



Criteria pollutant impacts of volatile chemical products informed by near-field modelling

Momei Qin^{1,2}  , Benjamin N. Murphy³ , Kristin K. Isaacs³, Brian C. McDonald⁴, Quanyang Lu^{5,6}, Stuart A. McKeen^{4,7}, Lauren Koval⁸, Allen L. Robinson^{5,6} , Christos Efstathiou⁹, Chris Allen⁹ and Haval O. T. Pye³  

Consumer, industrial and commercial product use is a source of exposure to potentially hazardous chemicals. In addition, cleaning agents, personal care products, coatings and other volatile chemical products (VCPs) evaporate and react in the atmosphere, producing secondary pollutants. Here, we show that high air emissions from VCP use (≥ 14 kg per person per yr, at least 1.7 \times higher than current operational estimates) are supported by multiple estimation methods and constraints imposed by ambient levels of ozone, hydroxyl radical reactivity and the organic component of fine particulate matter (PM_{2.5}) in Pasadena, California. A near-field model, which estimates human chemical exposure during or in the vicinity of product use, indicates that these high air emissions are consistent with organic product use up to ~75 kg per person per yr, and the inhalation of consumer products could be a non-negligible exposure pathway. After the PM_{2.5} yield is constrained to 5% by mass, VCPs produce ~41% of the photochemical organic PM_{2.5} ($1.1 \pm 0.3 \mu\text{g m}^{-3}$) and ~17% of the maximum daily 8 hr average ozone (9 ± 2 ppb) in summer in Los Angeles. Therefore, both toxicity and ambient criteria pollutant formation should be considered when organic substituents are developed for VCPs in pursuit of safer and more sustainable products and cleaner air.

Volatile chemical products (VCPs), including cleaning agents, personal care products, coatings, pesticides, printing inks and adhesives, contain complex mixtures of organic species that span a wide range of volatilities and oxidation states. The evaporation of components, such as organic solvents, results in emissions of volatile organic compounds (VOCs), primarily in indoor environments¹. Direct exposure to these VOCs can result in adverse health effects, including mucosal irritation, endocrine disruption, asthma and non-specific symptoms such as headache^{2–6}. Furthermore, these VOCs escape into ambient air, where they contribute to the formation of criteria pollutants such as fine particulate matter (PM_{2.5}) and ozone (O₃)¹, imposing further risks to public health^{7–9}. Because of the dramatic reduction in tailpipe emissions in the United States and Europe, recent estimates suggest that VCPs may be the largest source of urban VOC emissions¹. However, gaps remain in our understanding of the impacts of VCPs on air quality and human health.

Organic aerosol (OA) is a substantial component of ambient PM_{2.5} (ref. ¹⁰), with a large fraction (typically >60%) formed in the atmosphere from secondary chemistry^{11,12}. Regional chemical transport models struggle to simulate PM_{2.5}, specifically the organic component formed by photochemical activity (that is, secondary OA (SOA)) in urban areas^{13–18}. Li et al.¹⁹ report high PM_{2.5} yields for some individual intermediate volatility compounds found in VCPs and consumer product mixtures. Siloxanes, widely used in consumer products and commonly observed in urban air²⁰, can have PM_{2.5} yields of 20–50% (refs. ^{21,22}). Underestimated VCP emissions,

including species that are semivolatile to intermediately volatile^{1,23} or highly oxygenated²⁴, are probably responsible for some of the missing organic PM_{2.5} in regional air quality models.

Given the potential contribution of VCPs to both near-field (nominally indoor, during or in the vicinity of product use²⁵) and far-field (ambient air) exposure pathways, an approach that integrates their resulting impacts is needed. While most direct exposure to VCPs for people²⁶ will be in the indoor environment, ecosystem and other environmental exposure (for example, secondary pollutants) resulting from VCPs will be dominated by the far-field environment^{27–29}. Far-field exposure due to residential and commercial releases of VCPs and their resulting criteria pollutant formation are not considered in current chemical exposure estimates²⁶.

In this work, we show that a model developed for quantifying near-field chemical exposure due to consumer product use can be used to estimate ambient air emissions for far-field air quality modelling. Implementing higher ambient chemical emission estimates in a regional-scale chemical transport model results in organic PM_{2.5} and ozone predictions that are more consistent with ambient measurements. These simulations demonstrate that VCPs are important precursors to criteria pollutants and that ambient air measurements can be used to constrain near-field chemical exposure models.

Results

Estimating the impacts of VCP use on air quality and public health requires information on (1) organic product usage (the amount of

¹Oak Ridge Institute for Science and Education (ORISE) Research Participant at the Office of Research and Development, US Environmental Protection Agency, Research Triangle Park, NC, USA. ²Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control, Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, School of Environmental Science and Engineering, Nanjing University of Information Science & Technology, Nanjing, China. ³Office of Research and Development, US Environmental Protection Agency, Research Triangle Park, NC, USA. ⁴Chemical Sciences Laboratory, NOAA Earth System Research Laboratories, Boulder, CO, USA. ⁵Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA. ⁶Center for Atmospheric Particle Studies, Carnegie Mellon University, Pittsburgh, PA, USA. ⁷Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA. ⁸Oak Ridge Associated Universities (ORAU) at the Office of Research and Development, US Environmental Protection Agency, Research Triangle Park, NC, USA. ⁹General Dynamics Information Technology, Research Triangle Park, NC, USA.  e-mail: momei.qin@nuist.edu.cn; Pye.Havala@epa.gov

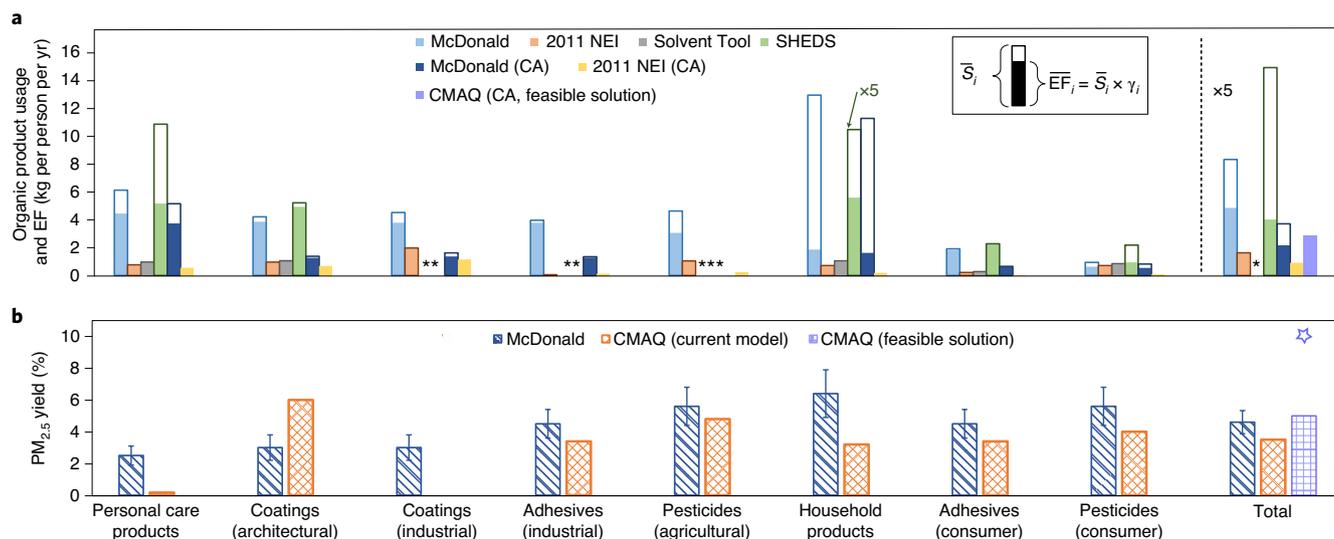


Fig. 1 | Comparison of organic product usage, EFs and PM_{2.5} yields across different methodologies. **a**, Organic product usage per person \bar{S}_i (total height of the bars including unfilled portion) and $\bar{E}F_i$ values attenuated by washing down the drain (filled portion of the bars). The \bar{S}_i and $\bar{E}F_i$ values labelled 'CA' are California-specific; the other values are nationwide estimates. \bar{S}_i for household products in SHEDS (one unfilled bar) and the total \bar{S}_i and $\bar{E}F_i$ (the rightmost bar cluster) values are $\times 5$ higher than the y axis as indicated by the $\times 5$ label; the asterisks indicate that data are not available. Also note that \bar{S}_i estimates are not shown for the California component of the 2011 NEI (only $\bar{E}F_i$ values are available for the California component). **b**, PM_{2.5} yields by mass organic gases reacted estimated by McDonald et al.¹ and upper bound estimates for prompt formation in the CMAQ modelling system. The PM_{2.5} yield of 5% by mass of organic gases was adopted in the CMAQ feasible solution; however, a PM_{2.5} yield of 10% (represented by the star on the far right) also resulted in PM_{2.5} SOA concentrations consistent with measurements.

product used that consists of carbon-containing compounds), (2) emission to air (described by a population-normalized emission factor, $\bar{E}F_i$; see Methods) and (3) transformation pathways leading to secondary products. Existing near-field chemical exposure models (for example, the High-Throughput Stochastic Human Exposure and Dose Simulation, SHEDS-HT, hereafter called SHEDS (ref. ³⁰)) help constrain the first two items, while far-field ambient chemical transport models (for example, the Community Multiscale Air Quality (CMAQ) modelling system) require the latter two as input. Thus, ambient air quality modelling, when evaluated against observed air concentrations, can improve the data, assumptions and methods on which near-field chemical exposure models rely.

Organic product usage and emissions to air. A first step in constraining the contribution of VCPs to air pollution is better quantification of their air emissions to the environment. Figure 1a compares different ambient air emission estimates (filled bars, see also Supplementary Tables 1 and 2 and Methods). In the United States, the evaporation of chemicals from organic products is estimated in the Solvent Sector as part of the US Environmental Protection Agency (EPA) National Emissions Inventory (NEI) and EPA Solvent Tool. Figure 1a indicates that these operational NEI methods result in a factor of ~ 3 – 4 lower emissions of VCP organic gases per person than a posteriori estimates (Methods) derived from the near-field SHEDS model³⁰ and values presented by McDonald et al.¹

Even though SHEDS and McDonald et al. use very different input data and approaches, the a posteriori SHEDS model and McDonald et al. arrive at approximately the same nationwide average air EFs of ~ 20 and ~ 24 kg per person per yr, respectively. The SHEDS-estimated values of $\bar{E}F_i$ for personal care are consistent with those of McDonald et al. (that is, 5.2 versus 4.5 kg per person per yr), and both methods estimate consumer adhesive emissions at ~ 2 kg per person per yr. However, SHEDS has the highest $\bar{E}F_i$ for household products, resulting in roughly similar amounts of organic gas emissions as personal care, while McDonald et al.

indicates that personal care emissions are the highest among VCP subsectors ($2\times$ higher than household cleaning). The $\bar{E}F_i$ values for California are generally lower than the nationwide ones in both the 2011 NEI (4.7 versus 8.2 kg per person per yr) and McDonald et al. (10.8 versus 24.3 kg per person per yr) estimates, potentially reflecting differences in down-the-drain adjustments applicable to certain categories³¹ and more stringent regulations for coatings and adhesives in California^{32,33}. Summing the per capita EFs for all sectors available in SHEDS and multiplying by population results in estimates of total residential (excluding industrial and commercial) consumer product emissions of organic gases to air of 1.4 Tgyr^{-1} in the 2011 NEI, compared with 6.1 Tgyr^{-1} ($\sim 4\times$) in SHEDS and 3.9 Tgyr^{-1} ($\sim 3\times$) in McDonald et al. for the United States. In addition to these numbers that suggest the NEI is low on a nationwide basis, we will show that an increase in the California component of the NEI to effectively 14 kg per person per yr (CMAQ feasible solution; the purple filled bar on the right in Fig. 1a) also results in improved estimates of criteria pollutants and other air quality endpoints in California, as predicted by a regional-scale chemical transport model.

Differences in organic product usage drive the differences in high versus low emission estimates. The Solvent Tool and EPA NEI methodology assume complete volatilization and air release of organic components in products ($\bar{S}_i = \bar{E}F_i$) yet have lower $\bar{E}F_i$ values than other methods that assume only a fraction of the organic product evaporates (Fig. 1a). This means that lower organic product usage (the total height of the bars in Fig. 1a) in the EPA national-level methodology results in lower $\bar{E}F_i$ values than in SHEDS and McDonald et al. These discrepancies in organic product usage are particularly large for sectors such as personal care, household products and consumer adhesives, with SHEDS and McDonald et al. indicating ~ 10 – $20\times$ higher organic product usage than the NEI methods. Total organic product usage for the VCP sector is estimated at ~ 75 , ~ 42 and ~ 8 kg per person per yr in the SHEDS, McDonald et al. and 2011 NEI methodologies, respectively.

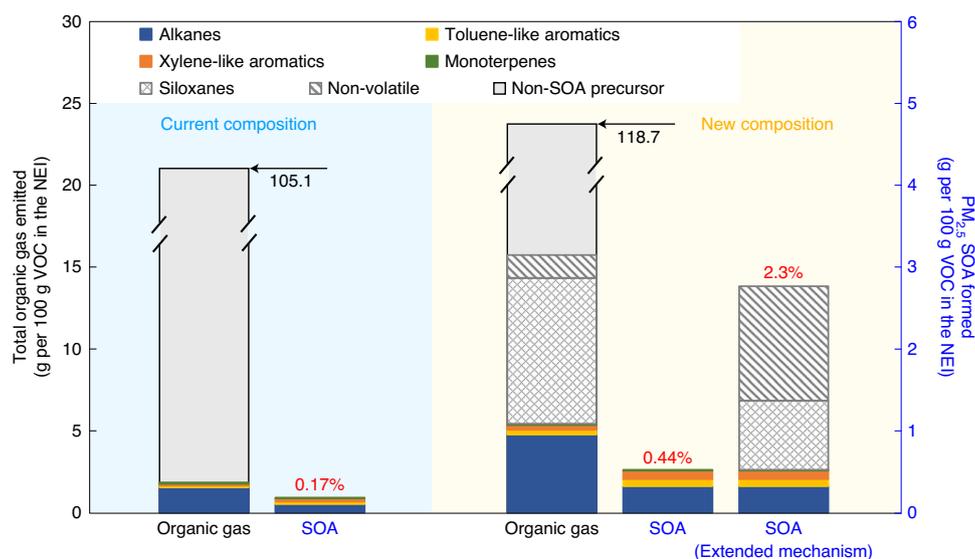


Fig. 2 | Organic gas emissions and PM_{2.5} SOA formation potentials for personal care products based on current and new emission composition, respectively, using parameters in CMAQ. The blue shaded portion indicates that the current operational method was used to specify emission composition, while updated information was used for the yellow panel. Emission values (in g) that do not correspond to the left y axis are displayed in black text. PM_{2.5} yields by mass of organic gases are indicated with red figures.

These differences indicate that constraining the organic product usage is of high priority for creating robust emission estimates.

The SHEDS and McDonald et al. approaches result in high air emissions, and volatilization assumptions modulate the magnitude of emissions. In this work, organic species in SHEDS were volatilized into air on the basis of species vapour pressure and nature of application, resulting in an effective volatilization fraction (γ , see Methods) of 27% in a posteriori estimates (Supplementary Table 3). For household products, a large fraction of the organic species (~86% in McDonald et al. and ~89% in SHEDS) were predicted to go down the drain, consistent with the types of product applications in that category, such as laundry detergent and cleaning products. SHEDS (a posteriori) indicates that half of the organic compounds in personal care products are ultimately emitted to air. This volatilization fraction is lower for SHEDS ($\gamma=0.48$) than for McDonald et al. ($\gamma=0.73$), reflecting different assumptions about the types of applications and/or composition of personal care products. In most other categories, the down-the-drain adjustment is less than a factor of 2 (that is, $0.5 < \gamma < 1$). Looking across all VCP categories available, ~58% and ~27% of the mass evaporated, according to the McDonald et al. and SHEDS methodologies, respectively.

A posteriori SHEDS estimates show an increased importance of inhalation exposure versus dermal uptake and/or ingestion from consumer product usage compared with a priori estimates before the modifications described in the Methods. The value of γ was increased from 2% in the a priori SHEDS chemical exposure model to 27% a posteriori, by making more reasonable assumptions for chemical mass retained on surfaces and skin (Methods). These results illustrate that volatilization assumptions, particularly for dermally applied personal care products, can substantially affect air release estimates. The 10 \times increase in γ between a priori and a posteriori implies a 10 \times increase in post-use inhalation exposure. Using Isaacs et al.³⁰ as a baseline and factoring in the a posteriori volatilization, the inhalation pathway is now estimated to be responsible for 2% of the mean daily intake dose for volatile chemicals in the near-field environment, making it competitive with previous estimates of dermal (7% of dose) and object-to-mouth (3% of dose) but not hand-to-mouth (near 90% of dose) exposure³⁰. This estimate of the role of inhalation in near-field chemical exposure applies

to consumers in residential buildings, and the relative importance of inhalation exposure could be higher for populations in other environments (for example, commercial or industrial settings). Volatilization should be further explored to more consistently represent the importance of chemicals in both chemical exposure and ambient air quality applications.

Speciation and PM_{2.5} yields in a regional air quality model. Since VOCs can be precursors for secondary PM_{2.5}, another step in determining the role of VCPs in air pollution is to identify transformation pathways leading to secondary products. The predicted ambient air PM_{2.5} yield depends on the composition of the emissions and the yields of individual species. Individual species in regional chemical transport models such as EPA's CMAQ modelling system are represented using simplified chemical mechanisms (Methods), which typically use a reduced number of surrogates to represent VOCs. Existing simplified chemical mechanisms were not developed to provide a robust representation of VCP chemistry (Fig. 2).

The challenges associated with representing PM_{2.5} are illustrated in Fig. 1b, which shows the prompt (rapidly formed, within several hours) sector-specific PM_{2.5} yields based on a current model system that employs these surrogate methods (see also Supplementary Tables 4–6). The PM_{2.5} yields for VCP subsectors in CMAQ vary by an order of magnitude (Supplementary Table 7), ranging from 0.17% for personal care products to as much as 16% for some industrial coatings, with an overall effective value of about 3.5% by mass in California (Methods). The CMAQ overall PM_{2.5} yield for VCPs is slightly lower than the estimate from McDonald et al. (that is, 4.6%). However, if the VCP sector is required to fully close the CMAQ model-measurement gap in PM_{2.5} SOA, then an even higher yield of 10% is required (discussed in the next section).

If the differences between the CMAQ and McDonald et al. (Fig. 1b) PM_{2.5} yield estimates are considered, in combination with the previously discussed emission magnitude discrepancies between SHEDS/McDonald et al. and the NEI (Fig. 1a), then adhesive, pesticide and personal care product uses account for the biggest potential underestimations in nationwide predictions of PM_{2.5} mass formation from VCPs at 75 Ggyr⁻¹, 40 Ggyr⁻¹ and 35 Ggyr⁻¹, respectively (Supplementary Table 7). This total missing PM_{2.5} from these three

subsectors (140 Ggyr^{-1}) is comparable to SOA formation from mobile sources ($\sim 170 \text{ Ggyr}^{-1}$) and accounts for 50% of VCP SOA ($\sim 300 \text{ Ggyr}^{-1}$). The potential underestimate of organic $\text{PM}_{2.5}$ from pesticides and adhesives is driven almost entirely by the emissions magnitude, while the aerosol yield ($15\times$ higher) causes the difference for personal care products. As a result, personal care is probably the category where $\text{PM}_{2.5}$ underestimates are driven largely by an incorrect characterization of VOC emission composition across the entire ambient air quality modelling system.

Figure 2 illustrates how the $\text{PM}_{2.5}$ yield from personal care emissions is underestimated in CMAQ and how it can be improved. A small fraction of the emitted organic gases ($\sim 1.7\%$) are mapped to traditional $\text{PM}_{2.5}$ SOA precursors including alkanes, aromatics and monoterpenes. Default emission processing and mapping results in an effective $\text{PM}_{2.5}$ mass yield of 0.17 g per 100 g of organic gases. SOA from personal care products in CMAQ is formed mainly from alkane oxidation, which contributes $\sim 57\%$ of the $\text{PM}_{2.5}$, followed by aromatic oxidation ($\sim 38\%$).

Resolving issues associated with the chemical composition of consumer product emissions, especially compounds with lower volatility that are poorly represented in the current CMAQ mechanism model surrogates³⁴, can increase the $\text{PM}_{2.5}$ yield for personal care products by a factor of 13 (from 0.17% to 2.3% by mass of organic gases, the fifth bar in Fig. 2). Updating VOC speciation profiles on the basis of more recent surveys of consumer products from 1997 to 2010 in California³⁵ results in a $\sim 3\times$ increase in amount of traditional SOA precursors.

The biggest cause of the underestimated $\text{PM}_{2.5}$ SOA formation from personal care VCPs is the lack of surrogate species representing low-vapour-pressure species with high $\text{PM}_{2.5}$ formation potential. For instance, *n*-heptadecane, a compound present in personal care products with a $\text{PM}_{2.5}$ yield of $>30\%$ (ref. 19), is mapped to an alkane surrogate that has low SOA formation ($\sim 3\%$ by mass in CMAQ). Additionally, substantial amounts of emitted organic gases are mapped to lumped species labelled as “non-reactive organic gases” and “nonvolatile” species (probably semivolatile or intermediate-volatility organic compounds considered ineffective ozone precursors) that are not considered in CMAQ regional model calculations of ozone or $\text{PM}_{2.5}$. Accounting for contributions of “non-reactive organic gases”, such as siloxanes^{22,36}, and treating the “nonvolatile” species as precursors, such as *n*-pentadecane with a $\sim 50\%$ (refs. 37,38) (lower bound) or 100% $\text{PM}_{2.5}$ yield (upper bound, Fig. 2), increases the $\text{PM}_{2.5}$ yield estimated for personal care products by a factor of 4–5 (from 0.44% to as much as 2.3%), suggesting that chemical mechanisms need to be extended to include these non-traditional species.

Constraining emissions and $\text{PM}_{2.5}$ yields with ambient observations.

Because the SHEDS and McDonald et al. inventories suggest that the 2011 NEI underrepresents VCPs, emission adjustments and $\text{PM}_{2.5}$ yields were implemented in CMAQ to investigate what combinations of parameters improved predictions of ozone, $\text{PM}_{2.5}$ and other metrics for the Los Angeles basin. CMAQ-simulated ambient $\text{PM}_{2.5}$ SOA concentrations in Pasadena, California, are driven by both the VCP emission magnitude and the $\text{PM}_{2.5}$ yield, as reflected by the normalized mean bias (NMB) between the measured and predicted SOA (Fig. 3). The current estimate of VCP emissions in the 2011 NEI (California) and CMAQ model (that is, VCP emissions of 4.7 kg per person per yr + effective $\text{PM}_{2.5}$ yield of $\sim 3.5\%$) results in an underestimation of photochemically driven prompt SOA of 53% ($\text{SOA}_{\text{prompt}} = \text{SOA}$ with background contributions removed; see Methods). Using the existing SOA parameterization and simply increasing VCP emissions across all subsectors by factors of 2–4 increases SOA production and improves model performance. We consider the 14 kg per person per yr of VCP emissions ($3\times$ NEI (California) emissions; $1.7\times$ higher than the 2011 NEI

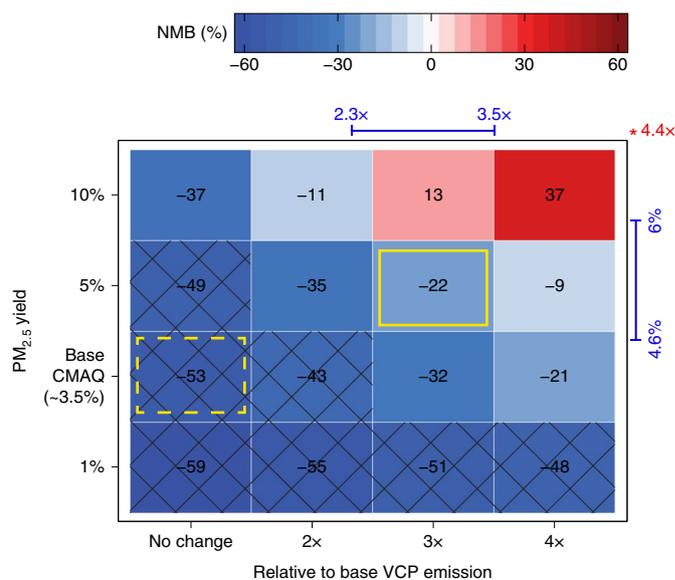


Fig. 3 | NMB between CMAQ-predicted and observed $\text{SOA}_{\text{prompt}}$ in Pasadena, California, for different combinations of VCP emission magnitude and VCP $\text{PM}_{2.5}$ yield adjustments.

The cross hatching indicates that the parameters are inconsistent with observations (specifically, $\text{NMB} \geq$ uncertainty in measured OAs, that is, 38% (ref. 67)). The recommended emissions (top) and $\text{PM}_{2.5}$ yields (right) in McDonald et al. are marked by the range indicated in the margin. The recommended emission magnitude based on SHEDS (consumer emissions only) is indicated with the red asterisk (top). Two cases, the current model (dashed) and feasible solution (solid), are indicated with yellow boxes. The base VCP emissions in the California component of the 2011 NEI are 4.7 kg per person per yr.

nationwide total $\overline{\text{EF}}_i$ estimate) and a 5% $\text{PM}_{2.5}$ yield as a plausible far-field modelling representation of the VCP sector (feasible solution). Yields could be higher or lower in combination with lower or higher emissions, and the feasible solution still underestimates SOA by 22%. With 2011 NEI emissions, some bias remains even with a $\text{PM}_{2.5}$ yield of 10% by mass from organic gases. Additional $\text{SOA}_{\text{prompt}}$ sources as well as background sources, probably of biogenic or non-fossil origin, may still be missing in CMAQ (Extended Data Fig. 1).

The feasible solution case, supported by the near-field SHEDS model emission estimates and updated VOC speciation profiles, demonstrates that VCPs probably play an important role in $\text{PM}_{2.5}$ SOA formation. The 14 kg per person per yr VCP emissions with a $\text{PM}_{2.5}$ yield of 5% per mass organic gas also substantially improve the predicted diurnal variation of prompt SOA with a peak concentration of $4.7 \mu\text{g m}^{-3}$ (versus $2.4 \mu\text{g m}^{-3}$ in the operational CMAQ modelling system) at Pasadena (Fig. 4). However, the simulated SOA was still lower than the observed $\text{SOA}_{\text{prompt}}$, which reached a diurnally averaged maximum of $6.6 \mu\text{g m}^{-3}$. The base CMAQ system estimated that VCPs contributed $0.2 \mu\text{g m}^{-3}$ of $\text{PM}_{2.5}$ SOA, compared with $1.0 \mu\text{g m}^{-3}$ from non-VCPs (Supplementary Table 8), which means that VCPs are not a major source of $\text{SOA}_{\text{prompt}}$ (only $\sim 20\%$) in the current models. In the feasible solution case, VCPs are predicted to contribute $1.1 \pm 0.3 \mu\text{g m}^{-3}$ of $\text{PM}_{2.5}$, making up 53% of the simulated total $\text{SOA}_{\text{prompt}}$ ($2.1 \mu\text{g m}^{-3}$) and $\sim 41\%$ of the observed $\text{SOA}_{\text{prompt}}$ at Pasadena ($2.6 \mu\text{g m}^{-3}$). The strong daytime increase in SOA requires higher $\text{PM}_{2.5}$ yields regardless of the emission magnitude (Extended Data Fig. 2).

VCP use results in large amounts of organic gases being emitted into ambient air and subsequently affects oxidant levels in addition to $\text{PM}_{2.5}$. In the feasible solution case, VCPs account for ~ 3.7 ppb

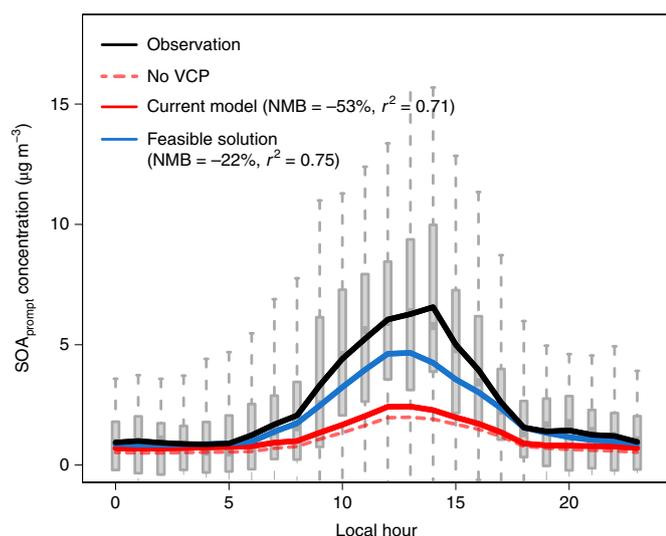


Fig. 4 | Diurnal variation of simulated and observed SOA_{prompt} at Pasadena, California. The simulation with VCP emissions zeroed out is indicated by the dashed line. The grey boxes indicate the 25th and 75th percentiles of the observations, with the median values marked with horizontal lines in the boxes, and the whiskers extend to the 10th and 90th percentiles. The solid lines indicate the means of the observations and simulations.

of organic gases in ambient air during the daytime (10:00–18:00). Increased organic gas emissions from VCPs trigger modest increases in hydroxyl radical (OH) and O₃ concentrations in the model (Extended Data Fig. 3 and Fig. 5a), which results in faster hydrocarbon reactions and more SOA. Increased VCP emissions increase the high bias in simulated OH concentrations (NMB of 29% versus 21% in the current model), but still within the uncertainty of the measurement (that is, 36%)³⁹. Organic gas emissions from VCPs also contribute 9 ± 2 ppb to the maximum daily 8 hr average (MDA8) O₃ with the feasible solution (versus 3 ppb in the current model; Fig. 5a). The higher VCP emissions improve the model simulations of O₃ compared with observations, resulting in the feasible solution regression line overlapping the 1:1 line at Pasadena.

OH reactivity is a measure of how quickly OH is lost via reaction with organic gases and is thus useful for assessing uncertainties in VOC emissions and examining their importance to O₃ and PM_{2.5} formation^{40–43}. OH reactivity was enhanced in the feasible solution case (15.5 s⁻¹ versus 12.8 s⁻¹ in the current model) and was brought to within 30% of the observation (21 s⁻¹, ref. ³⁹) with 3× the NEI (California) VCP emissions (Fig. 5b). OH reactivity contributed by VCPs (3.9 s⁻¹ in the feasible solution) is higher than predicted in the default CMAQ modelling system but lower than 4.8 s⁻¹ in McDonald et al. The OH reactivity of 4.5–4.7 s⁻¹ is explained by organic gases from non-VCP sources (such as mobile sources and biogenics). Additionally, inorganic species account for ~6.9 s⁻¹ of OH reactivity, higher than the observation-based value (5.8 s⁻¹) due to over-estimated NO₂ concentrations in the CMAQ simulations. Overall OH reactivity has better agreement with measurements with higher VCP emissions, but some missing OH reactivity in CMAQ remains to be addressed.

Integrating near-field and far-field exposure models. Near-field chemical exposure (such as SHEDS) and far-field air quality (such as CMAQ) models share common goals in terms of providing estimates of human (and ecosystem) exposure to pollutants. When connected through the CMAQ model, ambient air concentrations of OA, ozone and OH reactivity provide top-down constraints on air emissions associated with VCP use and suggest an EF of ~14 kg per

person per yr in California (feasible solution). This value is consistent with a posteriori SHEDS calculations of air emissions of ~20 kg per person per yr. SHEDS indicates that this ambient air emission rate is associated with ~75 kg per person per yr of organic product usage, notably higher than the current EPA national-level methodology, which indicates only 8.2 kg per person per yr. In addition to using ambient air concentrations to constrain emissions and organic product usage in near-field modelling, our work suggests additional avenues to better estimate exposure across scales. For example, the EPA has undertaken a major effort to curate hundreds of thousands of chemical structures via the CompTox Chemistry Dashboard⁴⁴; however, over 250 species in the ambient air emission speciation database used here (Methods) were not found in the Dashboard (as of 9 September 2019), although chemical identifier curation is ongoing. The curation of species emitted in air in the Dashboard will allow for quantitative structure–activity relationship models developed for chemical properties⁴⁵, toxicity⁴⁶ and exposure²⁶ to be applied and facilitate risk evaluations of these compounds. In addition, models like SHEDS can be used to create new speciation profiles for ambient air quality modelling or provide alternative EFs for use in creating VOC inventories. Ultimately, connecting near-field chemical exposure and far-field air quality models that consider direct exposure to emitted chemicals as well as their oxidation products (for example, PM_{2.5} and ozone) will result in improved tools for addressing public health questions.

Discussion

VCPs are ubiquitous in daily life for aesthetic, protective, health-care and other purposes. Because VCPs have the potential to directly expose people to harmful chemicals as well as escape to ambient air, their formulations have continuously evolved over decades in response to air quality and chemical safety drivers. Since the 1990s, VCPs have been targeted for control to reduce the formation of ozone as well as hazardous air pollutants in the United States via the Clean Air Act. This has led to an increase in the prevalence of water and other green solvents in VCPs. Product formulations will continue to evolve to improve performance, cost and regulatory compliance via the use of functional substitutes as chemicals with lower health risks and less potential for criteria pollutant formation are identified. However, designing product reformulations aimed at improving public health and the environment will require a wholistic perspective that accounts for multiple pathways of exposure to both the compounds themselves and their degradation products, which is not fully realized in current alternative assessment tools or air quality management. For example, a narrow focus on ozone has led to compounds such as siloxanes being exempt from VOC regulation despite their high potential for SOA formation. Multiple options, from solvent substitution to product capture and recycling, should be employed to reduce the environmental and health impacts of VCPs.

Methods

In this work, air emissions of VCPs were estimated with multiple inventory methods and a near-field chemical exposure model. The impacts of VCP emissions on ambient criteria pollutants were evaluated via air quality modelling, taking uncertainties in emissions and the model representation of pollutant formation into account (Extended Data Fig. 4).

VCP emission estimates for ambient air. The use of VCPs results in evaporative emissions of organic gases to the atmosphere. Air emissions from VCP usage (E) can be estimated on the basis of sector-specific EFs (EF_i) and activity data (A_i) for sector i :

$$E = \sum_i EF_i \times A_i \quad (1)$$

Emissions can be estimated using different activity surrogates such as employment, population or product use in a sector with appropriate EFs (Supplementary Table 2). If population-based EFs are reported by the method

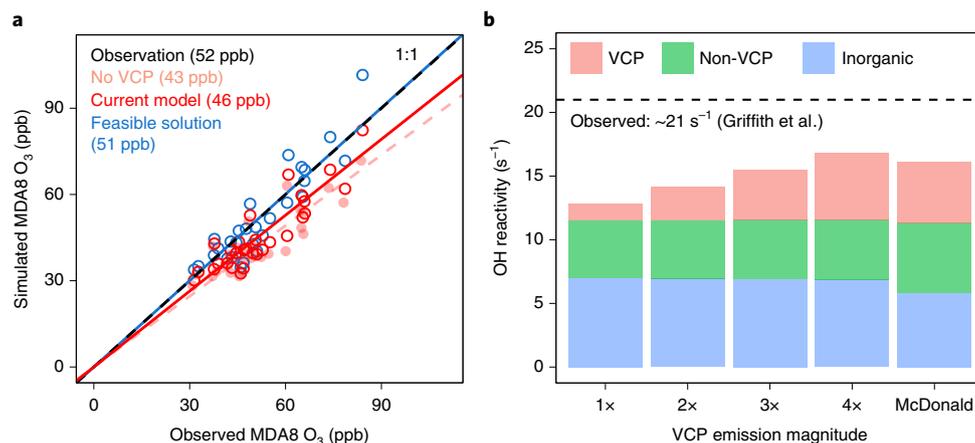


Fig. 5 | Simulated MDA8 O₃ and OH reactivity versus observations at Pasadena. a,b, MDA8 O₃ (**a**) and OH reactivity (**b**) with different VCP emission magnitudes. In **a**, the simulation with VCP emissions zeroed out is indicated with pink, and the averaged MDA8 O₃ concentrations are given in the parentheses.

(such as Solvent Tool and SHEDS), they are directly assigned to \overline{EF}_i (Fig. 1a). If population-based EFs are not provided, the values of \overline{EF}_i (in units of kg per person per yr) are computed using E divided by A_i in the form of population (using 312 million as the 2011 US population and 3.8 million for California). These \overline{EF}_i values quantify how many kilograms of organic gas are released to ambient air per person each year. EFs may consider how organic product usage (S_i) (mass per person per year for a population-based activity surrogate, that is, \overline{S}_i) is attenuated by washing down the drain (assumed here to be a permanent sink of VOCs) or otherwise sequestered from air. This down-the-drain adjustment is represented by γ_i , which represents the fraction of the organic product (generally a solvent) released into ambient air ($0 \leq \gamma_i \leq 1$):

$$EF_i = S_i \times \gamma_i \quad (2)$$

The work here is focused exclusively on the organic content of consumer products that can be released into the atmosphere. The water, electrolytes and other inorganic components of consumer products are not explicitly discussed or included in S_i , γ_i or EF_i . In this work, five different methods (see also Supplementary Table 2) for calculating VCP EFs are compared in terms of total sector emissions and category-specific emissions. These methods are described in the following sections.

The 2011 EPA NEI. The EPA NEI is compiled on the basis of air emission estimates developed by state and local agencies and submitted to the EPA, as required by the Air Emissions Reporting Requirements Rule 40 CFR § 51.1008 (2019). In the cases where state or local agencies do not provide emission estimates, the EPA provides them as a fallback⁴⁷. The California VCP emission estimates in the NEI are contributed by the state agency using its own methodology ('The California emission inventory'). The 2011 EPA NEI estimates are national-level VOC emission totals aggregated up from state-level reports (including not only non-point-source but also point-source emissions, which are available only for industrial sectors such as industrial surface coating, graphic arts and degreasing, and may use methods from individual jurisdictions or EPA methodology). The EPA NEI Solvent Sector emissions are used synonymously with VCP emissions in this work.

The fallback 2011 EPA NEI methodology is based on national-level estimates of solvent use from the Freedonia 2007 report⁴⁸ divided by national-level activity data such as population and employment, as well as some other sources (for example, the EPA Emissions Inventory Improvement Program technical report⁴⁹). The default EPA NEI methodology assumes complete volatilization of the organic components ($\gamma = 1$). The NEI inventories VOCs according to a strict definition based on what is likely to lead to ozone formation (40 CFR 51.100). Conversion between the inventoried VOCs and total organic gases (which may include non-reactive or other species) is based on other information (for example, SPECIATE; see below). The difference in regulatory VOCs versus total organic gases is small (<30% effect) compared with the differences between the 2011 NEI/Solvent Tool and other methods, which are generally on the order of a factor of 3 (Supplementary Table 9 and Fig. 1a).

The Solvent Tool. The Solvent Tool v.1.7 uses the same methods as the 2011 NEI and has been distributed by the EPA since the development of the 2014 NEI to provide a way for states to estimate emissions of VOCs from VCPs. EFs in the Solvent Tool v.1.7 for some categories were updated from 2011 using the updated

Freedonia 2013 report⁵⁰. The differences in EFs between the 2011 NEI and the Solvent Tool are negligible (Fig. 1a). EFs in the Solvent Tool apply to non-point emissions only⁵⁰.

The California emission inventory. The California Air Resources Board (CARB) estimates both total organic gas and VOC emissions from VCPs using the sales of products from consumer product surveys from product manufacturers and formulators, multiplied by the weight fraction of each organic species (VOC) compound (that is, the VOC content of the products) according to the speciation data from the survey⁵¹. The aggregated category (such as hair sprays or hand and body lotion) total organic gases (VOCs) are computed as a sum over each individual product organic gases (VOCs) total, followed by adjustment with values of γ in certain categories ($0 < \gamma < 1$).

The SHEDS model. The SHEDS model v.0.1.6 (ref. ³⁰) and its default inputs were used to estimate the residential release of VOCs from formulated consumer products. Ambient air quality (for example, CMAQ) models generally target the actual atmospheric mean state, while exposure models (for example, SHEDS) are designed to inform the risk-based prioritization of chemicals via conservative (upper-bound) exposure estimates for individual species^{26,30}. SHEDS uses consumer habits and practices (product prevalence in the marketplace, frequency of use and mass per use) for 324 individual product categories, and predicts the uses of these products for a population of simulated individuals representative of the United States in terms of age and gender (a standard population size of 10,000 was used here). Releases of individual chemicals were estimated by multiplying the mass of the product used by the median weight fractions from Material Safety Data Sheets obtained from the EPA's Chemical and Products Database⁵¹ and by values of γ (see below). Mean releases were estimated for all volatile chemicals for which weight fraction data were available. Organic gas emissions for general categories (Supplementary Table 10) were calculated by summing over all chemicals in the category. All volatilized product was assumed to be released into ambient air without considering air exchange rates or loss in the indoor compartment.

We use the terminology *a priori* to refer to SHEDS predictions using the standard configuration described by Isaacs et al.³⁰. Two critical adjustments were made to standard SHEDS operations (that is, *a priori* estimates), resulting in the *a posteriori* estimates shown in Fig. 1a: (1) the product prevalence in the marketplace was adjusted to be more mass conservative across all VOCs and (2) γ was estimated as a function of species vapour pressure and type of product application. By default, SHEDS estimates exposure to individual compounds, and if a species is present in one product, the model will always choose a non-zero value for that species' occurrence in that product type (100% chemical prevalence). For example, if only one shampoo contains siloxanes, that product is assumed to have 100% of the market share when siloxane exposure is being estimated. Similarly, if a different shampoo is the only one that contains polyethylene glycol, that shampoo will have 100% of the market share when polyethylene glycol exposure is being estimated. This assumption is made due to the lack of true knowledge of the fraction of products on the market containing the chemical and the market share of those products. However, it can result in the overcounting of mass when individual species are aggregated into a VOC or organic gas total. To remedy this, we opted to randomly sample product compositions, even across products in the Chemical and Products Database in which a species was not present (zero). The species were then aggregated up to a total. Furthermore, while SHEDS applies down-the-drain assumptions for products applied to residential and body surfaces, chemicals

applied dermally do not volatilize by default. Volatilization was parameterized on the basis of the type of usage (for example, applied to surfaces versus used in an aqueous environment) and the volatility of organic species (following the methodology of McDonald et al., Supplementary Table 11) to obtain emissions to ambient air.

Recent estimates in literature. McDonald et al.¹ estimated total organic gas emissions from VCPs using the annual sales of products from economic and commodity flow surveys, the organic content usage in each product based on energy and chemical production statistics, and the volatilization fraction of the organic content based on indoor air quality literature and emission control efficiency.

VCP emission processing in air quality models. The composition of VCP emissions (referred to as speciation) in air quality modelling is based on sector-specific speciation profiles from the EPA's SPECIATE database³². This database provides information to convert inventoried VOCs to specific individual chemical species. SPECIATE v.4.5 is the public version of the database in operational use at the time of this work, and consumer product composition is based mainly on the CARB 1997 consumer products survey. For SPECIATE v.5.0 (ref. 35) (released in summer 2019), the EPA updated the profiles for consumer products and architectural coating using CARB's 2010 consumer products inventory products update, which is based mainly on the 2003 and 2006 CARB consumer product surveys. Current and new emission compositions in Fig. 2 are based on SPECIATE v.4.5 and SPECIATE v.5.0, respectively. The speciation profiles for emission sources such as consumer products and architectural coatings, derived from CARB profiles, include a number of species (for example, bins) that are not defined in SPECIATE. Those were mapped to compounds or mixtures in SPECIATE v.5.0 using Carter's assignments³³ to avoid loss of mass. In addition, we updated the exempt status of VOCs relevant to VCPs in SPECIATE v.5.0 so that the ratio of total organic gases to inventoried non-exempt VOCs is better characterized. Mappings of chemical composites to model-ready surrogates, which are chemical-mechanism-dependent, follow Carter^{34,35}.

CMAQ model configurations and inputs. The CMAQ model is an ambient air chemical transport model used for research and regulatory purposes. For example, it is used by state and local agencies to design State Implementation Plans and other air quality management strategies. In this work, the CMAQ model uses VCP emissions as input to predict their impact on ambient concentrations of ozone, SOA and OH reactivity. CMAQ model v.5.3Beta (www.epa.gov/cmaq, <https://github.com/USEPA/CMAQ>) was used to perform simulations in California (Supplementary Fig. 1) with a horizontal resolution of 4 km and 35 vertical layers extending up to the 50 mbar level in the atmosphere. The simulation spanned the California Nexus of Air Quality and Climate Change (CalNex) campaign³⁴ (15 May to 15 June 2010) with the prior 14 days as a spin-up. The Pasadena ground site was located 18 km northeast of downtown Los Angeles and impacted by emissions in the Los Angeles metropolitan area with prevailing southwesterly winds during the daytime (advection of 3–5 h)¹⁵. The SAPRC07ic chemical mechanism (see the details in Pye et al.³⁶ and Xie et al.³⁶) was used to represent gas-phase chemistry, with AER07 for aerosol chemistry^{37–39}. The meteorological inputs were provided by Weather Research and Forecasting v.3.8.1, with model configurations identical to those in Lu et al.⁶⁰

The anthropogenic emissions used in CMAQ were based on the 2011 NEI, with a 25% reduction of mobile NO_x to achieve more consistency with the observed NO₂ at the Air Quality System sites within the domain (Supplementary Fig. 2). The biogenic emissions were calculated online using the Biogenic Emission Inventory model v.3.61, with the Biogenic Emission Landuse Database 4 and multiple data sources for vegetation representation including the US Forest Service Forest Inventory and Analysis v.5.1 for 2002–2013 and the US Department of Agriculture 2007 census of agriculture data⁴¹.

Treatment of primary OA and SOA from non-VCPs in CMAQ. Direct emissions of organic PM_{2.5} (primary OA) for area-wide anthropogenic emission sectors (gasoline vehicles, diesel vehicles, aircraft, cooking, biomass burning, dust, VCPs and other) were characterized as semivolatile with sector-specific volatility profiles for the mobile⁶⁰ and non-mobile¹⁴ sources. Atmospheric processing of primary OA is the same as described in Murphy et al.³⁷. A state-of-the-science PM_{2.5} SOA parameterization for mobile organic gases with intermediate volatility was implemented in CMAQ v.5.3, as in Lu et al.⁶⁰, in an effort to minimize uncertainty in model predictions of the mobile-sector SOA. Other sources of SOA in the base model include the oxidation of traditional VOCs including isoprene, monoterpenes, sesquiterpenes, aromatics and alkanes. For computational efficiency, SOA from VOCs of primarily anthropogenic origin (benzene, toluene, xylene, polycyclic aromatic hydrocarbons and alkanes) were parameterized in CMAQ v.5.3-AER07 as described in Supplementary Tables 4 and 5.

Treatment of SOA from VCPs in CMAQ. In addition to a base CMAQ v.5.3 simulation, a series of simulations were performed (Supplementary Table 12) with VCP organic gas emission perturbations (1×, 2×, 3× and 4× relative to the base

VCP emissions in the 2011 NEI). Three levels of fixed PM_{2.5} yields (α in Reaction 1) including 1%, 5% and 10% by mass were explored in concert with the emission perturbations. Due to the diversity of VOCs emitted and created in the atmosphere (Goldstein and Galbally⁶² estimated that 10⁴–10⁵ different species have been identified), lumped mechanisms (for example, SAPRC07) are created to represent the major routes to ozone and PM_{2.5} formation with a reduced number of organic species (usually ~100 surrogates). SOA yields for VCP subsectors in Figs. 1b and 2 were calculated offline (that is, not conducting CMAQ modelling) by mapping the emission composition to SAPRC07 surrogates with the surrogate-specific yields in CMAQ (Supplementary Table 6). Simulated SOA_{prompt} with base CMAQ v.5.3 was equivalent to a simulation with a fixed PM_{2.5} yield of ~3.5%. Organic gases from VCPs participated in gas-phase chemistry leading to ozone formation and oxidant modifications through traditional SAPRC mechanism species (that is, benzene, toluene, xylene, monoterpenes, etc.) as indicated by SPECIATE 4.5, while the SOA formation from VCPs was imposed by Reaction 1. In Reaction 1, all VCP organic gases were treated using a single surrogate, TOG_{VCP}, so that a specific yield ($\alpha = 1\%$, 5% and 10%) could be implemented. The reaction rate constant (k_{OH}) for VCP organic gases reacting with OH in Reaction 1 is identical to the averaged k_{OH} for mobile organic gases with intermediate volatility in CMAQ (ref. 60)—that is, $k_{OH} = 2.18 \times 10^{-11} \text{ cm}^3 \text{ per molecule per second}$, which is slightly higher than the McDonald et al. estimate of $1.25 \times 10^{-11} \text{ cm}^3 \text{ per molecule per second}$ (ref. 1).



CMAQ ambient air quality model evaluation. The CMAQ and Weather Research and Forecasting models were extensively evaluated for the CalNex domain for multiple pollutants (for example, ozone, total PM_{2.5} mass and constituents, and PM_{2.5} precursor vapours) in previous work^{14,16,57,60,63–65}. The evaluation in this study was designed to focus on rapidly formed photochemically driven SOA (that is, SOA_{prompt}), to isolate fresh SOA from VCPs in an urban atmosphere. Oxygenated OA (OOA), derived from factor analysis of OA data from the Aerodyne High-Resolution Time-of-Flight Aerosol Mass Spectrometer, is considered a measured proxy for SOA and can be further mathematically divided into two components: semivolatile OOA (SV-OOA) and low-volatility OOA (LV-OOA)¹². Generally, SV-OOA shows strong diurnal variation (peaking at 14:00–15:00) in Pasadena and is consistent with fresh SOA from urban areas, while LV-OOA (the more oxidized OOA) appears flat throughout the day and is expected to be aged SOA (ref. 66). The observed background OA was derived from the minimum in the LV-OOA diurnal circle¹⁵ and is probably associated with biogenic VOC oxidation and transport. Here, the average of the lower 50% of CMAQ-simulated SOA concentrations during 0:00–4:00 local time was used to estimate the background OA in simulations. SOA_{prompt} in both the model predictions and the observations, was derived by subtracting the background OA from the overall SOA concentration. The background adjustment could lead to negative values of prompt OA on an hourly basis, which has minor effects on diurnal cycles or period-averaged concentrations (Supplementary Fig. 3). A metric NMB was used for the evaluation of simulated SOA_{prompt} with CMAQ:

$$\text{NMB} = \frac{\sum_{i=1}^N (M_i - O_i)}{\sum_{i=1}^N O_i} \quad (3)$$

where M_i and O_i are the modelled and observed values at time i , respectively, and N is the number of hourly samples.

OH reactivity from VCP emissions in the simulations was estimated using the concentration of the surrogate TOG_{VCP} with single k_{OH} (Reaction 1). The values calculated with this method compared well with the values calculated using SAPRC07 speciation (Supplementary Fig. 4). The measured total OH reactivity was based on the observed decay of OH due to reactivity with ambient air in a turbulent-flow tube reactor.³⁹

Data availability

The VCP emission estimates in the form of population-based EFs are available in the Supplementary Information. The air quality modelling data for Pasadena during the 2010 CalNex campaign are available at <https://doi.org/10.23719/1506136>. The CalNex observations are publicly accessible at <https://www.esrl.noaa.gov/csl/groups/csl7/measurements/2010calnex/>. The full CMAQ outputs that support the findings of this study are archived on the EPA's high-performance computing system and can be obtained from the corresponding author upon request. Source data are provided with this paper.

Code availability

The source code of the operational CMAQ model is available at <https://github.com/USEPA/CMAQ>, with specific modifications applied in this work accessible at <https://doi.org/10.23719/1506136>. The code for the simulation data processing is available upon request from the corresponding author. The source code of the SHEDS-HT model is available at <https://github.com/HumanExposure/SHEDSHTPackage>. The modified SHEDS model including inputs and outputs from this study is available at https://github.com/HumanExposure/SHEDS_

Applications/. The Solvent Tool is available at ftp://newftp.epa.gov/air/nei/2014/doc/2014v2_supportingdata/nonpoint/Solvent_Tool_v1.7.zip.

Received: 18 November 2019; Accepted: 26 August 2020;
Published online: 5 October 2020

References

- McDonald, B. C. et al. Volatile chemical products emerging as largest petrochemical source of urban organic emissions. *Science* **359**, 760–764 (2018).
- Trantallidi, M. et al. Ephemeral III: health risk assessment of exposure to household consumer products. *Sci. Total Environ.* **536**, 903–913 (2015).
- Elberling, J. et al. Mucosal symptoms elicited by fragrance products in a population-based sample in relation to atopy and bronchial hyper-reactivity. *Clin. Exp. Allergy* **35**, 75–81 (2005).
- Gabb, H. A. & Blake, C. An informatics approach to evaluating combined chemical exposures from consumer products: a case study of asthma-associated chemicals and potential endocrine disruptors. *Environ. Health Perspect.* **124**, 1155–1165 (2016).
- Dodson, R. E. et al. Endocrine disruptors and asthma-associated chemicals in consumer products. *Environ. Health Perspect.* **120**, 935–943 (2012).
- Brinke, J. T. et al. Development of new volatile organic compound (VOC) exposure metrics and their relationship to “sick building syndrome” symptoms. *Indoor Air* **8**, 140–152 (1998).
- Pope, C. A. & Dockery, D. W. Health effects of fine particulate air pollution: lines that connect. *J. Air Waste Manage. Assoc.* **56**, 709–742 (2006).
- Schlesinger, R. B. et al. The health relevance of ambient particulate matter characteristics: coherence of toxicological and epidemiological inferences. *Inhal. Toxicol.* **18**, 95–125 (2006).
- Ebi Kristie, L. & McGregor, G. Climate change, tropospheric ozone and particulate matter, and health impacts. *Environ. Health Perspect.* **116**, 1449–1455 (2008).
- Zhang, Q. et al. Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes. *Geophys. Res. Lett.* **34**, L13801 (2007).
- Huang, R. J. et al. High secondary aerosol contribution to particulate pollution during haze events in China. *Nature* **514**, 218–222 (2014).
- Jimenez, J. L. et al. Evolution of organic aerosols in the atmosphere. *Science* **326**, 1525–1529 (2009).
- Ensberg, J. J. et al. Emission factor ratios, SOA mass yields, and the impact of vehicular emissions on SOA formation. *Atmos. Chem. Phys.* **14**, 2383–2397 (2014).
- Woody, M. C. et al. Understanding sources of organic aerosol during CalNex-2010 using the CMAQ-VBS. *Atmos. Chem. Phys.* **16**, 4081–4100 (2016).
- Hayes, P. et al. Modeling the formation and aging of secondary organic aerosols in Los Angeles during CalNex 2010. *Atmos. Chem. Phys.* **15**, 5773–5801 (2015).
- Baker, K. et al. Gas and aerosol carbon in California: comparison of measurements and model predictions in Pasadena and Bakersfield. *Atmos. Chem. Phys.* **15**, 5243–5258 (2015).
- Hodzic, A. et al. 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 (2010).
- Shrivastava, M. K. et al. Effects of gas particle partitioning and aging of primary emissions on urban and regional organic aerosol concentrations. *J. Geophys. Res. Atmos.* **113**, D18301 (2008).
- Li, W. et al. Potential of select intermediate-volatility organic compounds and consumer products for secondary organic aerosol and ozone formation under relevant urban conditions. *Atmos. Environ.* **178**, 109–117 (2018).
- Coggon, M. M. et al. Diurnal variability and emission pattern of decamethylcyclopentasiloxane (D5) from the application of personal care products in two North American cities. *Environ. Sci. Technol.* **52**, 5610–5618 (2018).
- Janeček, N. J. et al. Physical properties of secondary photochemical aerosol from OH oxidation of a cyclic siloxane. *Atmos. Chem. Phys.* **19**, 1649–1664 (2019).
- Wu, Y. & Johnston, M. V. Aerosol formation from OH oxidation of the volatile cyclic methyl siloxane (CVMS) decamethylcyclopentasiloxane. *Environ. Sci. Technol.* **51**, 4445–4451 (2017).
- Zhao, Y. et al. Intermediate-volatility organic compounds: a large source of secondary organic aerosol. *Environ. Sci. Technol.* **48**, 13743–13750 (2014).
- Karl, T. et al. Urban flux measurements reveal a large pool of oxygenated volatile organic compound emissions. *Proc. Natl Acad. Sci. USA* **115**, 1186–1191 (2018).
- Fantke, P. et al. Coupled near-field and far-field exposure assessment framework for chemicals in consumer products. *Environ. Int.* **94**, 508–518 (2016).
- Ring, C. L. et al. Consensus modeling of median chemical intake for the U.S. population based on predictions of exposure pathways. *Environ. Sci. Technol.* **53**, 719–732 (2019).
- Wambaugh, J. F. et al. High-throughput models for exposure-based chemical prioritization in the ExpoCast Project. *Environ. Sci. Technol.* **47**, 8479–8488 (2013).
- Barber, M. C., Isaacs, K. K. & Tebes-Stevens, C. Developing and applying metamodelling of high resolution process-based simulations for high throughput exposure assessment of organic chemicals in riverine ecosystems. *Sci. Total Environ.* **605**, 471–481 (2017).
- Nazaroff, W., Weschler, C. J., Little, J. C. & Hubal, E. A. C. Intake to production ratio: a measure of exposure intimacy for manufactured chemicals. *Environ. Health Perspect.* **120**, 1678–1683 (2012).
- Isaacs, K. K. et al. SHEDS-HT: an integrated probabilistic exposure model for prioritizing exposures to chemicals with near-field and dietary sources. *Environ. Sci. Technol.* **48**, 12750–12759 (2014).
- ARB solvent evaporation methodologies—*aerosol coatings & consumer products* (California Air Resources Board (CARB), 2020); https://ww3.arb.ca.gov/ei/areasrc/consumer_products_2020.pdf
- Rule 1113: *architectural coatings* (California Air Resources Board (CARB), 2013); <https://ww2.arb.ca.gov/sites/default/files/classic/technology-clearinghouse/rules/RuleID256.pdf>
- Rule 1168: *adhesive and sealant applications* (South Coast Air Quality Management District (SCAQMD), 2017); <http://www.aqmd.gov/docs/default-source/rule-book/reg-xi/rule-1168.pdf>
- Carter, W. P. Development of a database for chemical mechanism assignments for volatile organic emissions. *J. Air Waste Manage. Assoc.* **65**, 1171–1184 (2015).
- Final report SPECIATE version 5.0 database development documentation. Report EPA/600/R-19/098 (Abt Associates, 2019); https://www.epa.gov/sites/production/files/2019-07/documents/speciate_5.0.pdf
- Wu, Y. & Johnston, M. V. Molecular characterization of secondary aerosol from oxidation of cyclic methylsiloxanes. *J. Am. Soc. Mass. Spectrom.* **27**, 402–409 (2016).
- Jathar, S. H. et al. Unspecified organic emissions from combustion sources and their influence on the secondary organic aerosol budget in the United States. *Proc. Natl Acad. Sci. USA* **111**, 10473–10478 (2014).
- Presto, A. A., Miracolo, M. A., Donahue, N. M. & Robinson, A. L. Secondary organic aerosol formation from high-NO_x photo-oxidation of low volatility precursors: n-alkanes. *Environ. Sci. Technol.* **44**, 2029–2034 (2010).
- Griffith, S. M. et al. Measurements of hydroxyl and hydroperoxy radicals during CalNex-LA: model comparisons and radical budgets. *J. Geophys. Res. Atmos.* **121**, 4211–4232 (2016).
- Di Carlo, P. et al. Missing OH reactivity in a forest: evidence for unknown reactive biogenic VOCs. *Science* **304**, 722–725 (2004).
- Sheehy, P. M., Volkamer, R., Molina, L. T. & Molina, M. J. Oxidative capacity of the Mexico City atmosphere—part 2: a RO_x radical cycling perspective. *Atmos. Chem. Phys.* **10**, 6993–7008 (2010).
- Ren, X. R. et al. OH and HO₂ chemistry in the urban atmosphere of New York City. *Atmos. Environ.* **37**, 3639–3651 (2003).
- Shirley, T. R. et al. Atmospheric oxidation in the Mexico City Metropolitan Area (MCMA) during April 2003. *Atmos. Chem. Phys.* **6**, 2753–2765 (2006).
- Williams, A. J. et al. The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. *J. Cheminformatics* **9**, 61 (2017).
- Mansouri, K., Grulke, C. M., Judson, R. S. & Williams, A. J. Opera models for predicting physicochemical properties and environmental fate endpoints. *J. Cheminformatics* **10**, 10 (2018).
- Toxicity Estimation Software Tool (TEST) (US Environmental Protection Agency, 2016); <https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test>
- 2011 National Emissions Inventory version 2 technical support document (US Environmental Protection Agency, 2015); https://www.epa.gov/sites/production/files/2015-10/documents/nei2011v2_tsd_14aug2015.pdf
- US Industry Forecasts for 2012 & 2017: Solvents (Freedonia, 2008).
- Air Emissions Inventory Improvement Program (EIIP) technical report series (US Environmental Protection Agency, 1997); <https://www.epa.gov/air-emissions-inventories/air-emissions-inventory-improvement-program-eiip>.
- Solvent Utilization: Documentation for EPA's Nonpoint Emissions Estimation Tool (Abt Associates, 2017).
- Dionisio, K. L. et al. The Chemical and Products Database, a resource for exposure-relevant data on chemicals in consumer products. *Sci. Data* **5**, 180125 (2018).
- Simon, H. et al. The development and uses of EPA's SPECIATE database. *Atmos. Pollut. Res.* **1**, 196–206 (2010).
- Carter, W. P. L. Development of an improved chemical speciation database for processing emissions of volatile organic compounds for air quality models (2020); <https://intra.engr.ucr.edu/~carter/emidb/>
- Ryerson, T. et al. The 2010 California research at the nexus of air quality and climate change (CalNex) field study. *J. Geophys. Res. Atmos.* **118**, 5830–5866 (2013).

55. Pye, H. O. T. et al. Modeling the current and future roles of particulate organic nitrates in the southeastern United States. *Environ. Sci. Technol.* **49**, 14195–14203 (2015).
56. Xie, Y. et al. Understanding the impact of recent advances in isoprene photooxidation on simulations of regional air quality. *Atmos. Chem. Phys.* **13**, 8439–8455 (2013).
57. Murphy, B. N. et al. Semivolatile POA and parameterized total combustion SOA in CMAQv5.2: impacts on source strength and partitioning. *Atmos. Chem. Phys.* **17**, 11107–11133 (2017).
58. Pye, H. O. T. et al. On the implications of aerosol liquid water and phase separation for organic aerosol mass. *Atmos. Chem. Phys.* **17**, 343–369 (2017).
59. Xu, L. et al. Experimental and model estimates of the contributions from biogenic monoterpenes and sesquiterpenes to secondary organic aerosol in the southeastern United States. *Atmos. Chem. Phys.* **18**, 12613–12637 (2018).
60. Lu, Q. et al. Simulation of organic aerosol formation during the CalNex study: updated mobile emissions and secondary organic aerosol parameterization for intermediate-volatility organic compounds. *Atmos. Chem. Phys.* **20**, 4313–4332 (2020).
61. Bash, J. O., Baker, K. R. & Beaver, M. R. Evaluation of improved land use and canopy representation in BEIS v3.61 with biogenic VOC measurements in California. *Geosci. Model Dev.* **9**, 2191–2207 (2016).
62. Goldstein, A. H. & Galbally, I. E. Known and unexplored organic constituents in the Earth's atmosphere. *Environ. Sci. Technol.* **41**, 1514–1521 (2007).
63. Baker, K. R. et al. 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 (2013).
64. Kelly, J. T. et al. 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 (2014).
65. Markovic, M. et al. Measurements and modeling of the inorganic chemical composition of fine particulate matter and associated precursor gases in California's San Joaquin Valley during CalNex 2010. *J. Geophys. Res. Atmos.* **119**, 6853–6866 (2014).
66. Hayes, P. et al. Organic aerosol composition and sources in Pasadena, California, during the 2010 CalNex campaign. *J. Geophys. Res. Atmos.* **118**, 9233–9257 (2013).
67. Bahreini, R. et al. Organic aerosol formation in urban and industrial plumes near Houston and Dallas, Texas. *J. Geophys. Res. Atmos.* **114**, D00F16 (2009).

Acknowledgements

We thank EPA's SPECIATE workgroup, K. Seltzer, M. Strum, J. Snyder and T. Rao for useful discussions. This project was supported in part by an appointment to the Research Participation Program at the Office of Research and Development, US Environmental Protection Agency, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the US Department of Energy and the EPA. B.C.M. and S.A.M. were supported by NOAA award no. NA17OAR4320101 to the University of Colorado Boulder, Cooperative Institute for Research in Environmental Sciences. A.L.R. was supported by the Environmental Protection Agency STAR assistance agreement no. RD83587301. The US Environmental Protection Agency, through its Office of Research and Development, collaborated in the research described here. The research has been subjected to Agency administrative review and approved for publication but may not necessarily reflect official Agency policy. The views expressed in this Article are those of the authors and do not necessarily represent the views or policies of the US Environmental Protection Agency or of the National Oceanic and Atmospheric Administration.

Author contributions

H.O.T.P. and M.Q. conceived and designed the experiments. M.Q., B.N.M., K.K.I. and L.K. performed the experiments. M.Q., H.O.T.P., B.N.M., K.K.I., B.C.M. and S.A.M. analysed and interpreted the data. All of the authors contributed materials and/or analysis tools. M.Q. and H.O.T.P. wrote the manuscript with substantial contributions from all of the authors.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41893-020-00614-1>.

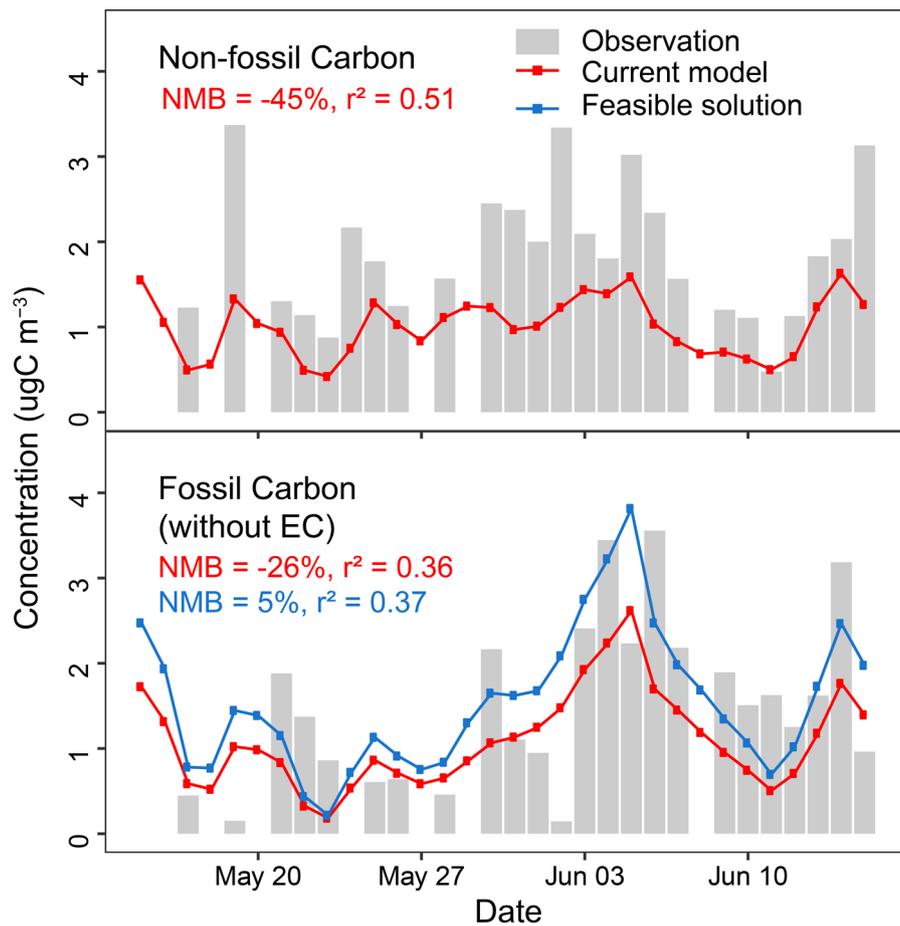
Supplementary information is available for this paper at <https://doi.org/10.1038/s41893-020-00614-1>.

Correspondence and requests for materials should be addressed to M.Q. or H.O.T.P.

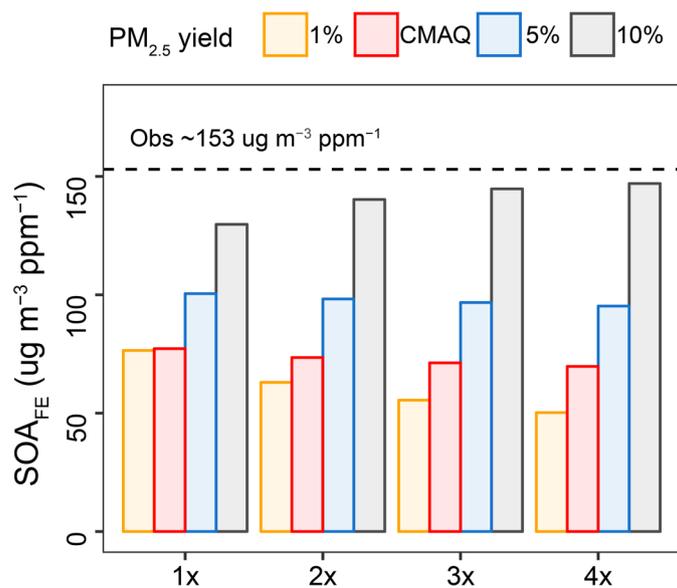
Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

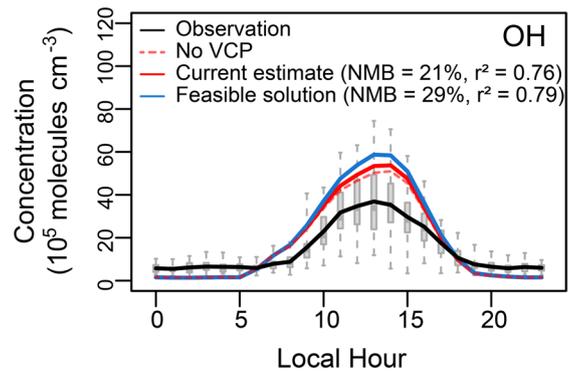
This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2020



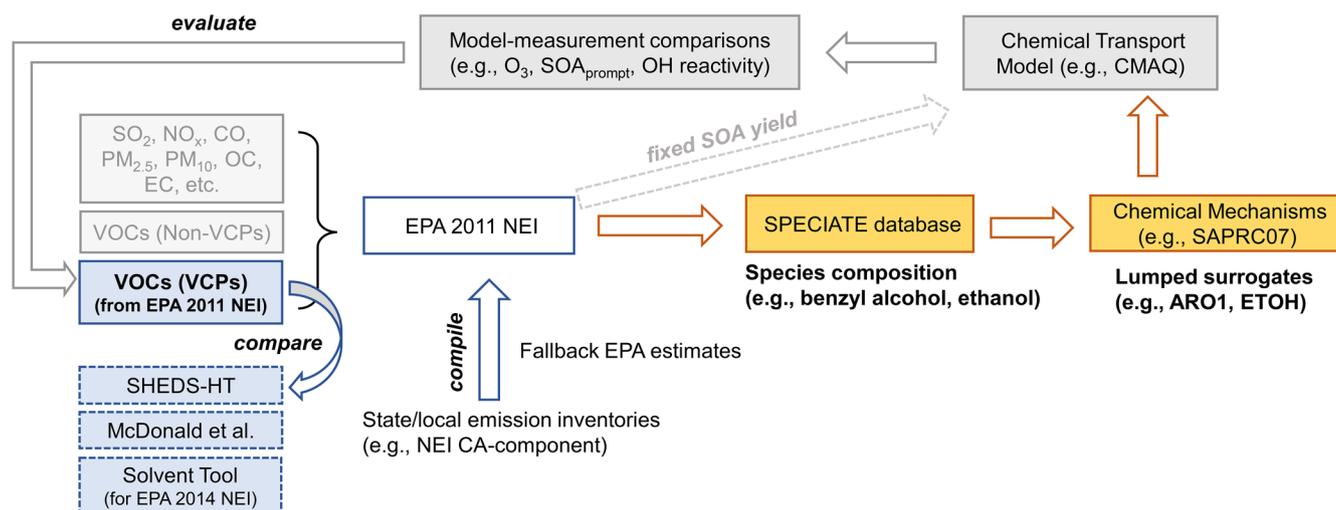
Extended Data Fig. 1 | Time series of simulated and observed fossil and non-fossil carbon at Pasadena. Simulated in the *current model* and the *feasible solution* case compared to observations in Woody et al.¹⁴.



Extended Data Fig. 2 | Simulated SOA formation efficiency (SOA_{FE}) over 8:30 am - 12:30 pm at Pasadena. SOA_{FE} quantifies SOA mass formation per volume of organic gases reacted over a time window, which brings together the organic $PM_{2.5}$ yield, precursor abundance, oxidant level, and reaction rate constant. See more details in Supplementary Notes.



Extended Data Fig. 3 | Simulated OH in the *current model* estimate and the *feasible solution* case compared to observations at Pasadena. The simulation without VCP emissions is indicated with the dashed line.



Extended Data Fig. 4 | Schematic of the methodology. This work integrated near-field (that is, SHEDS-HT) with far-field (that is, CMAQ modeling with 2011 NEI) modeling and top-down constraints based on ambient measurements. The blue boxes indicate emission estimates that were inter-compared. The emissions in EPA 2011 NEI, containing the estimate for VCP-emitted VOCs (solid outlined), went into air quality modeling. The yellow boxes indicate processing of the NEI including chemical speciation of emissions with the SPECIATE database, and mapping to CMAQ regional model surrogates using a certain chemical mechanism (for example, SAPRC07). For VCPs, SOA formation was parametrized with a fixed SOA yield (large dashed arrow), and thus emission processing was not required for VCP-emitted VOCs. See more details in Methods.