence the site, the diurnal cycles of many primary species with anthropogenic sources (e.g. CO, black carbon (BC), and benzene) appear to be dominated by advection of pollution from the southwest. Specifically, CO, BC, and benzene concentrations display strong peaks around noontime as shown in Figure 2 of Hayes et al. (2013), which is due to a transport time of several hours until the emissions from the morning rush hour arrive in Pasadena. At nighttime, winds were weak and were most frequently from the southwest or southeast, which is illustrated in the supporting information (Figure A-2) of Hayes et al. (2013). The site was influenced at that time by more local emissions than by advection from downtown Los Angeles. Aged emissions from the prior daytime may have influenced the site as well during nighttime.

2.2. SOA box model

The models in this work are summarized in Table 2. The box model used here accounts for SOA formed from gas-phase oxidation of two sets of precursors: (1) VOCs, and (2) P-S/IVOCs. Also included in the total model SOA is background SOA (BG-SOA), with a constant concentration of 2.1 μg m^{-3} that is derived from observations as described later in Section 2.4. BG-SOA is considered non-volatile in the model, which is consistent with observations that very aged SOA has low volatility (Cappa and Jimenez, 2010). For the remainder of the SOA the equilibrium partitioning between the particle and gas-phases is calculated using the reformulation of Pankow Theory by Donahue et al. (2006). The particle-phase fraction of species i, ξ_i, is calculated using its effective saturation concentration, C_i*, and the total concentration of the organic material available for partitioning, *[OA]*.

$$\xi_i = \left(1 + \frac{C_i^*}{[OA]}\right)^{-1} \; ; \; [OA] = \sum_i [SVOC]_i \, \xi_i \qquad (1)$$

We note that there is ongoing scientific research examining if OA adopts a liquid or solid/glassy phase with potentially slow diffusion properties, and the conditions that result in equilibrium or kinetically-limited partitioning are not yet clear (e.g. Cappa and Wilson, 2011; Perraud et al., 2012). For the purpose of this study however, field measurements from CalNex strongly suggest that organic aerosols undergo equilibrium partitioning in Pasadena (Zhang et al., 2012). In particular, for water-soluble organic carbon, a surrogate for SOA, the partitioning coefficient was observed to be correlated with the OA mass. A similar observation was made at a rural site in Colorado, USA, and the lack of kinetic limitations to equilibrium may be attributable to the higher ambient relative humidity, mostly greater than 30%, in both Pasadena and Colorado compared to some studies that have reported kinetic limitations (Yatavelli et al., 2014). Furthermore, we note that the diurnally averaged relative humidity in Pasadena was always greater than 60%, which laboratory studies have suggested is above the ~30% threshold where~~ich~~ particles form liquid phases (Renbaum-Wolff et al., 2013).

V-SOA in the box model includes products from the oxidation of 46 VOCs, and the V-SOA mass is distributed into a 4-bin VBS as shown Figure 1 (C^*=1, 10, 100, or 1000 μg m^{-3}). Furthermore, a table with the names of each VOC as well as the relevant model parameters is provided in the supporting information (Table SI-1). The reaction rates for most of the VOCs are taken from Atkinson and Arey (2003) and, when not available there, Carter (2010). Three terpene compounds (α-pinene, β-pinene, and limonene) were lumped for this model, and the rate constant of this lumped precursor species is the weighted average – by ambient concentrations – of the individual rate constants (Atkinson and Arey, 2003). In addition, the rates for naphthalene, 1-methylnaphthalene, and 2-methylnaphthalene oxidation are taken from Chan et al. (2009). The SOA yields for the VOCs are taken from Tsimpidi et al. (2010). For naphthalene and the methylnaphthalenes the yields are from data presented in Chan et al. (2009), which have been re-fitted to obtain yields for the 4-bin VBS utilized in this work. V-SOA is also allowed to 'age' after the initial reaction, and the subsequent gas-phase oxidation (with a rate constant of 10^{-11} cm^3 molec.$^{-1}$ s^{-1}, which was erroneously reported as 4 times higher in Tsimpidi et al. (2010))

leads to a 10× decrease in volatility as well as a 7.5% increase in mass due to added oxygen for each generation. This parameterization for V-SOA is abbreviated as "TSI" in the text.

It is possible that the SOA yields used for V-SOA, which are based on the chamber experiment literature, are several-fold too low due to, for example, losses of gas-phase species to chamber walls (Matsunaga and Ziemann, 2010; Zhang et al., 2014). To investigate this possibility a model variation – named "~~ROB+~~4xV" – is run wherein the SOA yields from aromatics are increased by a factor of four, based on recent chamber studies in which higher concentrations of aerosol seed were utilized in order to suppress losses to chamber walls, and an upper limit of a factor of 4 increase in V-SOA yields was estimated (Zhang et al., 2014). The multi-generation aging of secondary species produced from VOCs is turned off in this variation, since otherwise the SOA yields would reach extremely unrealistic levels (~400%).

SOA from P-S/IVOCs (SI-SOA) is simulated utilizing three different parameter sets. No duplication of precursors is expected between the Tsimpidi et al. (2010) parameterization and the three P-S/IVOCs parameterizations, with the possible exception of the naphthalenes (Robinson et al., 2007; Dzepina et al., 2009; Dzepina et al., 2011). However, since the naphthalenes contribute a very small amount to the total SOA mass (see below), the impact of double-counting their SOA contribution is negligible. The first two P-S/IVOCs parameterizations are from Robinson et al. (2007), hereinafter "ROB", and an alternate set published by Grieshop et al. (2009), hereinafter "GRI". The differences between the two parameterizations are highlighted in Figure 1. When compared to ROB, primary and secondary species in GRI have a lower gas-phase reactivity (2×10^{-11} versus 4×10^{-11} cm^3 molec^{-1} s^{-1}), a larger decrease in volatility per oxidation step (two orders of magnitude versus one), and more oxygen mass added to the products (40% versus 7.5% of the precursor mass). Furthermore, there are differences in the assumed enthalpies of vaporization, ΔH_{vap}, and molecular weights. Details of both parameterizations are given in Table SI-2 in the supporting information.

The third parameterization utilized for SI-SOA is that published by Pye and Seinfeld (2010), hereinafter "PYE", which is also illustrated in Figure 1. In PYE the SOA from primary SVOCs and primary IVOCs follow different treatments. The primary SVOCs emitted are represented by two lumped species with C^*=20 and 1646 μg m^{-3} and relative concentrations of 0.51 and 0.49, respectively. The gas phase reactivity (2×10^{-11} cm^3 molec^{-1} s^{-1}) and decrease in

volatility per oxidation step (two orders of magnitude) are identical to GRI. However, only one oxidation step is allowed in PYE. The oxygen mass added to the products is 50% of the precursor mass, which is higher than that for ROB and GRI. Another difference in PYE is the enthalpy of vaporization for all organic species, which is 42 kJ/mol. Lastly, the molecular weight utilized here is 250 g mol^{-1}, the same as ROB, although this parameter is not specified in Pye and Seinfeld (2010). In PYE also the concentration of SOA from primary IVOCs is estimated by scaling-up the concentration of SOA from naphthalene by a factor of 66.

Heterogeneous uptake of glyoxal onto aerosols can be a relevant source of SOA under some conditions (Volkamer et al., 2007; Dzepina et al., 2009). Previously published work on the glyoxal budget for CalNex indicates that this compound contributes only a small fraction of the SOA mass in the LA basin, however (Washenfelder et al., 2011; Knote et al., 2014b), and we do not consider it further in this study. In Pasadena, the urban SOA peaked in the afternoons, which were generally clear and sunny during the campaign. This observation is consistent with the conclusion that reactions occurring in clouds did not play a major role in SOA production during CalNex. In addition, a comparison of OA/ΔCO for three days that were cloudy against the remainder of the campaign shows no apparent difference in the magnitude of the ratio or its evolution with photochemical age (Figure SI-1), which further supports the conclusion that SOA production from clouds can be neglected in this study.

The design of the model used here includes several more elements that are general for V-SOA and SI-SOA. Only oxidation by hydroxyl radical (•OH) is considered since in urban regions other oxidants such as ozone, nitrate radical, and chlorine radical are expected to be minor contributors to SOA formation from urban VOCs (Dzepina et al., 2009; Dzepina et al., 2011; Hayes et al., 2013). Additionally, the model is run using "high-NO_x conditions," which is consistent with previously calculated branching ratios for the RO_2 + NO, RO_2 + HO_2, and RO_2 + RO_2 reactions (Hayes et al., 2013) and the dominance of the RO_2 + NO pathway. The primary and secondary species are assumed to mix into a single organic phase. This assumption is based on observations made off the coast of California that SOA condenses on primary particles (e.g., BC and POA) as indicated by the similar size distributions for these species across a range of photochemical ages (Cappa et al., 2012). In addition, the organic phase is taken to be separate from the inorganic phases, which is consistent with the relatively low O:C values observed

during CalNex (Hayes et al., 2013) and previous studies demonstrating that organic/inorganic phase separation occurs when O:C is less than 0.7 (Bertram et al., 2011). It should be noted that this statement holds true even after applying the updated calibration for AMS O:C (Canagaratna et al., 2015).

The temperature dependence of C^* is calculated with the Clausius-Clapeyron equation.

$$C_i^* = C_{i,o}^* \frac{T_0}{T} \exp\left[\frac{\Delta H_{vap}}{R}\left(\frac{1}{T_o} - \frac{1}{T}\right)\right] \quad (2)$$

Where $C_{i,o}^*$ is the effective saturation concentration of condensable compound i at the reference temperature T_o (K), and R is the ideal gas constant. The ambient temperature, T, was taken to be 18°C, which represents the average campaign temperature during CalNex. A sensitivity test exhibited less than a 4% change in predicted mass at a given time-of-day when using 14°C and 24°C, which are the minimum and maximum temperatures for the diurnal cycle. The error in predicted mass over this temperature range is small compared to other uncertainties in SOA modeling, and therefore the use of a constant temperature of 18°C to calculate C^* should introduce negligible errors.

2.3. Model set-up

This work utilizes a box approach wherein the model calculates the evolution of organic species in an air parcel as it undergoes photochemical aging. A schematic of the model set-up is shown in Figure 2. The calculation is run 24 times to predict the average diurnal cycle for the entire campaign (15 May – 15 June). For each of the 24 repetitions, the calculation always starts at hour zero and then runs to 12 h of photochemical aging (Panel 2). Next, the model output at the same photochemical age as that observed at the Pasadena ground site for the given time-of-day is saved for comparison against measurements (Panel 3). The initial concentrations of VOCs in the air parcel are calculated by multiplying the background-subtracted CO concentrations measured at Pasadena by the emission ratios, ΔVOC/ΔCO, previously determined for CalNex, which are consistent with those for other US urban areas (Warneke et al., 2007; Borbon et al.,

2013) (Panel 1). CO is an inert tracer of combustion emissions over these timescales and its formation from VOCs is very minor as well (Griffin et al., 2007). The CO background level represents the amount present from continental-scale transport and for which the co-emitted organic species have been lost by deposition (e.g. DeCarlo et al., 2010). The background was determined by examining CO measurements taken aboard the NOAA WP-3D aircraft off the Los Angeles coastline at altitudes less than 200 m as described in our previous paper (Hayes et al., 2013). Given that the model is set-up to predict the mean diurnal cycle of SOA during the entire CalNex-Pasadena measurement period, the mean diurnal cycle of the CO concentration is used for the calculation of the emissions. An important advantage of using CO as a conserved urban emissions tracer is that dilution of emissions in the air parcel is implicitly included in the model, since the reductions in CO concentration will lead to lower calculated initial precursor concentrations in that air parcel.

The biogenic VOCs are not expected to be emitted proportionally with CO, and therefore the approach described in the previous paragraph cannot be used to specify the biogenic VOC emissions. Rather, t~~T~~he emissions of biogenic VOCs were adjusted empirically to match the observed concentrations of isoprene and terpenes, after accounting for anthropogenic isoprene using Δ(isoprene)/ΔCO (Borbon et al., 2013). Only ~4% of the daily average isoprene is from anthropogenic sources. In addition, the diurnal profile of emissions was assumed to be proportional to ambient temperature.

The model consistency with the VOC measurements, including for biogenic VOCs, is evaluated by comparing the measured and modeled diurnal cycles. Some of t~~T~~he cycles compared are given in Figure SI-2 as an example. It is observed that the model is generally consistent with the VOC measurements.

For naphthalene and its analogs, emission ratios are not available in the literature, to our knowledge. To obtain the emission ratios the concentrations of the polycyclic aromatic hydrocarbons were plotted versus CO, and a linear orthogonal distance regression (ODR) analysis was carried out. The data were filtered and include only periods from 00:00 – 06:00 (local time) to minimize depletion by photochemical processing (Figure SI-3). The slope from the regression analysis was then used as the emission ratio. However, as observed in Figure SI-3, the diurnal cycles for naphthalene and its analogs are not well-reproduced by the model during

the daytime when using the early morning emission ratios. The sampling of these compounds was performed on a tar roof, and it is possible that the local concentrations in the vicinity of roof may be elevated during daytime due to volatilization of the roofing tar and not representative of concentrations throughout the Los Angeles basin. The naphthalene and methylnaphthalene concentrations are well correlated with temperature. However, it is also possible that the volatilization occurs over a larger city scale, and thus a variation of the model is run wherein the emission ratios are changed empirically along the diurnal cycle so that the model reproduces the measured diurnal cycle for each speciated naphthalene (Figure SI-3). The increases in emissions range between 1 and 3.5 times the original value, and the implications for SOA are discussed in Section 3.1.3.

The calculation of the initial P-S/IVOC concentrations requires a somewhat different procedure compared to the VOCs. Instead, the amount of initially emitted POA is calculated from measured ΔPOA/ΔCO ratios and the measured CO concentration in Pasadena. Then the total concentration of P-S/IVOCs is set so that the particle-phase P-S/IVOC concentration matches the amount of initially emitted POA, while constraining the volatility distribution to that of the corresponding parameterization, as done in previous studies (e.g. Dzepina et al., 2009).

The model consistency with respect to the POA measurement is shown in Figure SI-2. The comparison for POA is adequate, and a linear ODR analysis yields a slope of 1.01 (R = 0.76) when the GRI+TSI parameterization is used. Of these three model variants, PYE+TSI shows a larger positive bias. This is likely due to the relatively large amount of primary SVOCs placed in the C*=20 bin compared to ROB+TSI and GRI+TSI, which will result in more partitioning to the particulate phase as the total OA mass is increased (e.g. by SOA formation)

The initial VOCs and P-S/IVOCs are then oxidized in the air parcel. The aging of the air parcel is simulated separately 24 times with each simulation using measured parameters (e.g. ΔCO, photochemical age, POA) corresponding to one hour during the mean diurnal cycle. Following Dzepina et al. (2009) the evolution of the different compounds in each of the 24 aging simulations is calculated by discretizing the rate equations using Euler's method.

The photochemical age of the urban emissions at each time of day is determined from the ratio of 1,2,4-trimethylbenzene to benzene as described previously (Parrish et al., 2007; Hayes et al., 2013). We note that the photochemical age estimated from NO_y/NO_x is very similar (Hayes

et al., 2013), which is consistent with previous results from Mexico City for ages shorter than 1 day (C. A. Cantrell, Univ. of Colorado, personal communication, 2014). There are three important considerations that must be evaluated when using VOC concentration ratios as photochemical clocks.

First, trimethylbenzene and benzene are predominately from anthropogenic sources, and thus the photochemical clock only applies to the evolution of anthropogenic emissions. Previous work by Washenfelder et al. (2011) estimated that most biogenic VOCs were emitted mostly in the last quarter of the trajectory of the air parcel arriving at Pasadena at 16:00 PDT. This estimate was based on the vegetation coverage observed in visible satellite images of the upwind areas, as well as on the ratio of isoprene to its first-generation products (methyl vinyl ketone and methacrolein). However, in this work, the photochemical age for biogenic VOCs is kept the same as for the anthropogenic VOCs. This approach will overestimate the amount of photochemical aging – and the SOA from in-basin biogenic emissions – during daytime. The modeled biogenic SOA should thus be considered an upper limit. ~~The emissions of biogenic VOCs were adjusted empirically to match the observed concentrations of isoprene and terpenes, after accounting for anthropogenic isoprene using Δ(isoprene)/ΔCO (Borbon et al., 2013). Only ~4% of the daily average isoprene is from anthropogenic sources.~~ As discussed below, the amount of SOA from in-basin biogenic VOCs is very small. Thus, our SOA model results are not sensitive to the details of how SOA from biogenic VOCs emitted within the LA basin is modeled. We do not include oxidation of biogenic VOCs by O_3 or NO_3 in the box model, but these oxidants have only a minor role in SOA formation during the daytime when the peak for in-basin SOA concentration is observed. In particular, given the measured concentrations of oxidants (Hayes et al., 2013), oxidation of isoprene and terpenes by •OH is 37 and 5 times faster on average, respectively, than oxidation by O_3 during daytime.

The second consideration is that the purpose of using the ratio of VOC concentrations is to determine the •OH exposure for the air mass at the Pasadena site. (•OH exposure is the concentration integrated over time for an air parcel.) While the •OH exposure for the site is therefore well-constrained, the actual •OH concentration in the modeled air parcel as a function of time is not as well-constrained. Thus, the photochemical ages used here (Figure 3) are calculated using an average •OH concentration of 1.5×10^{6} molec cm^{-3}, as described in our

previous work (Hayes et al., 2013), and~~.~~ ~~T~~the model is run with the same concentration.~~, which is necessary to match the model and observed OH exposure. (•OH exposure is the concentration integrated over time for an air parcel.)~~ Insofar as the model produces the same •OH exposure as determined from measurements, which is always the case in this modeling study, the actual concentration of •OH used in the model is not expected to substantially influence the results. In other words, while the concentration •OH in the model is assumed to be 1.5×10^{6} molec cm^{-3}, the integral of the •OH concentration over time is constrained by the observed VOCs ratios. ~~Thus~~As expected, in the middle of the day the photochemical age will be longer than the transport age, and the opposite will be true during periods with low ambient •OH.

Third, photochemical age is a quantity developed as a metric for parcels of air arriving at a remote receptor site, and it is derived by assuming that the parcel is decoupled from fresh emissions as it is transported (Kleinman et al., 2007; Parrish et al., 2007). However, Pasadena is not a remote receptor site, and it is impacted by pollution that has been emitted recently as well as transported from more distant locations. The error in the calculated photochemical age that results from the mixing of nearby and far sources is evaluated in our previous work, and it may lead to underestimation of the actual photochemical age by ~10% (Hayes et al., 2013), which is relatively minor compared to the uncertainty in the OA measurement of ±30% (Middlebrook et al., 2012) and the possible biases in the different SOA parameterizations.

2.4. Model/measurement comparisons

The model is compared against the average diurnal cycles of various OA properties (e.g. concentration, O:C). The measurements utilized in this study are summarized in Table 3. In previous work the concentrations of five different OA components were determined using positive matrix factorization (PMF) of aerosol mass spectrometer (AMS) data, and the diurnal cycles of these components are shown in Figure 3 (Hayes et al., 2013). Hydrocarbon-like organic aerosol (HOA) and cooking-influenced organic aerosol (CIOA) are both thought to be dominated by POA. As discussed in Hayes et al. (2013), HOA is dominated by vehicle combustion emissions, and the CIOA is dominated by cooking sources. However, for the purpose of running the SOA model, HOA and CIOA are not treated separately, and instead their summed mass

concentrations are used as the POA concentration. It should be noted however that the amount of SOA from HOA or CIOA associated P-S/IVOCs can still be calculated under certain assumptions as discussion in Section 3.1.2 below. Low volatility oxygenated organic aerosol (LV-OOA) is a surrogate for highly aged secondary organic aerosol, and it displays a flat diurnal profile. Furthermore, recent ^{14}C measurements show that this component is largely composed of non-fossil carbon (Zotter et al., 2014). Both of these observations indicate that LV-OOA is transported into the Los Angeles Basin (Hayes et al., 2013).

Results from 3-D WRF-Chem simulations were also used to evaluate the concentration of BG-SOA. These simulations determined the BG-SOA by removing all the emissions in the Los Angeles region as shown in Figure SI-4, and it was observed that there are both biogenic and anthropogenic emissions in California that contribute to the background OA. In addition, background marine OA is thought to be very low during the CalNex measurement period, since concentrations of OA were less than 0.2 μg m^{-3} over the open ocean west of California for regions with low pollution influence (P. K. Quinn, NOAA, personal communication, 2012). As shown in Figure 3B, the background SOA concentration from the WRF-Chem simulation is similar to the concentration of LV-OOA. Given these observations as well as the ^{14}C results discussed in the previous paragraph, we use the LV-OOA component to constrain the amount of BG-SOA, and specifically, set the amount of BG-SOA to be the minimum of LV-OOA observed in the diurnal cycle (2.1 μg m^{-3}). Heo et al. (2015) recently concluded that the background SOA in the LA basin has an important component from biogenic emissions over the Central Valley, which is consistent with our results.

In contrast, semi-volatile oxygenated organic aerosol (SV-OOA) displays a distinct diurnal profile that peaks at a similar time as photochemical age, which is consistent with this component being a proxy for freshly formed SOA from urban emissions. The ^{14}C measurements also indicate that SV-OOA is predominately, 71% (±3%), composed of fossil carbon. (Note: to obtain this percentage it is assumed that the OC/OM ratio is the same for fossil and non-fossil SV-OOA.) As described above, the box model designed here is specifically focused on SOA formation from precursors emitted within the Los Angeles basin, and the ^{14}C measurements and diurnal cycle strongly indicate that SV-OOA concentration is a better surrogate for total urban SOA than the total OOA concentration. Lastly, there is a fifth component displayed in Figure 3B,

local organic aerosol (LOA) of primary origin and of uncertain sources, but this component comprises only ~5% of the aerosol mass. It is thought to be emitted very close to the site based on its very rapid time variations, and thus any co-emitted VOCs or S/IVOCs would have very little time to react and form SOA. Therefore, LOA is not considered further in this modeling study.

In principle, the box model could be run for multiple individual days. However, some datasets and published results used in this study are not available with sufficient time resolution for such an approach. In particular, the thermal desorption gas chromatograph mass spectrometry analysis for naphthalenes required adsorbent tube~~filter~~ samples that were composited over several days. In addition, both the apportionment of the SV-OOA and LV-OOA components between fossil and non-fossil sources (Zotter et al., 2014) as well as the analysis of the diesel fraction of OOA (Hayes et al., 2013) required analyzing datasets from multiple days as a single ensemble. To facilitate incorporating these datasets and published results into this study, we have chosen to run the box model so that it simulates the average diurnal cycle during the campaign. The measurements used here (Table 3) all had excellent coverage during the CalNex campaign, with each instrument reporting data for more than 75% of the total campaign duration. Thus, the measurements are expected to be representative of conditions during the campaign. An exception is the ^{14}C measurements, which were carried out on filters collected over 7 days. This limited sampling period is due to the time and resource intensive nature of the ^{14}C measurements (Zotter et al., 2014). In particular, the dates that the filters were collected were 30 May as well as 3, 4, 5, 6, 13, and 14 June 2010. Thus, these filters are more representative of the second half of the campaign that was more strongly influenced by pollution from the basin, compared to the first half of the campaign were regional advection played a more important role (Ryerson et al., 2013). Given the cost of the ^{14}C analyses, these days were chosen on the basis of the larger urban influence determined from the real-time measurements and are therefore better suited to constrain urban sources (the subject of this paper) than if the analyses had been performed on filters from randomly-chosen days. However, it is noted that the relative concentrations of the different components of the OA were similar when averaging the second half of the campaign or the entire campaign: 14% vs. 12% for HOA, 5% vs. 5% for LOA, 12% vs. 17% for COA, 28% vs. 34% for LV-OOA, 40% vs. 34% for SV-OOA. Thus, it appears reasonable to assume that the relative results from the ^{14}C analysis are representative of the entire campaign.

2.5. Modeling the SOA oxygen content

To simulate the oxygen-to-carbon ratio (O:C) of total OA, the box model utilizes the measured O:C for HOA, CIOA, and LV-OOA. The O:C values for HOA and CIOA are assumed to be constant because heterogeneous aging of primary aerosols is relatively slow, and thus the O:C should only vary by a relatively small amount due to this mechanism over the timescales considered here (Donahue et al., 2013). LV-OOA is predominately composed of aged background OA, and thus its O:C should not vary substantially either. The oxygen and carbon mass from HOA, CIOA, and LV-OOA are then added to the oxygen and carbon mass predicted in the model for freshly formed SOA.

The O:C of V-SOA is simulated using a modified version of the approach described in Dzepina et al. (2009). In that previous work the O:C of V-SOA was estimated to be 0.37 and constant. While this estimate is consistent with chamber experiments of aromatic precursors, it is conceptually difficult to reconcile with V-SOA aging wherein successive oxidation reactions are expected to reduce volatility and increase O:C. It is therefore assumed in the box model that O:C increases as follows: C* = 1000 μg m^{-3}, O:C = 0.25; C* = 100 μg m^{-3}, O:C = 0.30, C* = 10 μg m^{-3}; O:C = 0.40; C* = 1 μg m^{-3}, O:C = 0.60. This O:C distribution is taken from the first-generation distribution of Murphy et al. (2011), and in that work the O:C was simulated in a full 2-D VBS and depends on both volatility bin as well as oxidation generation. For the purpose of this study an intermediate approach is used wherein O:C depends on volatility bin only, and the first-generation distribution of Murphy et al. (2011) is applied to all oxidation generations of SOA. We note that only a small amount of V-SOA mass is from multi-generation oxidation (10 – 20%) for the relevant model conditions used for Pasadena. Thus, the O:C values predicted here will not be substantially different from a full 2-D VBS treatment.

The O:C ratio for SI-SOA is simulated following the approach described in Robinson et al. (2007). Conceptually, with each oxidation step the model adds 1 oxygen atom per 15 carbon atoms for ROB and 5.3 oxygen atoms per 15 carbons for GRI. This oxidation then gives an increase in mass of 7.5% or 40% for ROB and GRI, respectively, as discussed previously. (Note: It is assumed that H = 2×C + 2, which may not be strictly true, but an error of 1 or 2 hydrogen atoms per carbon does not substantially alter the calculated values for the mass increase.) With

this relationship O:C can be calculated for each generation of oxidation, and the OM:OC ratio can be calculated as well using the relationship OM:OC = 1 + (16/12)×O:C + (1/12)×H:C, in which H:C = 2 – 0.54×O:C (Murphy et al., 2011; Hayes et al., 2013; Canagaratna et al., 2015). Then the OM:OC ratio is used to convert the OM mass concentration in each generation bin to OC mass concentration, and the O:C ratio is used to convert the OC mass in each generation bin to O mass concentration. Finally, the O mass and OC mass are each summed and subsequently divided to obtain O:C.

2.6. Correction for changes in partitioning due to emissions into a shallower boundary layer upwind of Pasadena~~due to higher OA concentrations upwind of Pasadena~~

To account for changes in partitioning due to lower planetary boundary layer (PBL) heights, and thus, increased OA ~~particle~~ concentrations upwind of Pasadena, the concentrations of POA, V-SOA, and SI-SOA are increased upwind of Pasadena beyond the amount already simulated in the model. This correction is necessary because using CO as a conservative tracer of emissions does not account for how ~~particulate concentration~~the shallow boundary layer over Los Angeles in the morning influences partitioning between the gas and particle phases ~~upwind of Pasadena~~. Specifically, during the afternoon Pasadena is a receptor site for pollution from downtown Los Angeles that was generally emitted into a shallower boundary layer during the morning. The reduced vertical dilution will lead to higher concentrations of POA as well as any urban SOA formed, which in turn leads to higher partitioning to the particle phase and less gas phase oxidation of primary and secondary S/IVOCs.

The correction of the partitioning mass is estimated using three different methods depending on the time-of-day. First, for air parcels measured at 00:00 – 07:00 local time when the PBL height is essentially constant for an extended period and emissions are dominated by local sources (Hayes et al., 2013), no correction needs to be made. Second, for air parcels measured between 07:00 – 16:00 when the PBL is increasing as the air parcels are advected, a correction is applied that assumes the PBL increases linearly from the height measured in the early morning hours to the height measured for a given time of day. Third, for air parcels after

16:00, it is assumed that a residual layer aloft is decoupled from the ground after 16:00, resulting in no subsequent dilution.

The correction for the partitioning calculation described in the previous paragraph is an approximation, and two sensitivity studies are carried out to estimate the magnitude of the possible errors introduced by this approximation. The first study follows the approach described above, except that instead of linearly increasing the partitioning mass upwind of Pasadena the correction follows a step-function and increases the partitioning mass to its maximum value immediately upwind of the ground site. This test should overestimate the amount of partitioning to the particle-phase, since such a dramatic change in PBL height is not expected. The second sensitivity study simply applies no correction factor to the partitioning mass, and thus it underestimates the partitioning to the particle-phase. For the model runs with the ROB+TSI and GRI+TSI parameterizations the resulting changes in average predicted mass for the sensitivity studies are +4/-12% and +6/-7%, respectively. These changes are small, which indicates that the description of the boundary layer dilution does not have a major influence on the results.

2.7. WRF-CMAQ model runs

The Community Multiscale Air-Quality Model (WRF-CMAQ) version 5.0.1 (https://www.cmascenter.org/cmaq/) was applied with 4 km horizontal grid resolution and 34 vertical layers extending from the surface (layer 1 height ~38 m) to 50 mb for the time period matching the CalNex field campaign. Aqueous phase chemistry includes oxidation of sulfur and methyglyoxal (Carlton et al., 2008; Sarwar et al., 2013), gas phase chemistry is based on Carbon-Bond 05 with updates to toluene reactions (CB05-TU) (Yarwood, 2010), and inorganic chemistry is based on the ISORROPIA II thermodynamic model (Fountoukis and Nenes, 2007). WRF-CMAQ estimates SOA yields from VOC precursors including isoprene, monoterpenes, sesquiterpenes, xylenes, toluene, benzene, and methylglyoxal (Carlton et al., 2010). Note that WRF-CMAQ contains the SOA precursor species alkanes and glyoxal, but these are not explicit species in the CB05-TU gas phase mechanism (e.g., alkanes are mapped to “PAR”, or paraffins). SOA species oligomerize to non-volatile organic carbon grouped by anthropogenic and biogenic origin (Carlton et al., 2010).

The Weather Research and Forecasting model (WRF), Advanced Research WRF core (ARW) version 3.1 (Skamarock et al., 2008) was used to generate gridded meteorological fields used for input to WRF-CMAQ and the emissions model. Surface variables, flow patterns, and daytime mixing layer heights are generally well characterized during this time period (Baker et al., 2013). Hourly solar radiation and surface layer temperature estimated by the WRF model are used as input for the Biogenic Emission Inventory System (BEIS) version 3.14 to estimate hourly speciated VOC and NO_x emissions (Carlton and Baker, 2011).

Stationary point source emissions are based on continuous emissions monitor (CEM) data for 2010 where available and otherwise the 2008 version 2 National Emission Inventory (NEI). Area source emissions are also based on the 2008 version 2 NEI. Mobile sector (on-road and off-road) emissions are interpolated between 2007 and 2011 totals provided by the California Air Resources Board. Emissions from other areas of the United States and other countries are included through time and space variant lateral boundary inflow. Hourly boundary inflow concentrations are taken from a coarser WRF-CMAQ simulation covering the continental United States that used inflow estimates from a global GEOS-CHEM (version 8.03.02) model (http://acmg.seas.harvard.edu/geos/) simulation. Additional details regarding model setup and evaluation are provided elsewhere (Kelly et al., 2014).

2.8. WRF-Chem model runs

Weather Research and Forecasting Model coupled to Chemistry (WRF-Chem) is a fully coupled meteorology-chemistry model. WRF-Chem simulations were performed for May and June 2010 on a 12 km resolution domain, which covers a large part of the western United States. The model simulations include meteorological, gas, and aerosol phase chemical processes. The SOA scheme used in this study is based on the VBS approach. The SOA parameterization and other model parameterizations are described in detail by Ahmadov et al. (2012). Here the main objective of the WRF-Chem simulation was to estimate the OA contribution of the emission sources located upwind of the Los Angeles basin. Thus, all the anthropogenic emissions and biogenic VOC fluxes were set to zero over an area of 60 x 72 km covering the Los Angeles basin (Figure SI-4) in our simulation. The WRF-Chem simulated OA concentrations for the Pasadena site therefore provide an estimate of the BG-OA at this site.

3. Results and discussion

3.1. Modeling urban SOA mass concentration

3.1.1. Urban SOA concentration: model versus measurement comparisons

In Figure 4 the diurnal cycles of SV-OOA and urban SOA are shown. For all the model variations, the model V-SOA (light blue area) is substantially smaller than the observed SV-OOA concentrations (solid black line), even though the additional partitioning mass of SI-SOA is available for all model runs. ~~It is possible that the SOA yields used for V-SOA, which are based on the chamber experiment literature, are several-fold too low due to, for example, losses of gas-phase species to chamber walls. To investigate this possibility a model variation – named "ROB+4xV" – is run wherein the SOA yields from aromatics are increased by a factor of four, based on recent chamber studies in which higher concentrations of aerosol seed were utilized in order to suppress losses to chamber walls, and an upper limit of a factor of 4 increase in V-SOA yields was estimated (Zhang et al., 2014). The multi-generation aging of secondary species produced from VOCs is turned off in this variation.~~ ~~The result for ROB+4xV is shown in Figure 4.~~ Even in the~~is~~ model variation ROB+4xV where the V-SOA concentrations are substantially higher due to the higher VOC yields used, additional SOA precursors must be included to achieve model/measurement closure. This result is also true despite the inclusion of multi-generation V-SOA aging in ROB+TSI, GRI+TSI, and PYE+TSI, which increases the amount of SOA from VOCs to levels far beyond those observed in chambers, although over longer timescales than for the 4xV case. Previous work modeling SOA in Mexico City showed that either V-SOA aging or SI-SOA must be included in models to match observed SOA concentrations, but the inclusion of both resulted in an overprediction (Tsimpidi et al., 2010; Dzepina et al., 2011). In this study, the inclusion of aging only increases the concentration of V-SOA by 10 – 20% depending on the time of day due to the relatively low experimental photochemical ages. Thus, by testing models of SOA formation at short ages, our case study points towards the importance of additional SOA precursors such as P-S/IVOCs. Specifically, the contribution to total SOA from P-S/IVOCs in the box model is 65-75% (ROB+TSI), 80-87% (GRI+TSI), 80-92% (PYE+TSI), and 44-51% (ROB+4xV). The range indicates the variation in the contribution with the time-of-day. Thus, only in the ROB+4xV model variation is the

estimated contribution to SOA from VOCs generally larger than or equal to that from the P-S/IVOCs. We note however these percentages include only the urban SOA and not the background OA, which is likely also SOA as discussed above.

When comparing the four parameterizations for SOA formation, it is apparent that the GRI+TSI and ROB+4xV variations best reproduce the observations. The predicted SOA mass using GRI+TSI lies within the measurement uncertainty most of the day. In contrast, the ROB+TSI variation does not produce high enough concentrations of SOA, and the model is consistently lower than the measurements even after considering the measurement uncertainties. The PYE+TSI variation tends to over predict SOA concentrations especially at nighttime and in the morning, and also exhibits larger discrepancies with respect to measured POA concentrations (Figure SI-2). Finally, the performance of the ROB+4xV variation is similar to GRI+TSI, highlighting the uncertainties about the dominant SOA precursors in urban areas (i.e. VOCs vs. P-S/IVOCs).

In general, the measurements peak one hour later than the model, which may be due to the simple treatment of sources and transport in the modeled air mass, but the overall correlation is excellent: R = 0.93 - 0.94 for ROB+TSI, GRI+TSI, PYE+TSI, and ROB+4xV. This study contrasts with an earlier comparison of the ROB and GRI parameterizations for SI-SOA in Mexico City, which showed that GRI produces more SOA than observed (Dzepina et al., 2011). Although the same modeling method was used to quantify the emissions and properties of P-S/IVOCs in both studies, the sources, composition, and SOA yields of P-S/IVOCs in urban areas are poorly characterized, and differences in those between the two urban areas may explain the differences in model performance for Pasadena and Mexico City.

In addition, the effective SOA yields predicted in the box model for P-S/IVOCs can be compared against those determined in previous modeling and smog chamber studies. The effective yield is a function of photochemical aging, and thus for the purpose of this comparison we focus on the effective box model yields for 12:00 – 15:00 when there was a moderate amount of photochemical aging (5 h at an average OH concentration of 1.5×10^6 molecules cm^{-3}) comparable to the degree of aging typically achieved in chambers. During this period the effective yields for P-S/IVOCs were 12%, 27%, and 36% for ROB+TSI, GRI+TSI, and PYE+TSI, respectively. Zhao et al. (2014) recently carried out a modeling study of SOA formed

in Pasadena that was constrained with an extensive set of IVOC data and found an overall SOA yield for IVOCs of 29%, which falls within the range of effective yields for P-S/IVOCs that are predicted by the box model for the three different parameterizations. Jathar et al. (2014) also recently estimated from chamber studies an effective SOA yield of 10 – 40% for unspeciated organic emissions from combustion sources, which is also consistent with the P-S/IVOC yields from our box model. For reference, the effective yields for the aromatic VOCs m-xylene, toluene, and benzene under high-NO_X conditions in chamber studies range from 4 – 28% depending on the precursor identity and chamber conditions (Ng et al., 2007). Similar chamber studies on 12-carbon alkanes determined effective yields ranging from 11% – 160%, where the highest yield corresponded to a cyclic alkane (Loza et al., 2014). In general, it appears that the effective yields resulting from the box model for the lower photochemical ages used here are similar to those determined from other chamber and modeling studies.

It is also possible to compare the predicted IVOC concentrations in the box model versus the concentrations measured by Zhao et al. (2014). The comparison is summarized in Table SI-3 of the supporting information. In total, the initial IVOC concentrations in the box model are two times higher compared to those determined from measurements (16 μg m^{-3} versus 8(±1) μg m^{-3}). In addition, there is a larger difference for the $C^* = 10^3$ bin (2.5 μg m^{-3} versus 0.2(±0.1) μg m^{-3}). At the same time, the model used by Zhao et al. to predict urban SOA is lower than the measurements by 50% on the urban scale, whereas as the box model used here does not exhibit such a low bias. Given these differences we have run two sensitivity studies to explore how the model predictions depend on the IVOC emissions that are discussed in the following sections. The first sensitivity study reduces the emission of P-S/IVOCs from cooking emissions to zero (Section 3.1.2), and the second sensitivity study reduces all IVOC emissions by one-half (Section 3.1.5). Both of these variations greatly improve the agreement between the modeled and measured IVOC concentrations.

3.1.2. Total SOA concentration: fossil vs contemporary carbon

As described above, on average 71(±3)% of the SV-OOA is composed of fossil carbon (Zotter et al., 2014), and it is important to evaluate whether this percentage is consistent with the model results. As shown in Figure 4, the V-SOA from in-basin biogenics is very small, and V-

SOA is overwhelmingly from fossil carbon sources since it is dominated by aromatic precursors (see 3.1.3 below) and the main source of aromatic hydrocarbons in the Los Angeles basin is vehicle emissions (Borbon et al., 2013). For SI-SOA, two types of POA, and thus, primary P-S/IVOCs are included in this study. Since HOA is dominated by vehicle emissions, it is most likely composed of fossil carbon. On the other hand, CIOA will have a majority of modern carbon. In previous work we noted that 0 – 50% of the CIOA mass may be from non-cooking sources and, specifically, from vehicles (Hayes et al., 2013). Furthermore, recent results have shown that cooking emissions can form substantial amounts of SOA (El Haddad et al., 2012). If P-S/IVOCs emitted with HOA are 100% fossil carbon, P-S/IVOCs emitted with CIOA are 25(±25)% fossil, and both emission sources form SI-SOA with the same efficiency, then the corresponding amount of fossil urban SOA in the model would be 65(±9)%, 63(±12)%, 62(±12)%, and 78(±7)% for ROB+TSI, GRI+TSI, PYE+TSI, and ROB+4xV, respectively. It should be noted that these percentages do not include BG-SOA because the ^{14}C results from Zotter et al. (2014) correspond to SV-OOA. None of these predictions are significantly different from the ^{14}C measurements. An important caveat is that P-S/IVOCs from ~~CIOA~~ cooking sources are modeled using the same parameters as P-S/IVOCs from ~~HOA~~ vehicle sources. It is possible that cooking and vehicle emissions do not exhibit the same SOA-forming properties, but it is not clear which would be a more potent SOA precursor as there are no parameterizations specific to cooking emissions available in the literature. Thus, the ROB, GRI, and PYE parameterizations are used for all P-S/IVOCs regardless of their source, and the amount of SOA from HOA (or CIOA) associated P-S/IVOCs can be calculated as simply the product of the total SI-SOA and the ratio HOA/POA (or CIOA/POA), where the hourly HOA, CIOA, and SI-SOA concentrations are used. It should also be noted that in Los Angeles gasoline contains nearly 10% ethanol made from corn and thus modern carbon (de Gouw et al., 2012), but it is thought that ethanol and its combustion products are not incorporated into aerosols (Lewis et al., 2006).

It should be noted that the fossil/modern split from the box model that is described above depends on the initial P-S/IVOCs concentrations and volatility distribution assumed in the model. These parameters are not well constrained for cooking emissions, as discussed in further detail in Section 3.1.4 below. In addition, as discussed in the previous section (3.1.1) the concentration of primary IVOCs in the box model is higher than that measured. Thus, ~~A~~as an extreme sensitivity study, the model variations were also run under the assumption that ~~CIOA~~

cooking sources did not emit any P-S/IVOCs or, in the case of the PYE+TSI variation, any SVOCs (Figure SI-5). In this sensitivity study there is improved model/measurement agreement for the primary IVOCs as shown in Table SI-3. The GRI+TSI, PYE+TSI, and ROB+4xV variations reasonably reproduce the SV-OOA concentrations with some periods outside the measurement uncertainties. In contrast, the ROB+TSI variation without cooking-related P-S/IVOCs predicts concentrations that are too low. Regardless of the parameterization, a strong urban source of non-fossil SOA precursors, such as cooking emissions, must be included to obtain agreement with the ^{14}C measurements; otherwise the modeled SOA is overwhelmingly fossil. Clearly, there are still large uncertainties in SOA formation from cooking emissions. Further studies are needed to constrain models and to identify potential additional urban sources of non-fossil SOA, although our results suggest that cooking emissions are a potentially important source of anthropogenic non-fossil SOA.

3.1.3. SOA concentration apportionment to precursor compounds

The diurnal cycles of V-SOA mass concentration produced from individual VOCs are shown in Figure 5A. Among the VOCs the five largest contributors to V-SOA are methyl-substituted aromatics such as xylenes, trimethylbenzenes, and toluene. When SOA concentrations peak, these compounds account for ~70% of the predicted V-SOA mass. In Figure 5B the precursor-specific model predictions are compared against results from a methodology developed by the U.S. EPA that apportions SOA to specific precursors using molecular tracers measured in ambient aerosol samples (Kleindienst et al., 2012). For methylbenzenes (i.e. aromatics containing one or more methyl substituents) the tracer molecule utilized is 2,3-dihydroxy-4-oxopentanoic acid, and for naphthalene, 1-methylnaphthalene, and 2-methylnaphthalene the tracer molecule is phthalic acid and its associated methyl-containing analogs. Several tracers are used for isoprene (Edney et al., 2005) and monoterpenes (Jaoui et al., 2005; Claeys et al., 2007; Szmigielski et al., 2007), and they are listed in Table SI-4~~3~~ in the supporting information.

For the methylbenzenes, the model/tracer comparison is good, indicating consistency between model predictions and ambient measurements. The similarity further validates the model, although it is noted that if V-SOA 'aging' is eliminated from the model the model/tracer

comparison improves further and the difference becomes less than 5%. We note that this comparison cannot constrain whether chamber yields have been reduced by vapor losses, since the same effect would have occurred when measuring the yields included in the model and when measuring the SOA/tracer ratio used for the tracer estimate.

For the biogenic VOCs, isoprene and the monoterpenes, the tracer estimate indicates several-fold higher concentrations than predicted in the model. This difference is not surprising since the background SOA is thought to have a major contribution from isoprene and monoterpene oxidation in areas north of the Los Angeles Basin, and in the model BG-SOA from different VOCs is not resolved. In other words, the model results in Figure 5B represent only the in-basin biogenic SOA and are lower limits for total SOA from isoprene and monoterpenes. Moreover, the tracer estimates in Figure 5B are likely lower limits as well because the tracers may be lost by subsequent physical or chemical processes occurring in very aged aerosol transported to the measurement location (Hallquist et al., 2009). If the tracer molecule is oxidized or oligomerized, then it will be effectively lost with respect to the tracer method, even if its mass stays in the particle phase. If a semi-volatile tracer evaporates during atmospheric transport or from the filter after sampling, it is also lost from the point of view of the tracer method even though a chemical reaction has not occurred. It appears that the model/measurement comparison for the biogenic VOCs is qualitatively consistent given the known limitations of both approaches. However, the amount of SOA from biogenic VOCs as determined by the tracer method is only ~10% of the BG-SOA (0.22 versus 2.1 $\mu g\ m^{-3}$) even though the BG-SOA is predominantly biogenic as previously noted. The most likely explanation for the difference in mass concentration is the loss mechanisms described above. Other possible sources for the background such as biomass burning or marine OA are known to be very low at this location (Hayes et al., 2013), and more than 69% of the LV-OOA stems from non-fossil sources (Zotter et al., 2014).

Figure 5B also shows a comparison for the naphthalenes. The tracer estimates are over an order-of-magnitude higher than the model predictions when using the SOA yields from the literature (which are ~20% for the conditions of our study) and the emission ratios determined from the regression analysis of nighttime measurements shown in Figure SI-3. The model is also run using the empirically adjusted emission ratios that better match the observed concentrations

of the naphthalenes. ~~One can observe that t~~The model for this variation is still much lower than the tracer estimate. As an additional sensitivity study, we also run the model with the adjusted emissions and a yield of 150% that places all the oxidized mass in the C^*=1 μg m^{-3} volatility bin. This last variation represents an upper limit estimate of SOA from naphthalenes, in which nearly all of the mass plus the added oxygen partitions to the particle phase, which is much higher than laboratory observations. The tracer estimate, however, is still about a factor of two higher than the model. It is known that the tracer estimate is an upper limit, because the tracer compound, phthalic acid, may not be a unique tracer, and it potentially could be emitted from primary sources (Kleindienst et al., 2012). However, there may be other alkylated or functionalized PAHs that are not explicitly accounted for in the box model, and some of them might produce this tracer. ~~Thus, when considering this limitation it is concluded the model/measurement comparison is consistent.~~

The best estimate from the model with the adjusted emissions results in 0.7% of the predicted SOA being formed from the measured naphthalenes. Utilizing the upper limit of the model results for the PAHs~~naphthalene~~, including that from the parameterization with a purposefully high yield, it is apparent that naphthalene, 1-methyl naphthalene, and 2-methyl naphthalene ~~and its analogs~~ account for less than 4% of the SOA mass. While previous work has suggested that PAHs are important precursors for SOA in the Los Angeles Basin~~SoCAB~~ (Hersey et al., 2011) these earlier findings were qualitative and based on the observation of phthalic acid in samples. The work presented here, both the modeling results as well as the tracer results, quantitatively demonstrates that SOA from identified ~~naphthalene, 1-methylnaphthalene, and 2-methylnaphthalene~~PAHs is~~are~~ relatively ~~unimportant~~ small but not negligible when compared to ~~other precursors such as methylbenzenes~~the total SOA concentration. An upper limit for the contribution of~~, with~~ this group of precursors ~~accounting for at most~~is 8(±3)% of the SOA. ~~(~~This percentage is calculated using the tracer method in which the SOA concentration from PAHs is higher than in the box model and a 30% uncertainty for the SV-OOA concentration.~~)~~ Lastly, we note that no suitable tracers for alkane oxidation have been identified yet, which prevents carrying out similar model/tracer comparisons with respect to the P-S/IVOCs, ~~which~~since these compounds are thought to be composed primarily, although not exclusively, of alkanes.

3.1.4. SOA concentration apportionment to gasoline vehicles, diesel vehicles, cooking activities, and in-basin biogenic sources

In addition to apportioning the amount of SOA formed from individual compounds there is also considerable recent interest in the apportionment of SOA between diesel and gasoline vehicle emissions as well as other urban sources (Bahreini et al., 2012; Gentner et al., 2012; Hayes et al., 2013; Ensberg et al., 2014). The SOA model developed here can be used to address this important problem, and in Figure 6 the urban SOA mass calculated in the model is apportioned between diesel vehicles, gasoline vehicles, cooking sources, and in-basin biogenic emissions. The SOA mass is apportioned to each source using the following method, which can be described in five steps. First, the background is set to 2.1 μg m^{-3}. Second, the in-basin biogenic SOA is calculated as described in the methods section. Third, for the diesel contribution, since it is estimated that 70(±10)% of HOA is emitted from diesel vehicles (Hayes et al., 2013), it is assumed in the model that 70% of the P-S/IVOCs co-emitted with HOA are from diesel vehicles as well. (The remainder is assumed to be from gasoline vehicles.) While VOCs emissions from diesel vehicles are low (Warneke et al., 2012) in the Los Angeles Basin, VOCs have still been measured in diesel fuel. Specifically, using the measurements of Gentner et al. (2012) given in Tables S9 and S10 of that paper, the percentage of each VOC included in our model emitted from diesel vehicles is calculated. The precursor-specific SOA concentrations, as shown in here in Figure 5, are then multiplied by these percentages to determine the fraction of V-SOA attributable to diesel emissions, which is 3%. It should be noted that for all the VOCs included here except 1,3-butadiene, styrene, and anthropogenic isoprene, the corresponding concentrations in gasoline and diesel fuel are published in Gentner et al. (2012). Fourth, the cooking contribution is calculated by assuming that 75% of the P-S/IVOCs co-emitted with CIOA are from cooking activities. This percentage is chosen since it lies halfway between 50 and 100%, which is the current constraint from measurements on the amount of CIOA from cooking sources as discussed above and in Hayes et al. (2013). Fifth, the gasoline fraction is taken to be the SOA formed from all the remaining VOCs as well as the remaining P-S/IVOCs.

As can be seen in Figure 6, for the urban SOA (i.e. excluding the background OA) diesel, gasoline, and cooking emissions all contribute substantially to SOA formation, with the sum of gasoline and cooking being much larger than diesel for all model variants. In contrast, the in-

basin biogenic contribution is small. The analogous results when the background is included are shown in the supporting information (Figure SI-6). The formation of SOA from diesel emissions accounts for 16 – 27% of the urban SOA in the model depending on the variant used. This result is very similar to the percentage reported in our previous work, 19(+17/-21)%, which was determined using measurements of OOA weekly cycles (Hayes et al., 2013). In addition, the existence of a diesel contribution in the model is consistent with PMF analysis of FTIR spectra of OA filter samples collected in Pasadena, in which, one SOA component exhibited relative peak intensities in the C-H stretching region that suggest some ~~qualitatively a~~ contribution from diesel emissions (Guzman-Morales et al., 2014), although the percentage of SOA from diesel could not be determined in this previous work.~~.~~ The results of our work stand in contrast to those of Gentner et al. (2012) however, wherein the contribution of diesel and gasoline to vehicular SOA were estimated to be 70% and 30%, respectively. This discrepancy may be due to the assumption used by Gentner et al. that effectively all vehicle emissions are unburned fuel, whereas recent experiments have indicated that important SOA precursors exist in gasoline vehicle emissions that are not present in unburned gasoline when after-treatment devices such as catalytic converters are used (Jathar et al., 2013).

Also shown in Figure 6 is a bar graph summarizing the result from each parameterization grouped by fossil and non-fossil sources as well as the fossil fraction of SV-OOA determined by Zotter et al. (2014). The results of the two studies are consistent, with cooking and in-basin biogenic SOA accounting for between 23 – 38% of the in-basin SOA mass in the models. These two sources represent the modern fraction in the box model.

The uncertainties shown in Figure 6 (in parentheses) are calculated by propagating the uncertainty in the amount of HOA from diesel sources, as well as the uncertainty in the amount of CIOA from cooking sources under the assumption that the P-S/IVOCs co-emitted with these primary aerosols have similar uncertainties. It is also noted that another source of uncertainty is the selection of the ROB+TSI, GRI+TSI, PYE+TSI, or ROB+4xV model variation. The model variant used has an important impact on the apportionment, but the greatest amount of urban SOA formed from diesel emissions when considering all the uncertainties described in this paragraph is still only 31%.

The uncertainties in Figure 6 do not however account for certain assumptions that were made in order to perform the source apportionment. In particular, it was assumed that the P-S/IVOCs to POA ratio as well as the volatility distribution of P-S/IVOCs is the same for all sources, which is likely not the case. However, to our knowledge there is insufficient information in the literature to prescribe different volatility distributions for different sources.

Lastly, the percentage of SOA attributed to cooking emission in this work also requires discussion. Compared to gasoline or diesel vehicles there is relatively little data on the SOA forming potential of cooking emissions, but nevertheless there is both direct and indirect data supporting the SOA forming potential of cooking emissions. First, it is clear from numerous source apportionment studies that cooking emissions are a source of organic matter in the atmosphere (e.g. Robinson et al., 2006; Mohr et al., 2011; Hayes et al., 2013). Second, molecular speciation of cooking emissions has demonstrated that cooking activities emit a variety of volatile and semi-volatile compounds that are known SOA precursors (Schauer et al., 1999, 2002). Third, chamber studies have demonstrated SOA formation from cooking emissions. The latter results have been presented at several major conferences, but have not yet been published in the peer-reviewed literature (El Haddad et al., 2012). Thus, it is reasonable to conclude that SOA models should include the SOA resulting from chemical processing of cooking emissions, but there is a lack of chamber yields that could be used to develop specific SOA parameterizations. We have therefore assumed that SOA from cooking emission can be described using the same parameterizations as used for SOA from vehicular P-S/IVOCs. We have also performed a sensitivity study where we assume that cooking emissions do not produce any SOA. Ultimately, the percentage of SOA from cooking emissions reported here should be considered a first-order estimate that should be updated when additional data regarding SOA from cooking emissions becomes available.

3.1.5. Evolution of SOA concentration for 3 days

It is of high interest to explore the evolution of the different parameterizations discussed here at greater photochemical ages than those observed at the Pasadena site, since this behavior can lead to different results in regional and global modeling studies, and since similar combinations of parameterizations were found to overpredict regional SOA downwind of

Mexico City (Dzepina et al., 2011). To explore this question, the evolution of SOA concentration was simulated for 3 days using each of the four major variations (ROB+TSI, GRI+TSI, PYE+TSI, ROB+4xV). The same simulation was carried out for the SIMPLE model and it is discussed below in Section 3.3. The results are shown in Figure 7, and in order to facilitate comparisons the SOA concentrations are normalized to the CO concentration, after subtracting the CO background (DeCarlo et al., 2010). These simulations are for continuous aging at a reference •OH concentration of 1.5×10^{6} molec cm^{-3}, and thus, they do not attempt to simulate a diurnal variation in the amount of photochemical aging. This approach is used because it facilitates the comparison against field measurements described below. Furthermore, the box model does not account for how dilution downwind of Los Angeles may increase SOA evaporation and thus the rate of oxidation via increased partitioning to the gas phase. However, this phenomenon would only lead to small changes in total model SOA, and that should not change the conclusions discussed in this section (Dzepina et al., 2011). Also shown in Figure 7 is the same ratio, SOA/ΔCO, determined previously from measurements at the Pasadena site (Hayes et al., 2013). At photochemical ages less than 0.25 days, GRI+TSI and ROB+4xV perform the best (Table 4), which is consistent with the comparisons against the diurnal average of SOA since the diurnal photochemical age peaks at about 0.25 days (Figure 3A). However, for higher photochemical ages between 0.25 and 0.5 days the performance of ROB+TSI improves.

We also note that all of the parameterizations used in this section produce SOA/ΔCO ratios substantially larger (by factors of 2 or more) than those observed globally for aged air masses (i.e. photochemical ages greater than one day at an average OH concentration of 1.5×10^{6} molec cm^{-3}). For reference the range of OA/ΔCO ratios reported by de Gouw and Jimenez (2009) for aged urban SOA across multiple sites is indicated by the gray regions in Figure 7. This OA/ΔCO ratio includes both POA and SOA, but POA is a small contributor to OA/ΔCO for very aged air. Also shown in Figure 7 is the SOA/ΔCO ratio observed by Bahreini et al. (2012) from the NOAA P3 aircraft in the Los Angeles Basin outflow where air masses were aged from 1 – 2 days. This ratio is similar to the range taken from de Gouw and Jimenez (2009). The differences between the modeled and the measured SOA/ΔCO at higher photochemical ages may be important for regional and global models as they suggest an overestimation of urban SOA downwind of polluted regions. One possible explanation for the higher predicted values is the lack of a fragmentation mechanism in the parameterizations, which would reduce the SOA mass

by producing higher volatility products. Indeed, decreases in SOA concentration at high photochemical ages have been observed in flow-tube studies, although typically at photochemical ages much longer than 3 days (George and Abbatt, 2010). Also dry deposition in the regional models may decrease over-prediction depending on how it is implemented (Knote et al., 2014a).

A third explanation is the potential overestimation of IVOC emissions in the box model. As discussed in Section 3.1.1, the initial concentration of primary IVOCs in the model is a factor of 2 higher than the values determined from field measurements by Zhao et al (2014). To investigate this possibility, a sensitivity study was run in which the initial concentrations of primary IVOCs in the volatility bins C*= 10^3, 10^4, 10^5, and 10^6 were decreased by one-half. The results of this sensitivity study are shown in Figure 8. In general, ROB+TSI, GRI+TSI, PYE+TSI and ROB+4xV all show better agreement with measurements at long photochemical ages, although all four variants still overestimate the measurements. For shorter photochemical ages (in the urban scale) ROB+TSI under-predicts the SOA concentration, whereas GRI+TSI and ROB+4xV both predict SOA/ΔCO ratios that are not significantly different from the measured values (Hayes et al., 2013), and lastly PYE+TSI overestimates the SOA concentration. Thus, IVOCs emissions that are too high in the box model may be responsible for some, but not all, of the overestimation of SOA concentrations at long photochemical ages.

For reference, we note that when the IVOC concentrations are halved the four variations all predict less SI-SOA for the Pasadena ground site (Figure SI-7), but the contribution of P-S/IVOCs to SOA formation remains important: 59 – 73% (ROB+TSI), 72 – 80% (GRI+TSI), 79 – 92% (PYE+TSI), 38 – 48% (ROB+4xV). Furthermore, all four variations still predict a fossil fraction of urban SOA consistent with the ^{14}C measurements at the Pasadena site: 66(±9)%, 64(±10)%, 61(±12)%, 78(±6)%, respectively. Note that in calculating these fossil fractions the IVOCs emissions from cooking and gasoline/diesel were reduced by the same amount (i.e. one-half).

3.1.6. Comparison of WRF-CMAQ versus measurements and box model

The comparison of the SOA predicted for Pasadena by the WRF-CMAQ model is shown in Figure ~~8~~9A. Unlike the box model, the 3-D WRF-CMAQ model simulates the production and

transport of SOA both within and outside the Los Angeles Basin. It is therefore most appropriate to compare the WRF-CMAQ model output with OOA (SV-OOA + LV-OOA) rather than just SV-OOA as is done for the box model that focused only on the urban area. The WRF-CMAQ SOA is well correlated with the measured OOA (R= 0.73), but the SOA mass concentration in the model is ~25 times lower than the observed amount. This discrepancy is observed despite the fact that the VOCs show reasonable agreement (Supporting Information Figure SI-8~~7~~, Panels A – C). The difference of a factor of 25 in the SOA concentrations is also observed consistently across different photochemical ages (Supporting Information Figure SI-8~~7~~, Panel D). Furthermore, the performance of WRF-CMAQ is good for the inorganic aerosol species (Supporting Information Figures SI-9~~8~~ and SI-10~~9~~) as well as for elemental carbon and different meteorological parameters (Baker et al., 2013; Kelly et al., 2014). These comparisons indicate that while the model appears to be accurately simulating the transport to Pasadena and photochemical aging, the amount of SOA formed from urban precursors is greatly underestimated by WRF-CMAQ. Given the importance of P-S/IVOCs as SOA precursors in the box model, the lack of these species in WRF-CMAQ explains a substantial fraction of the difference between the models.

To further examine both WRF-CMAQ and the box model results, we modify the SOA module of the box model to be similar to the treatment of urban SOA in WRF-CMAQ as described by Carlton et al. (2010). First, for the box model P-S/IVOCs are not included, since these species are not in WRF-CMAQ. Second, the BG-SOA in the box model is adjusted to 0.1 μg m^{-3} so that the concentrations of SOA in the two models are similar in the early morning hours when the background dominates. Third, the box model uses a different approach for simulating V-SOA identical to that described by Dzepina et al. (2009). Briefly, instead of the VBS, an empirical 2-product parameterization wherein the oxidized products cannot undergo aging is used (Koo et al., 2003).

$$VOC + \bullet OH \rightarrow \alpha_1(SVOC_1) + \alpha_2(SVOC_2) \qquad (3)$$

The effective saturation concentration for each lumped product, *SVOC*, is then used to calculate the equilibrium partitioning between gas- and particle-phases as shown earlier in Equation 1.

Also, in Equation 3, α is the yield for each VOC. Note that the Koo et al. (2003) parameterization produces substantially lower V-SOA yields than the Tsimpidi et al. (2010) parameterization used in the rest of this work. The latter parameterization was updated using more recent chamber results (with higher yields) such as those of Ng et al. (2007). The use of the older Koo et al. (2003) parameterization for this specific comparison was motivated by the fact that the parameterization for urban SOA in the version of CMAQ used here is based on the same older data and does not contain the higher updated yields in, for example, Tsimpidi et al. (2010).

The results of the comparison of WRF-CMAQ with the modified box model are shown in Figure ~~8~~9B. With those modifications the results are very similar. This good agreement indicates that the differences between the default box model and WRF-CMAQ are not due to differences in transport or another variable, but rather the intrinsic differences in the SOA modules. In addition, the comparison between the two models suggests that 3-D air quality models need to include either SOA from P-S/IVOCs, additional precursor sources, and/or increased V-SOA yields to accurately predict SOA concentrations in the Los Angeles Basin and other urban areas.

3.2. Comparison of predicted and measured SOA oxygen content

The diurnal cycle of O:C of total OA is shown in Figure ~~9~~10, along with the estimated ±30% uncertainty of the O:C determination (Aiken et al., 2007; 2008). A recent re-evaluation of the AMS elemental analysis has found an underestimation of oxygen content for multi-functional oxidized organics (Canagaratna et al., 2015). Thus, the updated calibration factors have been used in the work here, and they increase the measured O:C and H:C by factors of 1.28 and 1.1, respectively. The model predictions of O:C are shown for both the ROB+TSI and GRI+TSI variations. The measured O:C is similar or higher than the models, and exhibits only small changes during the day. The minimum after noon in the measured O:C is due to the arrival of POA above Pasadena as well as the production of fresh SOA. The second minimum in the evening is due to emissions of CIOA, which has relatively low oxygen content.

When the model is run with the ROB+TSI variation for O:C evolution in SOA the model diurnal cycle is generally lower than the field data. Similar to the comparison of mass concentration, the GRI+TSI model variation better reproduces the O:C observations. As a control the model is also run without SI-SOA, which, interestingly, also does an excellent job of

reproducing the observations. Two conclusions can be drawn from the results shown in Figure ~~9~~10. First, the SI-SOA in the ROB parameterization appears to be not sufficiently oxidized, which drives down the predicted O:C, and, in general, SOA production and oxidation in Pasadena is very rapid and is therefore best described by the GRI parameterization. Second, both SI-SOA from the GRI parameterization and V-SOA have an O:C of ~ 0.45, which is not very different from the weighted mean of HOA, CIOA, and LV-OOA (O:C ~ 0.6), and, as a result, the total OA O:C is relatively constant for the different times of day. This consideration also explains why O:C does not change substantially when the SI-SOA is included or excluded in the model.

Lastly, we note that there are large uncertainties in the parameters used to predict O:C such as the distribution of O:C values as a function of volatility for V-SOA or the amount of oxygen mass added to the SI-SOA for each oxidation reaction. It is therefore not very meaningful to conclude from this study that one parameterization performs better than another. Rather, it is apparent that when using several different previously published SOA parameterizations (i.e. ROB+TSI, GRI+TSI, and TSI alone) it is possible to reproduce the observed O:C at the Pasadena ground site.

3.3. A simple parameterization for SOA formation in polluted urban regions

While medium-complexity parameterizations of SOA formation and evolution such as those used above represent some important details of SOA chemistry and properties, there is a need for very computationally inexpensive SOA parameterizations that still retain good accuracy for use in regional, global, and climate models. Such a parameterization was recently reported by Hodzic and Jimenez (2011), and was designed to predict properties of urban SOA in global and climate modeling studies (referred to as the "SIMPLE" parameterization hereinafter). The model represents SOA precursors as a single surrogate lumped species, termed here 'VOC*', which is emitted proportionally to anthropogenic CO. The model converts VOC* to SOA by reaction with •OH with a specified rate constant. The SOA formed in the SIMPLE model is non-volatile and does not partition to the gas-phase, consistent with the low volatility observed for aged SOA in field studies (e.g., Cappa and Jimenez, 2010).

We replaced the SOA parameterizations discussed above with the SIMPLE parameterization just described, and ran the box model for a large number of possible parameter value combinations (i.e. emission ratio of VOC*/CO and •OH rate constant). Figure 11~~0~~A shows the difference between model and measurement over that parameter space. The diurnal cycle predicted by the SIMPLE parameterization with the optimum parameters for Pasadena is shown in Figure 11~~0~~B. The SIMPLE model with the optimized parameters performs comparably to the more complex parameterizations used in this work. At the same time, the SIMPLE parameterization is unable to capture perfectly the location of the peak in time because it depends solely on CO and photochemical age. The CO concentration at the site peaks at 12:00 and photochemical age peaks at 13:00 (Figure 3A) while the measured SOA has a broad peak between 14:00 – 16:00. The fact that SOA does not peak at the same time as CO and photochemical age indicates the assumption in SIMPLE that VOC*/CO does not vary in time is probably not completely accurate. Still, the performance of the SIMPLE parameterization for urban SOA is sufficient for many applications and certainly far better than many models currently used.

Interestingly, the optimal model parameters for Mexico City and Pasadena are very similar.~~,~~ In other words, when tuning the model separately for each city, the parameters obtained are identical within the estimated uncertainties. ~~which~~This result suggests~~ the~~ SIMPLE~~ model~~, with the parameters reported for Mexico City or Pasadena, can be applied to other polluted urban regions as well. In addition, the optimal parameters for Pasadena (and Mexico City) are consistent with the OA/ΔCO ratios observed for highly aged air masses by Bahreini et al. (2012) from the NOAA P3 aircraft in the LA basin outflow, as well as for other urban areas as summarized by de Gouw and Jimenez (2009) and shown in Figure 7. However, it should be noted that a range of SIMPLE parameter combinations still remains in which the different combinations perform similarly in the model/measurement comparison, and this range is indicated by the dashed box in Figure 11~~0~~A.~~ This lack of a precise constraint is due to, in part, the limited range of photochemical ages observed at most stationary field sites. Thus~~ While the SIMPLE model is promising, additional work should be carried out to verify the optimal SIMPLE model parameters including analysis of data for a broad range of ages, e.g., by utilizing results from ambient air processed by oxidation flow reactors (Ortega et al., 2013). Also, the accuracy of the SIMPLE model for predicting urban SOA under a variety of atmospheric

conditions should be explored (e.g. VOC/NO_X or relative amounts of gasoline versus diesel emissions.) Finally, we note that the SIMPLE model parameterizes urban SOA, and is not applicable to biogenic SOA.

Hodzic and Jimenez (2011) also proposed an approach for predicting the oxygen content of SOA that utilized the equation O:C = 1 – 0.6exp(-*A*/1.5), where *A* is the photochemical age in days. (Note: the photochemical age was calculated using a reference •OH concentration of 1.5×10^{6} molec cm^{-3}.) As shown in Figure 12~~1~~, this parameterization compares well with the O:C from measurements. However, the parameterization of Hodzic and Jimenez does not take into account the new AMS O:C calibrations factors, as described in the preceding section. In order to account for this change, the equation proposed by Hodzic and Jimenez must be multiplied by a factor of 1.28. Thus, the updated parameterization is O:C = 1.28(1 – 0.6exp(-*A*/1.5)), and the corresponding O:C values are shown in Figure 12~~1~~. The updated simple parameterization also exhibits good agreement with measurements. (Note: The O:C predicted by the updated model does not increase by a factor 1.28 relative to the original version because the SOA from the Hodzic and Jimenez parameterization is mixed with HOA, CIOA, and BG-SOA to determine the total OA O:C shown in Figure 12~~1~~.)

3.4 Update of the U.S. and Global Urban SOA budgets

As shown in Figure 7, the SIMPLE parameterization asymptotically approaches a SOA/ΔCO value of 80 μg m^{-3} ppm^{-1}, which can be used to estimate an urban SOA budget. The SIMPLE parameterization is better for estimating this budget than the more complex parameterizations, because the SIMPLE parameterization is consistent with the observations of de Gouw and Jimenez (2009) that were made at multiple locations. For the ROB+TSI, GRI+TSI, PYE+TSI, and ROB+4xV model variants, values of SOA/ΔCO between 150 and 220 μg m^{-3} ppm^{-1} are predicted at long photochemical ages, and such high values have never been observed, to our knowledge, downwind of anthropogenic-dominated sources. These four more complex parameterizations are based on laboratory data at short photochemical ages, and thus, applying them to long photochemical ages is an extrapolation. The SIMPLE parameterization is imperfect, but based on the available evidence it appears that the SIMPLE model is the most accurate at long photochemical ages and better suited for estimating the urban SOA budget.

For the U.S., the annual urban CO emissions reported in the 2011 NEI are 44 Tg yr^{-1} (EPA, 2013), which when multiplied by SOA/ΔCO gives a~~n~~ national urban SOA source of 3.1 Tg yr^{-1}. ~~For reference, from the same database the national biogenic VOCs emissions are 37 Tg yr^{-1}. Then using an approximate yield of 10%, the biogenic VOCs would represent a SOA source of 3.7 Tg yr^{-1}. While there are several major uncertainties in this analysis – the accuracy of the SIMPLE model at greater photochemical ages, yields for biogenic VOCs, etc. – it is evident that the amount of urban SOA is not negligible compared to the amount of biogenic SOA in the U.S.~~ The same estimate can be performed for global urban SOA, since similar ratios of SOA/ΔCO have been observed in other countries such as downwind of Mexico City and China (DeCarlo et al., 2010; Hu et al., 2013). Using the EDGAR v4.2 inventory of 371 Tg yr^{-1} of urban/industrial CO for 2008 (JRC, 2011), we estimate a global pollution SOA source of 26 Tg yr^{-1}, or about 17% of the estimated global SOA source of 150 Tg yr^{-1} (Hallquist et al., 2009; Heald et al., 2010; Heald et al., 2011; Spracklen et al., 2011). We note that 1/3 of that SOA would be non-fossil, if a similar cooking fraction is observed globally as in this study, which is expected given the identification of similar fractions of cooking POA in many field studies globally (Wang et al., 2009; Mohr et al., 2011; Sun et al., 2011).

4. Conclusions

SOA in Pasadena during CalNex has been modeled using three different methods: (1) a box model, (2) a 3-D dimensional model, namely, WRF-CMAQ, and (3) a simple two-parameter model. Model/measurement comparisons clearly indicate that SOA formed from P-S/IVOCs, or a similar source, must be included in the models to accurately predict SOA concentrations in Pasadena. In other words, SOA from VOC oxidation is not sufficient to explain the observed concentrations, even when the highest SOA yields are used. Specifically, the parameterizations utilized were the Tsimpidi et al. (2010) parameterization with aging or a modified version of that parameterization in which the SOA yields for aromatic VOCs were multiplied by four as recently suggested by Zhang et al. (2014).

Three parameterizations for SOA formation from P-S/IVOCs were tested. It was found that the parameterization reported by Grieshop et al. (2009) best predicts ~~both~~ SOA concentration at the urban site ~~and SOA oxygen content~~. In contrast, the parameterization of

Robinson et al. (2007) predicts too little SOA~~ and too low oxygen content~~. These results contrast earlier modeling studies of Mexico City that showed the Robinson parameterization performed better when compared against the measured SOA concentration. The reason for the difference is not clear~~a different mixture of P-S/IVOCs at the two locations~~. Both the Mexico City and Pasadena studies suggest ~~indicate~~ that the Grieshop parameterization more accurately predicts SOA oxygen content, but this conclusion is also dependent on model parameters that are not well-constrained. Additionally, we tested the parameterization proposed in Pye and Seinfeld (2010) for the formation of SOA from P-S/IVOCs, which produces similar results but tends to overpredict SOA concentrations especially at nighttime and in the morning for this case study. The relative importance of VOCs and P-S/IVOCs as contributors to urban SOA over different time and length scales remains unclear. Depending on the parameterization used in the box model, the amount of urban SOA from VOCs can range between 15 – 53% of the total predicted SOA for the Pasadena ground site. This range is 16 – 58% in a sensitivity study in which the IVOC emissions are reduced by one-half. All the parameterizations used in the box model overpredict urban SOA at photochemical ages larger than one day compared to field observations, which has implications for their use in regional and global models. However, when the IVOC emissions in the box model are reduced by one-half to better match the measurements of Zhao et al. (2014) the predictions of SOA at long photochemical are improved although still too high, while the model/measurement comparison at short photochemical ages is still within the measurement uncertainties for the GRI+TSI and ROB+4xV variations.

This work represents the first chemically explicit evaluation of WRF-CMAQ SOA mass predictions in the Western U.S. or California. This model provides excellent predictions of secondary inorganic particle species but underestimates the observed SOA mass by a factor of about 25. The discrepancy is likely attributable to the VOC-only parameterization used that has relatively low yields and does not include SOA from P-S/IVOCs or a similar source.

SOA source apportionment was also carried out using the box model results. Among the explicitly modeled VOCs, the precursor compounds that contribute the most SOA mass are all methylbenzenes. In contrast, measured PAHs ~~(i.e.~~including naphthalene, 1-methylnaphthalene, and 2-methylnaphthalene~~)~~ are relatively minor precursors and contribute 0.7% of the SOA mass. In addition, the amount of urban SOA from diesel vehicles, gasoline vehicles, and cooking-related emissions is estimated to be 16 – 27%, 35 – 61%, and 19 – 35%, respectively,~~.~~ with an

almost factor of 2 difference in the estimated contribution depending on the box model variant used. A significant amount of background SOA appears to be formed outside the Los Angeles Basin and transported to the Pasadena site. The percentage estimated from diesel in the model is in agreement with our previous study that estimated the diesel contribution to be 0 – 36% by analyzing the weekly cycle of OOA concentrations (Hayes et al., 2013). The fraction of fossil and non-fossil urban SOA from the different models is generally consistent with the measurements. Importantly, a large source of urban non-fossil SOA most likely due to cooking is identified, while biogenic SOA formed from urban-scale emissions makes a small contribution.

The final portion of this work adapts the SIMPLE two parameter model of Hodzic and Jimenez (2011) to predict SOA properties for Pasadena. The simple model successfully predicts SOA concentration and oxygen content with accuracy similar to the more complex parameterizations. Furthermore, the optimal parameters for the SIMPLE model are very similar in both Mexico City and Pasadena, which indicates that this computationally inexpensive approach may be useful for predicting pollution SOA in global and climate models. Pollution SOA is estimated to account for 17% of global SOA, and we note that ~1/3 of urban SOA may be non-fossil mainly due to the impact of cooking and other sources.

Acknowledgements

This work was partially supported by CARB 08-319 and CARB 11-305, US DOE (BER, ASR program) DE-SC0006035, DE-SC0006711, and DE-SC0011105, NSF AGS-1243354 and AGS-1360834, and NOAA NA13OAR4310063. PLH is also grateful for a fellowship from the CIRES Visiting Fellows Program ~~as well as~~and PLH and PKM acknowledge support from a NSERC Discovery Grant and the Université de Montréal. The authors thank Chris J. Hennigan (UMBC) and Allen L. Robinson (CMU) for providing the naphthalene and methylnaphthalene data. We also thank John S. Holloway (NOAA) for providing CO data, Roya Bahreini (University of California-Riverside) and Ann M. Middlebrook (NOAA) for providing OA data from the NOAA P3, as well as Stuart A. McKeen (NOAA) for helpful discussions. RA is supported by the US Weather Research Program within NOAA/OAR Office of Weather and Air Quality. The US Environmental Protection Agency through its Office of Research and Development collaborated in the research described here. The manuscript is being subjected to

external peer review and has not been cleared for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

Ahmadov, R., McKeen, S.A., Robinson, A.L., Bahreini, R., Middlebrook, A.M., de Gouw, J.A., Meagher, J., Hsie, E.Y., Edgerton, E., Shaw, S. and Trainer, M. (2012) A volatility basis set model for summertime secondary organic aerosols over the eastern United States in 2006. J. Geophys. Res.-Atmos. 117, D06301.

Aiken, A.C., DeCarlo, P.F. and Jimenez, J.L. (2007) Elemental analysis of organic species with electron ionization high-resolution mass spectrometry. Anal. Chem. 79, 8350-8358.

Aiken, A.C., Decarlo, P.F., Kroll, J.H., Worsnop, D.R., Huffman, J.A., Docherty, K.S., Ulbrich, I.M., Mohr, C., Kimmel, J.R., Sueper, D., Sun, Y., Zhang, Q., Trimborn, A., Northway, M., Ziemann, P.J., Canagaratna, M.R., Onasch, T.B., Alfarra, M.R., Prevot, A.S.H., Dommen, J., Duplissy, J., Metzger, A., Baltensperger, U. and Jimenez, J.L. (2008) O/C and OM/OC ratios of primary, secondary, and ambient organic aerosols with high-resolution time-of-flight aerosol mass spectrometry. Environ. Sci. Technol. 42, 4478-4485.

Atkinson, R. and Arey, J. (2003) Atmospheric degradation of volatile organic compounds. Chem. Rev. 103, 4605-4638.

Bahreini, R., Middlebrook, A.M., de Gouw, J.A., Warneke, C., Trainer, M., Brock, C.A., Stark, H., Brown, S.S., Dube, W.P., Gilman, J.B., Hall, K., Holloway, J.S., Kuster, W.C., Perring, A.E., Prevot, A.S.H., Schwarz, J.P., Spackman, J.R., Szidat, S., Wagner, N.L., Weber, R.J., Zotter, P. and Parrish, D.D. (2012) Gasoline emissions dominate over diesel in formation of secondary organic aerosol mass. Geophys. Res. Lett. 39, L06805.

Baker, K.R., Misenis, C., Obland, M.D., Ferrare, R.A., Scarino, A.J. and Kelly, J.T. (2013) Evaluation of surface and upper air fine scale WRF meteorological modeling of the May and June 2010 CalNex period in California. Atmos. Environ. 80, 299-309.

Bertram, A.K., Martin, S.T., Hanna, S.J., Smith, M.L., Bodsworth, A., Chen, Q., Kuwata, M., Liu, A., You, Y. and Zorn, S.R. (2011) Predicting the relative humidities of liquid-liquid phase separation, efflorescence, and deliquescence of mixed particles of ammonium sulfate, organic material, and water using the organic-to-sulfate mass ratio of the particle and the oxygen-to-carbon elemental ratio of the organic component. Atmos. Chem. Phys. 11, 10995-11006.

Borbon, A., Gilman, J.B., Kuster, W.C., Grand, N., Chevaillier, S., Colomb, A., Dolgorouky, C., Gros, V., Lopez, M., Sarda-Esteve, R., Holloway, J., Stutz, J., Petetin, H., McKeen, S., Beekmann, M., Warneke, C., Parrish, D.D. and de Gouw, J.A. (2013) Emission ratios of anthropogenic volatile organic compounds in northern mid-latitude megacities: Observations versus emission inventories in Los Angeles and Paris. J. Geophys. Res.-Atmos. 118, 2041-2057.

Canagaratna, M.R., Jimenez, J.L., Kroll, J.H., Chen, Q., Kessler, S.H., Massoli, P., Hildebrandt Ruiz, L., Fortner, E., Williams, L.R., Wilson, K.R., Surratt, J.D., Donahue, N.M., Jayne, J.T. and Worsnop, D.R. (2015) Elemental ratio measurements of organic compounds using aerosol mass spectrometry: characterization, improved calibration, and implications. Atmos. Chem. Phys. 15, 253-272.

Cappa, C.D. and Jimenez, J.L. (2010) Quantitative estimates of the volatility of ambient organic aerosol. Atmos. Chem. Phys. 10, 5409-5424.

Cappa, C.D., Onasch, T.B., Massoli, P., Worsnop, D.R., Bates, T.S., Cross, E.S., Davidovits, P., Hakala, J., Hayden, K.L., Jobson, B.T., Kolesar, K.R., Lack, D.A., Lerner, B.M., Li, S.M., Mellon, D., Nuaaman, I., Olfert, J.S., Petaja, T., Quinn, P.K., Song, C., Subramanian, R., Williams, E.J. and Zaveri, R.A. (2012) Radiative Absorption Enhancements Due to the Mixing State of Atmospheric Black Carbon. Science 337, 1078-1081.

Cappa, C.D. and Wilson, K.R. (2011) Evolution of organic aerosol mass spectra upon heating: implications for OA phase and partitioning behavior. Atmos. Chem. Phys. 11, 1895-1911.

Carlton, A.G. and Baker, K.R. (2011) Photochemical Modeling of the Ozark Isoprene Volcano: MEGAN, BEIS, and Their Impacts on Air Quality Predictions. Environ. Sci. Technol. 45, 4438-4445.

Carlton, A.G., Bhave, P.V., Napelenok, S.L., Edney, E.D., Sarwar, G., Pinder, R.W., Pouliot, G.A. and Houyoux, M. (2010) Model Representation of Secondary Organic Aerosol in CMAQv4.7. Environ. Sci. Technol. 44, 8553-8560.

Carlton, A.G., Turpin, B.J., Altieri, K.E., Seitzinger, S.P., Mathur, R., Roselle, S.J. and Weber, R.J. (2008) CMAQ model performance enhanced when in-cloud SOA is included: comparisons of OC predictions with measurements. Environ. Sci. Technol. 42, 8798-8802.

Carter, W.P.L. (2010) Development of the SAPRC-07 chemical mechanism. Atmos. Environ. 44, 5324-5335.

Chan, A.W.H., Kautzman, K.E., Chhabra, P.S., Surratt, J.D., Chan, M.N., Crounse, J.D., Kürten, A., Wennberg, P.O., Flagan, R.C. and Seinfeld, J.H. (2009) Secondary organic aerosol formation from photooxidation of naphthalene and alkylnaphthalenes: implications for oxidation of intermediate volatility organic compounds (IVOCs). Atmos. Chem. Phys. 9, 3049-3060.

Chen, Q., Farmer, D.K., Schneider, J., Zorn, S.R., Heald, C.L., Karl, T.G., Guenther, A., Allan, J.D., Robinson, N., Coe, H., Kimmel, J.R., Pauliquevis, T., Borrmann, S., Poschl, U., Andreae, M.O., Artaxo, P., Jimenez, J.L. and Martin, S.T. (2009) Mass spectral characterization of submicron biogenic organic particles in the Amazon Basin. Geophys. Res. Lett. 36, L20806.

Claeys, M., Szmigielski, R., Kourtchev, I., Van der Veken, P., Vermeylen, R., Maenhaut, W., Jaoui, M., Kleindienst, T.E., Lewandowski, M., Offenberg, J.H. and Edney, E.O. (2007) Hydroxydicarboxylic Acids: Markers for Secondary Organic Aerosol from the Photooxidation of α-Pinene. Environ. Sci. Technol. 41, 1628-1634.

de Gouw, J.A., Gilman, J.B., Borbon, A., Warneke, C., Kuster, W.C., Goldan, P.D., Holloway, J.S., Peischl, J., Ryerson, T.B., Parrish, D.D., Gentner, D.R., Goldstein, A.H. and Harley, R.A. (2012) Increasing atmospheric burden of ethanol in the United States. Geophys. Res. Lett. 39, L15803.

de Gouw, J.A. and Jimenez, J.L. (2009) Organic Aerosols in the Earth's Atmosphere. Environ. Sci. Technol. 43, 7614-7618.

DeCarlo, P.F., Ulbrich, I.M., Crounse, J., de Foy, B., Dunlea, E.J., Aiken, A.C., Knapp, D., Weinheimer, A.J., Campos, T., Wennberg, P.O. and Jimenez, J.L. (2010) Investigation of the sources and processing of organic aerosol over the Central Mexican Plateau from aircraft measurements during MILAGRO. Atmos. Chem. Phys. 10, 5257-5280.

Docherty, K.S., Stone, E.A., Ulbrich, I.M., DeCarlo, P.F., Snyder, D.C., Schauer, J.J., Peltier, R.E., Weber, R.J., Murphy, S.M., Seinfeld, J.H., Grover, B.D., Eatough, D.J. and Jimenez, J.L. (2008) Apportionment of Primary and Secondary Organic Aerosols in Southern California during the 2005 Study of Organic Aerosols in Riverside (SOAR-1). Environ. Sci. Technol. 42, 7655-7662.

Dockery, D.W. and Pope, C.A. (1994) Acute respiratory effects of particulate air-pollution. Annu. Rev. Publ. Health 15, 107-132.

Donahue, N.M., Chuang, W., Epstein, S.A., Kroll, J.H., Worsnop, D.R., Robinson, A.L., Adams, P.J. and Pandis, S.N. (2013) Why do organic aerosols exist? Understanding aerosol lifetimes using the two-dimensional volatility basis set. Environmental Chemistry 10, 151-157.

Donahue, N.M., Robinson, A.L., Stanier, C.O. and Pandis, S.N. (2006) Coupled partitioning, dilution, and chemical aging of semivolatile organics. Environ. Sci. Technol. 40, 2635-2643.

Dzepina, K., Cappa, C.D., Volkamer, R.M., Madronich, S., DeCarlo, P.F., Zaveri, R.A. and Jimenez, J.L. (2011) Modeling the Multiday Evolution and Aging of Secondary Organic Aerosol During MILAGRO 2006. Environ. Sci. Technol. 45, 3496-3503.

Dzepina, K., Volkamer, R.M., Madronich, S., Tulet, P., Ulbrich, I.M., Zhang, Q., Cappa, C.D., Ziemann, P.J. and Jimenez, J.L. (2009) Evaluation of recently-proposed secondary organic aerosol models for a case study in Mexico City. Atmos. Chem. Phys. 9, 5681-5709.

Edney, E.O., Kleindienst, T.E., Jaoui, M., Lewandowski, M., Offenberg, J.H., Wang, W. and Claeys, M. (2005) Formation of 2-methyl tetrols and 2-methylglyceric acid in secondary organic aerosol from laboratory irradiated isoprene/NOX/SO2/air mixtures and their detection in ambient PM2.5 samples collected in the eastern United States. Atmos. Environ. 39, 5281-5289.

El Haddad, I., Platt, S., Slowik, J.G., Mohr, C., Crippa, M., Temime-Roussel, B., Detournay, A., Marchand, N., Baltensperger, U. and Prevot, A.S.H. (2012) *Contributions of Cooking Emissions to Primary and Secondary Organic Aerosol in Urban Atmospheres*, American Association for Aerosol Research 31st Annual Conference, Minneapolis, Minnesota http://aaarabstracts.com/2012/AbstractBook.pdf

Ensberg, J.J., Hayes, P.L., Jimenez, J.L., Gilman, J.B., Kuster, W.C., de Gouw, J.A., Holloway, J.S. and Seinfeld, J.H. (2014) Emission factor ratios, SOA mass yields, and the impact of vehicular emissions on SOA formation. Atmos. Chem. Phys. 14, 2383-2397.

EPA (2013) *National Emissions Inventory*. Environmental Protection Agency http://www.epa.gov/ttn/chief/net/2011inventory.html

Ervens, B. and Volkamer, R. (2010) Glyoxal processing by aerosol multiphase chemistry: towards a kinetic modeling framework of secondary organic aerosol formation in aqueous particles. Atmos. Chem. Phys. 10, 8219-8244.

Fountoukis, C. and Nenes, A. (2007) ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K+-Ca2+-Mg2+-Nh(4)(+)-Na+-SO42--NO3--Cl--H2O aerosols. Atmos. Chem. Phys. 7, 4639-4659.

Gentner, D.R., Isaacman, G., Worton, D.R., Chan, A.W.H., Dallmann, T.R., Davis, L., Liu, S., Day, D.A., Russell, L.M., Wilson, K.R., Weber, R., Guha, A., Harley, R.A. and Goldstein, A.H. (2012) Elucidating secondary organic aerosol from diesel and gasoline

vehicles through detailed characterization of organic carbon emissions. Proc. Natl. Acad. Sci. U. S. A. 109, 18318-18323.

George, I.J. and Abbatt, J.P.D. (2010) Chemical evolution of secondary organic aerosol from OH-initiated heterogeneous oxidation. Atmos. Chem. Phys. 10, 5551-5563.

Grieshop, A.P., Logue, J.M., Donahue, N.M. and Robinson, A.L. (2009) Laboratory investigation of photochemical oxidation of organic aerosol from wood fires 1: measurement and simulation of organic aerosol evolution. Atmos. Chem. Phys. 9, 1263-1277.

Griffin, R.J., Chen, J.J., Carmody, K., Vutukuru, S. and Dabdub, D. (2007) Contribution of gas phase oxidation of volatile organic compounds to atmospheric carbon monoxide levels in two areas of the United States. J. Geophys. Res.-Atmos. 112, D10S17.

Guzman-Morales, J., Frossard, A.A., Corrigan, A.L., Russell, L.M., Liu, S., Takahama, S., Taylor, J.W., Allan, J., Coe, H., Zhao, Y. and Goldstein, A.H. (2014) Estimated contributions of primary and secondary organic aerosol from fossil fuel combustion during the CalNex and Cal-Mex campaigns. Atmos. Environ. 88, 330-340.

Hallquist, M., Wenger, J.C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N.M., George, C., Goldstein, A.H., Hamilton, J.F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M.E., Jimenez, J.L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T.F., Monod, A., Prevot, A.S.H., Seinfeld, J.H., Surratt, J.D., Szmigielski, R. and Wildt, J. (2009) The formation, properties and impact of secondary organic aerosol: current and emerging issues. Atmos. Chem. Phys. 9, 5155-5236.

Hayes, P.L., Ortega, A.M., Cubison, M.J., Froyd, K.D., Zhao, Y., Cliff, S.S., Hu, W.W., Toohey, D.W., Flynn, J.H., Lefer, B.L., Grossberg, N., Alvarez, S., Rappenglück, B., Taylor, J.W., Allan, J.D., Holloway, J.S., Gilman, J.B., Kuster, W.C., de Gouw, J.A., Massoli, P., Zhang, X., Liu, J., Weber, R.J., Corrigan, A.L., Russell, L.M., Isaacman, G., Worton, D.R., Kreisberg, N.M., Goldstein, A.H., Thalman, R., Waxman, E.M., Volkamer, R., Lin, Y.H., Surratt, J.D., Kleindienst, T.E., Offenberg, J.H., Dusanter, S., Griffith, S., Stevens, P.S., Brioude, J., Angevine, W.M. and Jimenez, J.L. (2013) Organic aerosol composition and sources in Pasadena, California during the 2010 CalNex campaign. J. Geophys. Res.-Atmos., 9233-9257.

Heald, C.L., Coe, H., Jimenez, J.L., Weber, R.J., Bahreini, R., Middlebrook, A.M., Russell, L.M., Jolleys, M., Fu, T.M., Allan, J.D., Bower, K.N., Capes, G., Crosier, J., Morgan, W.T., Robinson, N.H., Williams, P.I., Cubison, M.J., DeCarlo, P.F. and Dunlea, E.J. (2011) Exploring the vertical profile of atmospheric organic aerosol: comparing 17 aircraft field campaigns with a global model. Atmos. Chem. Phys. 11, 12673-12696.

Heald, C.L., Ridley, D.A., Kreidenweis, S.M. and Drury, E.E. (2010) Satellite observations cap the atmospheric organic aerosol budget. Geophys. Res. Lett. 37, L24808.

Heo, J., de Foy, B., Olson, M.R., Pakbin, P., Sioutas, C. and Schauer, J.J. (2015) Impact of regional transport on the anthropogenic and biogenic secondary organic aerosols in the Los Angeles Basin. Atmos. Environ. 103, 171-179.

Hersey, S.P., Craven, J.S., Schilling, K.A., Metcalf, A.R., Sorooshian, A., Chan, M.N., Flagan, R.C. and Seinfeld, J.H. (2011) The Pasadena Aerosol Characterization Observatory (PACO): chemical and physical analysis of the Western Los Angeles basin aerosol. Atmos. Chem. Phys. 11, 7417-7443.

Hodzic, A. and Jimenez, J.L. (2011) Modeling anthropogenically-controled secondary organic aerosols in a megacity: a simplified framework for global and climate models. Geosci. Model Dev. 4, 901-917.

Hodzic, A., Jimenez, J.L., Madronich, S., Aiken, A.C., Bessagnet, B., Curci, G., Fast, J., Lamarque, J.F., Onasch, T.B., Roux, G., Schauer, J.J., Stone, E.A. and Ulbrich, I.M. (2009) Modeling organic aerosols during MILAGRO: importance of biogenic secondary organic aerosols. Atmos. Chem. Phys. 9, 6949-6981.

Hodzic, A., Jimenez, J.L., Madronich, S., Canagaratna, M.R., DeCarlo, P.F., Kleinman, L. and Fast, J. (2010a) Modeling organic aerosols in a megacity: potential contribution of semi-volatile and intermediate volatility primary organic compounds to secondary organic aerosol formation. Atmos. Chem. Phys. 10, 5491-5514.

Hodzic, A., Jimenez, J.L., Prevot, A.S.H., Szidat, S., Fast, J.D. and Madronich, S. (2010b) Can 3-D models explain the observed fractions of fossil and non-fossil carbon in and near Mexico City? Atmos. Chem. Phys. 10, 10997-11016.

Hu, W.W., Hu, M., Yuan, B., Jimenez, J.L., Tang, Q., Peng, J.F., Hu, W., Shao, M., Wang, M., Zeng, L.M., Wu, Y.S., Gong, Z.H., Huang, X.F. and He, L.Y. (2013) Insights on organic aerosol aging and the influence of coal combustion at a regional receptor site of central eastern China. Atmos. Chem. Phys. 13, 10095-10112.

IPCC (2013) *Climate Change 2013: The Physical Scientific Basis*. Intergovernmental Panel on Climate Change: Working Group I, Geneva Switzerland

Jaoui, M., Kleindienst, T.E., Lewandowski, M., Offenberg, J.H. and Edney, E.O. (2005) Identification and Quantification of Aerosol Polar Oxygenated Compounds Bearing Carboxylic or Hydroxyl Groups. 2. Organic Tracer Compounds from Monoterpenes. Environ. Sci. Technol. 39, 5661-5673.

Jathar, S.H., Gordon, T.D., Hennigan, C.J., Pye, H.O.T., Pouliot, G., Adams, P.J., Donahue, N.M. and Robinson, A.L. (2014) Unspeciated organic emissions from combustion sources and their influence on the secondary organic aerosol budget in the United States. Proc. Natl. Acad. Sci. U. S. A. 111, 10473-10478.

Jathar, S.H., Miracolo, M.A., Tkacik, D.S., Donahue, N.M., Adams, P.J. and Robinson, A.L. (2013) Secondary Organic Aerosol Formation from Photo-Oxidation of Unburned Fuel: Experimental Results and Implications for Aerosol Formation from Combustion Emissions. Environ. Sci. Technol. 47, 12886-12893.

Jimenez, J.L., Canagaratna, M.R., Donahue, N.M., Prevot, A.S.H., Zhang, Q., Kroll, J.H., DeCarlo, P.F., Allan, J.D., Coe, H., Ng, N.L., Aiken, A.C., Docherty, K.S., Ulbrich, I.M., Grieshop, A.P., Robinson, A.L., Duplissy, J., Smith, J.D., Wilson, K.R., Lanz, V.A., Hueglin, C., Sun, Y.L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J.M., Collins, D.R., Cubison, M.J., Dunlea, E.J., Huffman, J.A., Onasch, T.B., Alfarra, M.R., Williams, P.I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J.Y., Zhang, Y.M., Dzepina, K., Kimmel, J.R., Sueper, D., Jayne, J.T., Herndon, S.C., Trimborn, A.M., Williams, L.R., Wood, E.C., Middlebrook, A.M., Kolb, C.E., Baltensperger, U. and Worsnop, D.R. (2009) Evolution of Organic Aerosols in the Atmosphere. Science 326, 1525-1529.

JRC (2011) *Emission Database for Global Atmospheric Research*. European Commission's Joint Research Centre http://edgar.jrc.ec.europa.eu/overview.php?v=42

Kelly, J.T., Baker, K.R., Nowak, J.B., Murphy, J.G., Markovic, M.Z., VandenBoer, T.C., Ellis, R.A., Neuman, J.A., Weber, R.J. and Roberts, J.M. (2014) Fine-scale simulation of ammonium and nitrate over the South Coast Air Basin and San Joaquin Valley of California during CalNex-2010. J. Geophys. Res.-Atmos. 119, 3600-3614.

Kleindienst, T.E., Jaoui, M., Lewandowski, M., Offenberg, J.H. and Docherty, K.S. (2012) The formation of SOA and chemical tracer compounds from the photooxidation of naphthalene and its methyl analogs in the presence and absence of nitrogen oxides. Atmos. Chem. Phys. 12, 8711-8726.

Kleinman, L.I., Daum, P.H., Lee, Y.N., Senum, G.I., Springston, S.R., Wang, J., Berkowitz, C., Hubbe, J., Zaveri, R.A., Brechtel, F.J., Jayne, J., Onasch, T.B. and Worsnop, D. (2007) Aircraft observations of aerosol composition and ageing in New England and Mid-Atlantic States during the summer 2002 New England Air Quality Study field campaign. J. Geophys. Res.-Atmos. 112, D09310.

Knote, C., Hodzic, A. and Jimenez, J.L. (2014a) The effect of dry and wet deposition of condensable vapors on secondary organic aerosols concentrations over the continental US. Atmos. Chem. Phys. Discuss. 14, 13731-13767.

Knote, C., Hodzic, A., Jimenez, J.L., Volkamer, R., Orlando, J.J., Baidar, S., Brioude, J., Fast, J., Gentner, D.R., Goldstein, A.H., Hayes, P.L., Knighton, W.B., Oetjen, H., Setyan, A., Stark, H., Thalman, R., Tyndall, G., Washenfelder, R., Waxman, E. and Zhang, Q. (2014b) Simulation of semi-explicit mechanisms of SOA formation from glyoxal in aerosol in a 3-D model. Atmos. Chem. Phys. 14, 6213-6239.

Koo, B.Y., Ansari, A.S. and Pandis, S.N. (2003) Integrated approaches to modeling the organic and inorganic atmospheric aerosol components. Atmos. Environ. 37, 4757-4768.

Kroll, J.H., Ng, N.L., Murphy, S.M., Flagan, R.C. and Seinfeld, J.H. (2006) Secondary organic aerosol formation from isoprene photooxidation. Environ. Sci. Technol. 40, 1869-1877.

Lewis, C.W., Volckens, J., Braddock, J.N., Crews, W.S., Lonneman, W.A. and McNichol, A.P. (2006) Absence of 14C in PM2.5 Emissions from Gasohol Combustion in Small Engines. Aerosol Sci. Technol. 40, 657-663.

Lim, H.J., Carlton, A.G. and Turpin, B.J. (2005) Isoprene forms secondary organic aerosol through cloud processing: Model simulations. Environ. Sci. Technol. 39, 4441-4446.

Loza, C.L., Craven, J.S., Yee, L.D., Coggon, M.M., Schwantes, R.H., Shiraiwa, M., Zhang, X., Schilling, K.A., Ng, N.L., Canagaratna, M.R., Ziemann, P.J., Flagan, R.C. and Seinfeld, J.H. (2014) Secondary organic aerosol yields of 12-carbon alkanes. Atmos. Chem. Phys. 14, 1423-1439.

Martin-Reviejo, M. and Wirtz, K. (2005) Is benzene a precursor for secondary organic aerosol? Environ. Sci. Technol. 39, 1045-1054.

Matsunaga, A. and Ziemann, P.J. (2010) Gas-Wall Partitioning of Organic Compounds in a Teflon Film Chamber and Potential Effects on Reaction Product and Aerosol Yield Measurements. Aerosol Sci. Technol. 44, 881-892.

McKeen, S., Chung, S.H., Wilczak, J., Grell, G., Djalalova, I., Peckham, S., Gong, W., Bouchet, V., Moffet, R., Tang, Y., Carmichael, G.R., Mathur, R. and Yu, S. (2007) Evaluation of several PM(2.5) forecast models using data collected during the ICARTT/NEAQS 2004 field study. J. Geophys. Res.-Atmos. 112.

Middlebrook, A.M., Bahreini, R., Jimenez, J.L. and Canagaratna, M.R. (2012) Evaluation of Composition-Dependent Collection Efficiencies for the Aerodyne Aerosol Mass Spectrometer using Field Data. Aerosol Sci. Technol. 46, 258-271.

Mohr, C., DeCarlo, P.F., Heringa, M.F., Chirico, R., Slowik, J.G., Richter, R., Reche, C., Alastuey, A., Querol, X., Seco, R., Peñuelas, J., Jiménez, J.L., Crippa, M., Zimmermann, R., Baltensperger, U. and Prévôt, A.S.H. (2011) Identification and quantification of organic aerosol from cooking and other sources in Barcelona using aerosol mass spectrometer data. Atmos. Chem. Phys. 11, 1649-1665.

Murphy, B.N., Donahue, N.M., Fountoukis, C. and Pandis, S.N. (2011) Simulating the oxygen content of ambient organic aerosol with the 2D volatility basis set. Atmos. Chem. Phys. 11, 7859-7873.

Murphy, D.M., Cziczo, D.J., Froyd, K.D., Hudson, P.K., Matthew, B.M., Middlebrook, A.M., Peltier, R.E., Sullivan, A., Thomson, D.S. and Weber, R.J. (2006) Single-particle mass spectrometry of tropospheric aerosol particles. J. Geophys. Res.-Atmos. 111, D23S32.

Ng, N.L., Kroll, J.H., Chan, A.W.H., Chhabra, P.S., Flagan, R.C. and Seinfeld, J.H. (2007) Secondary organic aerosol formation from m-xylene, toluene, and benzene. Atmos. Chem. Phys. 7, 3909-3922.

Ortega, A.M., Day, D.A., Cubison, M.J., Brune, W.H., Bon, D., de Gouw, J.A. and Jimenez, J.L. (2013) Secondary organic aerosol formation and primary organic aerosol oxidation from biomass-burning smoke in a flow reactor during FLAME-3. Atmos. Chem. Phys. 13, 11551-11571.

Parrish, D.D., Stohl, A., Forster, C., Atlas, E.L., Blake, D.R., Goldan, P.D., Kuster, W.C. and de Gouw, J.A. (2007) Effects of mixing on evolution of hydrocarbon ratios in the troposphere. J. Geophys. Res.-Atmos. 112, D10S34.

Perraud, V., Bruns, E.A., Ezell, M.J., Johnson, S.N., Yu, Y., Alexander, M.L., Zelenyuk, A., Imre, D., Chang, W.L., Dabdub, D., Pankow, J.F. and Finlayson-Pitts, B.J. (2012) Nonequilibrium atmospheric secondary organic aerosol formation and growth. Proc. Natl. Acad. Sci. U. S. A. 109, 2836-2841.

Pye, H.O.T. and Seinfeld, J.H. (2010) A global perspective on aerosol from low-volatility organic compounds. Atmos. Chem. Phys. 10, 4377-4401.

Renbaum-Wolff, L., Grayson, J.W., Bateman, A.P., Kuwata, M., Sellier, M., Murray, B.J., Shilling, J.E., Martin, S.T. and Bertram, A.K. (2013) Viscosity of alpha-pinene secondary organic material and implications for particle growth and reactivity. Proc. Natl. Acad. Sci. U. S. A. 110, 8014-8019.

Robinson, A.L., Donahue, N.M., Shrivastava, M.K., Weitkamp, E.A., Sage, A.M., Grieshop, A.P., Lane, T.E., Pierce, J.R. and Pandis, S.N. (2007) Rethinking organic aerosols: Semivolatile emissions and photochemical aging. Science 315, 1259-1262.

Robinson, A.L., Subramanian, R., Donahue, N.M., Bernardo-Bricker, A. and Rogge, W.F. (2006) Source apportionment of molecular markers and organic aerosol. 3. Food cooking emissions. Environ. Sci. Technol. 40, 7820-7827.

Ryerson, T.B., Andrews, A.E., Angevine, W.M., Bates, T.S., Brock, C.A., Cairns, B., Cohen, R.C., Cooper, O.R., de Gouw, J.A., Fehsenfeld, F.C., Ferrare, R.A., Fischer, M.L., Flagan, R.C., Goldstein, A.H., Hair, J.W., Hardesty, R.M., Hostetler, C.A., Jimenez, J.L., Langford, A.O., McCauley, E., McKeen, S.A., Molina, L.T., Nenes, A., Oltmans, S.J., Parrish, D.D., Pederson, J.R., Pierce, R.B., Prather, K., Quinn, P.K., Seinfeld, J.H., Senff, C.J., Sorooshian, A., Stutz, J., Surratt, J.D., Trainer, M., Volkamer, R., Williams, E.J. and Wofsy, S.C. (2013) The 2010 California Research at the Nexus of Air Quality and Climate Change (CalNex) field study. J. Geophys. Res.-Atmos. 118, 5830-5866.

Sarwar, G., Fahey, K., Kwok, R., Gilliam, R.C., Roselle, S.J., Mathur, R., Xue, J., Yu, J. and Carter, W.P.L. (2013) Potential impacts of two SO2 oxidation pathways on regional sulfate concentrations: Aqueous-phase oxidation by NO2 and gas-phase oxidation by Stabilized Criegee Intermediates. Atmos. Environ. 68, 186-197.

Schauer, J.J., Kleeman, M.J., Cass, G.R. and Simoneit, B.R.T. (1999) Measurement of emissions from air pollution sources. 1. C-1 through C-29 organic compounds from meat charbroiling. Environ. Sci. Technol. 33, 1566-1577.

Schauer, J.J., Kleeman, M.J., Cass, G.R. and Simoneit, B.R.T. (2002) Measurement of emissions from air pollution sources. 4. C-1-C-27 organic compounds from cooking with seed oils. Environ. Sci. Technol. 36, 567-575.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X., Wang, W. and Powers, J.G. (2008) A description of the Advanced Reserch WRF version 3. NCAR Technical Note NCAR/TN-475+STR.

Slowik, J.G., Stroud, C., Bottenheim, J.W., Brickell, P.C., Chang, R.Y.W., Liggio, J., Makar, P.A., Martin, R.V., Moran, M.D., Shantz, N.C., Sjostedt, S.J., van Donkelaar, A., Vlasenko, A., Wiebe, H.A., Xia, A.G., Zhang, J., Leaitch, W.R. and Abbatt, J.P.D. (2010) Characterization of a large biogenic secondary organic aerosol event from eastern Canadian forests. Atmos. Chem. Phys. 10, 2825-2845.

Spracklen, D.V., Jimenez, J.L., Carslaw, K.S., Worsnop, D.R., Evans, M.J., Mann, G.W., Zhang, Q., Canagaratna, M.R., Allan, J., Coe, H., McFiggans, G., Rap, A. and Forster, P. (2011) Aerosol mass spectrometer constraint on the global secondary organic aerosol budget. Atmos. Chem. Phys. 11, 12109-12136.

Sun, Y.L., Zhang, Q., Schwab, J.J., Demerjian, K.L., Chen, W.N., Bae, M.S., Hung, H.M., Hogrefe, O., Frank, B., Rattigan, O.V. and Lin, Y.C. (2011) Characterization of the sources and processes of organic and inorganic aerosols in New York city with a high-resolution time-of-flight aerosol mass apectrometer. Atmos. Chem. Phys. 11, 1581-1602.

Szmigielski, R., Surratt, J.D., Gómez-González, Y., Van der Veken, P., Kourtchev, I., Vermeylen, R., Blockhuys, F., Jaoui, M., Kleindienst, T.E., Lewandowski, M., Offenberg, J.H., Edney, E.O., Seinfeld, J.H., Maenhaut, W. and Claeys, M. (2007) 3-methyl-1,2,3-butanetricarboxylic acid: An atmospheric tracer for terpene secondary organic aerosol. Geophys. Res. Lett. 34, L24811.

Tsimpidi, A.P., Karydis, V.A., Zavala, M., Lei, W., Molina, L., Ulbrich, I.M., Jimenez, J.L. and Pandis, S.N. (2010) Evaluation of the volatility basis-set approach for the simulation of organic aerosol formation in the Mexico City metropolitan area. Atmos. Chem. Phys. 10, 525-546.

Tunved, P., Hansson, H.C., Kerminen, V.M., Strom, J., Dal Maso, M., Lihavainen, H., Viisanen, Y., Aalto, P.P., Komppula, M. and Kulmala, M. (2006) High natural aerosol loading over boreal forests. Science 312, 261-263.

Volkamer, R., Jimenez, J.L., Martini, F.S., Dzepina, K., Zhang, Q., Salcedo, D., Molina, L.T., Worsnop, D.R. and Molina., M.J. (2006) Secondary Organic Aerosol Formation from Anthropogenic Air Pollution: Rapid and Higher than Expected. Geophys. Res. Lett. 33, L17811.

Volkamer, R., San Martini, F., Molina, L.T., Salcedo, D., Jimenez, J.L. and Molina, M.J. (2007) A missing sink for gas-phase glyoxal in Mexico City: Formation of secondary organic aerosol. Geophys. Res. Lett. 34, L19807.

Volkamer, R., Ziemann, P.J. and Molina, M.J. (2009) Secondary Organic Aerosol Formation from Acetylene (C(2)H(2)): seed effect on SOA yields due to organic photochemistry in the aerosol aqueous phase. Atmos. Chem. Phys. 9, 1907-1928.

Wang, Q., Shao, M., Zhang, Y., Wei, Y., Hu, M. and Guo, S. (2009) Source apportionment of fine organic aerosols in Beijing. Atmos. Chem. Phys. 9, 8573-8585.

Warneke, C., de Gouw, J.A., Holloway, J.S., Peischl, J., Ryerson, T.B., Atlas, E., Blake, D., Trainer, M. and Parrish, D.D. (2012) Multiyear trends in volatile organic compounds in Los Angeles, California: Five decades of decreasing emissions. J. Geophys. Res.-Atmos. 117, D00V17.

Warneke, C., McKeen, S.A., de Gouw, J.A., Goldan, P.D., Kuster, W.C., Holloway, J.S., Williams, E.J., Lerner, B.M., Parrish, D.D., Trainer, M., Fehsenfeld, F.C., Kato, S., Atlas, E.L., Baker, A. and Blake, D.R. (2007) Determination of urban volatile organic compound emission ratios and comparison with an emissions database. J. Geophys. Res.-Atmos. 112, D10S47.

Washenfelder, R.A., Young, C.J., Brown, S.S., Angevine, W.M., Atlas, E.L., Blake, D.R., Bon, D.M., Cubison, M.J., de Gouw, J.A., Dusanter, S., Flynn, J., Gilman, J.B., Graus, M., Griffith, S., Grossberg, N., Hayes, P.L., Jimenez, J.L., Kuster, W.C., Lefer, B.L., Pollack, I.B., Ryerson, T.B., Stark, H., Stevens, P.S. and Trainer, M.K. (2011) The glyoxal budget and its contribution to organic aerosol for Los Angeles, California, during CalNex 2010. J. Geophys. Res.-Atmos. 116, D00V02.

Watson, J.G. (2002) Visibility: Science and regulation. J. Air Waste Manage. Assoc. 52, 628-713.

Yarwood, G., Jung, J., Whitten, G.Z., Heo, G., Mellberg, J., Estes, E. (2010) *Updates to the Carbon Bond Mechanism for Version 6 (CB6). Presented at the 9th Annual CMAS Conference, Chapel, Hill, NC*. ENVIRON International Corporation, Novato http://www.camx.com/publ/pdfs/CB05_Final_Report_120805.pdf

Yatavelli, R.L.N., Stark, H., Thompson, S.L., Kimmel, J.R., Cubison, M.J., Day, D.A., Campuzano-Jost, P., Palm, B.B., Hodzic, A., Thornton, J.A., Jayne, J.T., Worsnop, D.R. and Jimenez, J.L. (2014) Semicontinuous measurements of gas–particle partitioning of organic acids in a ponderosa pine forest using a MOVI-HRToF-CIMS. Atmos. Chem. Phys. 14, 1527-1546.

Zhang, Q., Jimenez, J.L., Canagaratna, M.R., Allan, J.D., Coe, H., Ulbrich, I., Alfarra, M.R., Takami, A., Middlebrook, A.M., Sun, Y.L., Dzepina, K., Dunlea, E., Docherty, K., DeCarlo, P.F., Salcedo, D., Onasch, T., Jayne, J.T., Miyoshi, T., Shimono, A., Hatakeyama, S., Takegawa, N., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Williams, P., Bower, K., Bahreini, R., Cottrell, L., Griffin, R.J., Rautiainen, J., Sun, J.Y., Zhang, Y.M. and Worsnop, D.R. (2007) Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes. Geophys. Res. Lett. 34, L13801.

Zhang, X., Cappa, C.D., Jathar, S.H., McVay, R.C., Ensberg, J.J., Kleeman, M.J. and Seinfeld, J.H. (2014) Influence of vapor wall loss in laboratory chambers on yields of secondary organic aerosol. Proc. Natl. Acad. Sci. U. S. A. 111, 5802-5807.

Zhang, X., Liu, J., Parker, E.T., Hayes, P.L., Jimenez, J.L., de Gouw, J.A., Flynn, J.H., Grossberg, N., Lefer, B.L. and Weber, R.J. (2012) On the gas-particle partitioning of soluble organic aerosol in two urban atmospheres with contrasting emissions: 1. Bulk water-soluble organic carbon. J. Geophys. Res.-Atmos. 117, D00V16.

Zhao, Y., Hennigan, C.J., May, A.A., Tkacik, D.S., de Gouw, J.A., Gilman, J.B., Kuster, W.C., Borbon, A. and Robinson, A.L. (2014) Intermediate-Volatility Organic Compounds: A Large Source of Secondary Organic Aerosol. Environ. Sci. Technol. 48, 13743-13750.

Zotter, P., El-Haddad, I., Zhang, Y., Hayes, P.L., Zhang, X., Lin, Y.-H., Wacker, L., Schnelle-Kreis, J., Abbaszade, G., Zimmermann, R., Surratt, J.D., Weber, R., Jimenez, J.L., Szidat, S., Baltensperger, U. and Prévôt, A.S.H. (2014) Diurnal cycle of fossil and nonfossil carbon using radiocarbon analyses during CalNex. J. Geophys. Res.-Atmos. 119, 6818-6835.

Table 1. Definitions of acronyms frequently used in this article.

AMS	Aerosol Mass Spectrometer
BG-SOA	Background secondary organic aerosols
CalNex	California research at the nexus of air quality and climate change field campaign
CIOA	Cooking-influenced organic aerosol
GRI	Grieshop et al. (2009) parameterization for secondary organic aerosol formation from P-S/IVOCs
IVOCs	Intermediate volatility organic compounds
NEI	National Emissions Inventory
OA	Organic aerosol
ODR	Orthogonal distance regression
PAH	Polycyclic aromatic hydrocarbon
PBL	Planetary Boundary Layer
P-S/IVOCs	Primary semi-volatile and intermediate volatility organic compounds
PYE	Pye and Seinfeld (2010) parameterization for secondary organic aerosols formation from P-S/IVOCs
ROB	Robinson et al. (2007) parameterization for secondary organic aerosol formation from P-S/IVOCs
SI-SOA	Secondary organic aerosol from primary semi-volatile and intermediate volatility organic compounds
SOA	Secondary organic aerosol
SVOCs	Semi-volatile organic compounds
TSI	Tsimpidi et al. (2010) parameterization for secondary organic aerosol formation from VOCs
V-SOA	Secondary organic aerosol formed from the oxidation of volatile organic compounds
VBS	Volatility basis set
VOCs	Volatile organic compounds
WRF-CMAQ	Weather Research Forecasting – Community multiscale air quality model
ΔCO	Enhanced CO concentration over the background concentration (105 ppb).

Table 2. Summary of the SOA models and their major variants used in this work.

Model Name	Variation	Notes	References	Figures
Box Model	1 (ROB + TSI)	VOCs: Tsimpidi et al. parameterization with aging. P-S/IVOCs: Robinson et al. parameterization, and all SOA treated within VBS framework.	Tsimpidi et al. *Atmos. Chem. Phys.* **2010,** 525-546. Robinson et al. *Science* **2007,** 1259-1262.	4, 6, 7, 8, 10, SI-2, SI-5, SI-6, SI-7
	2 (GRI + TSI)	VOCs: Tsimpidi et al. parameterization with aging. P-S/IVOCs: Grieshop et al. parameterization, and all SOA treated within VBS framework.	Tsimpidi et al. *Atmos. Chem. Phys.* **2010,** 525-546. Grieshop et al. *Science* **2009,** 1263-1277.	4, 5, 6, 7, 8, 10, SI-2, SI-5, SI-6, SI-7
	3 (PYE + TSI)	VOCs: Tsimpidi et al. parameterization with aging. P-S/IVOCs: Pye and Seinfeld parameterization.	Tsimpidi et al. *Atmos. Chem. Phys.* **2010,** 525-546. Pye and Seinfeld *Atmos. Chem. Phys.* **2010,** 4377-4401.	4, 6, 7, 8, SI-2, SI-5, SI-6, SI-7
	4 (ROB + 4xV)	VOCs: Tsimpidi et al. parameterization without aging and aromatic yield multiplied by 4. P-S/IVOCs: Robinson et al. parameterization, and all SOA treated within VBS framework.	Tsimpidi et al. *Atmos. Chem. Phys.* **2010,** 525-546. Robinson et al. *Science* **2007,** 1259-1262. Zhang et al. *PNAS* **2014**.	4, 6, 7, 8, SI-5, SI-6, SI-7
	5	VOCs: Koo et al. and Ng et al. wherein SOA is treated in a lumped product parameterization.	Koo et al. *Atmos. Environ.* **2003,** 4757-4768. Ng et al. *Atmos. Chem. Phys.* **2007,** 3909-3922.	8

Table 2 (Continued). Summary of the SOA models and their major variants used in this work.

Model Name	Variation	Notes	References	Figures
WRF-CMAQ	v5.0.1	4 anthropogenic VOC and 3 biogenic VOC precursors and GLY/MGLY. 12 semi-volatile partitioning species and 7 non-volatile SOA species	Carlton et al. *Environ. Sci. Technol.* **2010,** 8553-8560.	9, SI-8, SI-9, SI-10
SIMPLE	N/A	Single lumped precursor and single lumped, non-volatile SOA product.	Hodzic et al. *Geosci. Model Dev.* **2011,** 901-917.	7, 8, 11, 12
WRF-Chem	N/A	4-bin VBS framework with aging, 7 anthropogenic VOC classes and 4 biogenic VOC classes	Ahmadov et al. *J. Geophys. Res.-Atmos.* **2012,** D06301.	3, SI-4

Table 3. Measurements acquired at the Pasadena ground site during CalNex and used in this study.

Measurement	Technique	Uncertainty	Reference
Bulk aerosol mass concentrations for organics, nitrate, sulfate, and ammonium as well as the concentrations of organic aerosol components	High-resolution Aerosol Mass Spectrometry (AMS) and Positive Matrix Factorization (PMF) analysis	±30%	Hayes et al. 2013
Oxygen-to-carbon ratio	High-resolution Aerosol Mass Spectrometer (AMS) and Elemental Analysis (EA)	±30%	Hayes et al. 2013
Speciated VOCs	Gas chromatography – mass spectrometry	±5 – 25% (hydrocarbons) ±20 – 35% (oxygenates)	Borbon et al. 2013
CO	VUV Fluorescence	±4%	Hayes et al. 2013
Modern and fossil fraction of organic carbon	^{14}C	See text	Zotter et al. 2014
Concentration of SOA from specific precursor compounds	U.S. E.P.A. tracer method and measurement of oxygenates from filter samples using GC-MS	See text	Kleindienst et al. 2012
Concentration of naphthalene and its derivatives	Thermal desorption gas chromatography mass spectrometry	±30%	Presto et al. 2011 Presto et al. 2012

Table 4. Slope of SOA/ΔCO as reported by Hayes et al. (2013), and as predicted in the four major box model variations. For the box model, the slopes are obtained by performing a linear ODR analysis on the data shown in Figure 7.

Variation	SOA / ΔCO slope between 0 and 0.25 Days	SOA / ΔCO slope between 0.25 and 0.5 Days
Observed (Hayes et al. 2013)	108 $\mu g\ m^{-3}$	
ROB + TSI	69 $\mu g\ m^{-3}\ ppmv^{-1}$	88 $\mu g\ m^{-3}\ ppmv^{-1}$
GRI + TSI	110 $\mu g\ m^{-3}\ ppmv^{-1}$	130 $\mu g\ m^{-3}\ ppmv^{-1}$
PYE + TSI	168 $\mu g\ m^{-3}\ ppmv^{-1}$	153 $\mu g\ m^{-3}\ ppmv^{-1}$
ROB + 4xV	105 $\mu g\ m^{-3}\ ppmv^{-1}$	123 $\mu g\ m^{-3}\ ppmv^{-1}$

Figure Captions

Figure 1. Schematic of the major SOA parameterizations used in the box model. The different regions of the volatility scale are indicated on the top axis: low-volatility organic compounds (LVOCs), semi-volatile organic compounds (SVOCs), intermediate volatility organic compounds (IVOCs), and volatile organic compounds (VOCs). The fraction in the particle phase, F_p (top panel), increases with decreasing volatility (i.e. C^*) according to Equation 1. The campaign average OA concentration, 7 μg m^{-3}, has been used to calculate the partitioning. The parameterization of Tsimpidi et al. (2010) distributes the VOC oxidation products into four volatility bins, and subsequent oxidation reactions are allowed as indicated by the curved arrows. The two parameterizations for P-S/IVOC oxidation from Robinson et al. (2007) and Grieshop et al. (2009) are illustrated as well. Lastly, the parameterization of Pye and Seinfeld (2010) is shown in which SVOCs are treated as four lumped species (pink), and IVOCs are treated using the yields and volatility distribution for naphthalene oxidation (yellow). For clarity the arrows indicating IVOC aging are not shown.

Figure 2. Schematic of the SOA model set-up used in this work. Model inputs are indicated by hollow arrows whereas steps in the modeling process are indicated by solid arrow. All the steps in the dashed box are repeated for each hour of the day.

Figure 3. (A) Average diurnal cycle of CO (red) and photochemical age (blue) for the Pasadena ground site during CalNex. Note: A background of 105 ppbv has been subtracted from the CO concentration. **(B)** Average diurnal cycle of the five OA components identified by PMF analysis, as well as the background OA calculated from WRF-Chem. The five components are semi-volatile oxygenated organic aerosol (SV-OOA), cooking-influenced organic aerosol (CIOA), hydrocarbon-like organic aerosol (HOA), local organic aerosol (LOA), and low volatility organic aerosol (LV-OOA).

Figure 4. Model/measurement comparisons for urban SOA mass concentration plotted by time of day. The model results are shown for the **(ROB+TSI)**, **(GRI+TSI)**, **(PYE+TSI)**, and **(ROB+4xV)** variations. The model variations are described in Table 2. In all panels the SV-OOA determined from measurements at the Pasadena ground site is shown. The uncertainty for the AMS measurement used to determine the SV-OOA concentration is indicated by the dashed lines (Middlebrook et al., 2012).

Figure 5. (A) Predicted SOA mass from precursor VOCs. For clarity only the five largest contributors to the SOA mass are shown. Note that SI-SOA from P-S/IVOCs is not included in this panel. **(B)** Campaign average concentrations of SOA from specific precursors as determined in the box model as well as by the U.S. EPA tracer method (Kleindienst et al., 2012). Comparisons are shown for methylbenzenes, naphthalenes, isoprene, and monoterpenes. For the naphthalenes the bar for "adjusted emissions" indicates the model variation where the naphthalene emissions are increased in order to match the measured concentrations in Pasadena as shown in Figure SI-3. The adjusted emissions are also used for the variation with a yield of 1.5 at C*=1. Note: The GRI parameterization is used to predict the SI-SOA for these results.

Figure 6. The estimated fractional contribution to SOA mass concentration from gasoline vehicles, diesel vehicles, cooking emissions, and in-basin biogenic emissions. The results for the four model variations are displayed as pie charts as well as a bar chart. The bar chart also shows the percentage of SOA that is from fossil or modern sources as determined by Zotter et al. (2014). The modern sources are indicated by hollow bars and fossil sources are indicated by solid bars. Background SOA is not included in this figure, but the analogous figure with background SOA is given in Figure SI-6 of the supporting information.

Figure 7. SOA concentration predicted by the ROB+TSI, GRI+TSI, PYE+TSI, and ROB+4xV parameterizations for up to 3 days of photochemical aging at a reference •OH concentration of 1.5×10^{6} molec cm^{-3}. Also shown in the four panels is the same result for the SIMPLE model using the optimized parameters. Note that the SOA concentrations have been normalized to the background subtracted CO concentration to account for changes in emission strengths, and the processed data are identified by the symbol SOA/ΔCO. In addition, the SOA/ΔCO data determined for the Pasadena site from the measurements of Hayes et al. (2013) are shown as well as similar airborne measurements for the Los Angeles basin outflow performed by Bahreini et al. (2012) aboard the NOAA P3 (black marker). The Bahreini et al. data point corresponds to an average of all data between 1 – 2 days of photochemical aging. The OA/ΔCO ratio reported by de Gouw and Jimenez (2009) is also indicated (gray box) to serve as an estimate of S~~O~~OA/ΔCO in highly aged air masses. ~~For clarity, the uncertainty in the SOA determined from measurements (±30%) is not shown.~~

Figure 8. SOA concentration predicted by the ROB+TSI, GRI+TSI, PYE+TSI, and ROB+4xV parameterizations for up to 3 days of photochemical aging at a reference •OH concentration of 1.5×10^{6} molec cm^{-3}. These predictions correspond to the sensitivity study in which the concentration of IVOCs in the volatility bins $C^{*}=10^{3}$ – 10^{6} were reduced by one-half. Also shown in the three panels is the same result for the SIMPLE model using the optimized parameters (see Section 3.3 for further discussion). Note that the SOA concentrations have been normalized to the background subtracted CO concentration to account for changes in emission strengths, and the processed data are identified by the symbol SOA/ΔCO. In addition, the SOA/ΔCO data determined for the Pasadena site from the measurements of Hayes et al. (2013) are shown (black line) as well as similar airborne measurements downwind of Pasadena performed by Bahreini et al. (2012) aboard the NOAA P3 (black marker). The Bahreini et al. point corresponds to an average of all LA Basin outflow data between 1 – 2 days of photochemical aging. The OA/ΔCO ratio reported by de Gouw and Jimenez (2009) is also indicated (gray box) to serve as an estimate of SOA/ΔCO in highly aged air masses.

Figure 9~~8~~. (A) Scatter plot of SOA predicted by the WRF-CMAQ model versus the OOA determined from measurements at the Pasadena ground site. Also shown in this panel is an ODR linear regression analysis of the data with the y-intercept fixed to zero. **(B)** SOA diurnal cycles from the WRF-CMAQ and box model. The box model was run using an empirical two product parameterization (i.e., Model Variant 5 in Table 2) wherein the oxidized products cannot undergo aging (Dzepina et al., 2009).

Figure 10~~9~~. Model/measurement comparison for O:C of total OA versus time of day. The left panel contains the results when using the ROB+TSI model variation, whereas the right panel contains the results when using the GRI+TSI model variation. In both panels the O:C of OA measured at the Pasadena ground site is shown along with the O:C uncertainty. Shown in both panels is the model O:C when including only the SOA from VOCs (blue line), and the model O:C when including the SOA from both VOCs and P-S/IVOCs (pink line).

Figure 11~~0~~. (A) Image plot of the root mean square error between the SIMPLE urban SOA parameterization concentration and the measured SV-OOA as a function of both the lumped precursor emission ratio and the oxidation rate constant. The gray stars indicate the parameter pairs that result in the minimum errors for Pasadena (this study) and Mexico City (Hodzic and

Jimenez, 2011). The dashed box approximately indicates the range of possible optimal parameter combinations. For reference an emission ratio of 80 μg m^{-3} $ppmv^{-1}$ equals 0.069 g g^{-1}. **(B)** Diurnal cycle of SV-OOA with corresponding uncertainty (grey dashed lines). The diurnal cycle of SOA predicted by the SIMPLE model is shown as well.

Figure 12~~1~~. Model/measurement comparison of O:C of OA versus time of day for the SIMPLE urban SOA parameterization. The original parameterization proposed by Hodzic and Jimenez (2011) is O:C = 1 – 0.6exp(-*A*/1.5), where *A* is the photochemical age. The updated SIMPLE parameterization is O:C = 1.28(1 – 0.6exp(-*A*/1.5)), which accounts for the updated AMS O:C calibration factors.