SAPRC07TIC_AE7I_VCP assignment rules:

Note: Mapping for SOA and radical chemistry are independently treated. Therefore, double mapping may occur. Rules are based on contents of SMILES, k_{OH}, calculated log(C*), number of oxygens (n_O), number of carbons (n_C), and estimated SOA yield.

SOA Chemistry -
1. If SMILES string contains "Si" (i.e. it's a siloxane or silane), add it fully (fraction = 1.0) to **SILOX**.
2. If a species has a SPECIATE_ID = 9001-9032 (i.e. it's new), follow these rules to assign it fully (fraction = 1.0) to one of the new surrogates. IVOP3/4/5/6/5ARO/6ARO definitions are based on (Lu et al., 2020). If a species is mapped to IVOC, NVOL, or NROG (US EPA, 2019), follow these rules to re-assign the existing fraction.
    1. If the species has an estimated SOA yield = 0.0%, assign it to **NONR**.
    2. If the species has $n_O > 0$ and an estimated SOA yield > 0.0%, assign it to **SOAOXY**.
    3. If the species has $n_O = 0$, an estimated SOA yield > 0.0%, and has log(C*) > 6.5, assign it to **SOAALK** (both aromatic and aliphatic species).
    4. If the species has $n_O = 0$, an estimated SOA yield > 0.0%, is aromatic, and has 5.5 < log(C*) < 6.5, assign it to **IVOCP6ARO**.
    5. If the species has $n_O = 0$, an estimated SOA yield > 0.0%, is aromatic, and has log(C*) < 5.5, assign it to **IVOCP5ARO**.
    6. If the species has $n_O = 0$, an estimated SOA yield > 0.0%, is aliphatic, and has 5.5 < log(C*) < 6.5, assign it to **IVOCP6**.
    7. If the species has $n_O = 0$, an estimated SOA yield > 0.0%, is aliphatic, and has 4.5 < log(C*) < 5.5, assign it to **IVOCP5**.
    8. If the species has $n_O = 0$, an estimated SOA yield > 0.0%, is aliphatic, and has 3.5 < log(C*) < 4.5, assign it to **IVOCP4**.
    9. If the species has $n_O = 0$, an estimated SOA yield > 0.0%, is aliphatic, and has log(C*) < 3.5, assign it to **IVOCP3**.
3. If an existing species is mapped to ALK4/5 and has an estimated SOA yield > 0.0%, assign it fully (fraction = 1.0) to **SOAALK**.
4. Four species are uniquely treated:
    1. Divinyl Benzene (SPECIATE_ID = 2081): assign fully (fraction = 1.0) to **ARO2MN**.
    2. Styrene (SPECIATE_ID = 698): assign fully (fraction = 1.0) to **ARO2MN**.
    3. Dimethyl Succinate (SPECIATE_ID = 420): assign fully (fraction = 1.0) to **SOAOXY**.
    4. Fragrances (SPECIATE_ID = 467): assign fully (fraction = 1.0) to **IVOCP6** and fully (fraction = 1.0) to **ALK5**.
5. Maintain all other original SAPRC07TC_AE7 assignments.
6. Replace all ARO2 assignments with ARO2MN to match CMAQ's EmissCtrl naming.

Radical Chemistry -
1. If SMILES string contains "Si" (i.e. it's a siloxane or silane), ignore. According to the SAPRC07 database (Carter, 2010), siloxanes have a negative MIR. No applicable surrogate available in SAPRC07.
2. If a species has a SPECIATE_ID = 9001-9032 (i.e. it's new), follow these rules to assign it fully (fraction = 1.0) to ALK1/2/3/4/5. If a species is mapped to IVOC, NVOL, or NROG, follow these rules to re-assign the existing fraction. k_{OH} rules are based on definitions in CMAQ (US EPA Office of Research and Development, 2020).
    a. If $1.35E\text{-}13 < k_{OH} < 3.38E\text{-}13$ cm^3 molec^{-1} sec^{-1}, assign to **ALK1**.
    b. If $3.38E\text{-}13 < k_{OH} < 1.69E\text{-}12$ cm^3 molec^{-1} sec^{-1}, assign to **ALK2**.
    c. If $1.69E\text{-}12 < k_{OH} < 3.38E\text{-}12$ cm^3 molec^{-1} sec^{-1}, assign to **ALK3**.
    d. If $3.38E\text{-}12 < k_{OH} < 6.77E\text{-}12$ cm^3 molec^{-1} sec^{-1}, assign to **ALK4**.
    e. If $6.77E\text{-}12 > k_{OH}$, assign to **ALK5**.
3. Maintain all other original SAPRC07TC_AE7 assignments.

Figure S1. SOA mass yield vs. log(C*) for 401 VCPy species categorized by their SAPRC07TC_AE7_VCP assignments. White shading indicates the range of VOCs with log(C*) > 6.5, orange shading indicates the range of IVOCs with 2.5 < log(C*) < 6.5, and green shading indicates the range of SVOCs with log(C*) < 2.5. SOA yield increases with decreasing volatility.

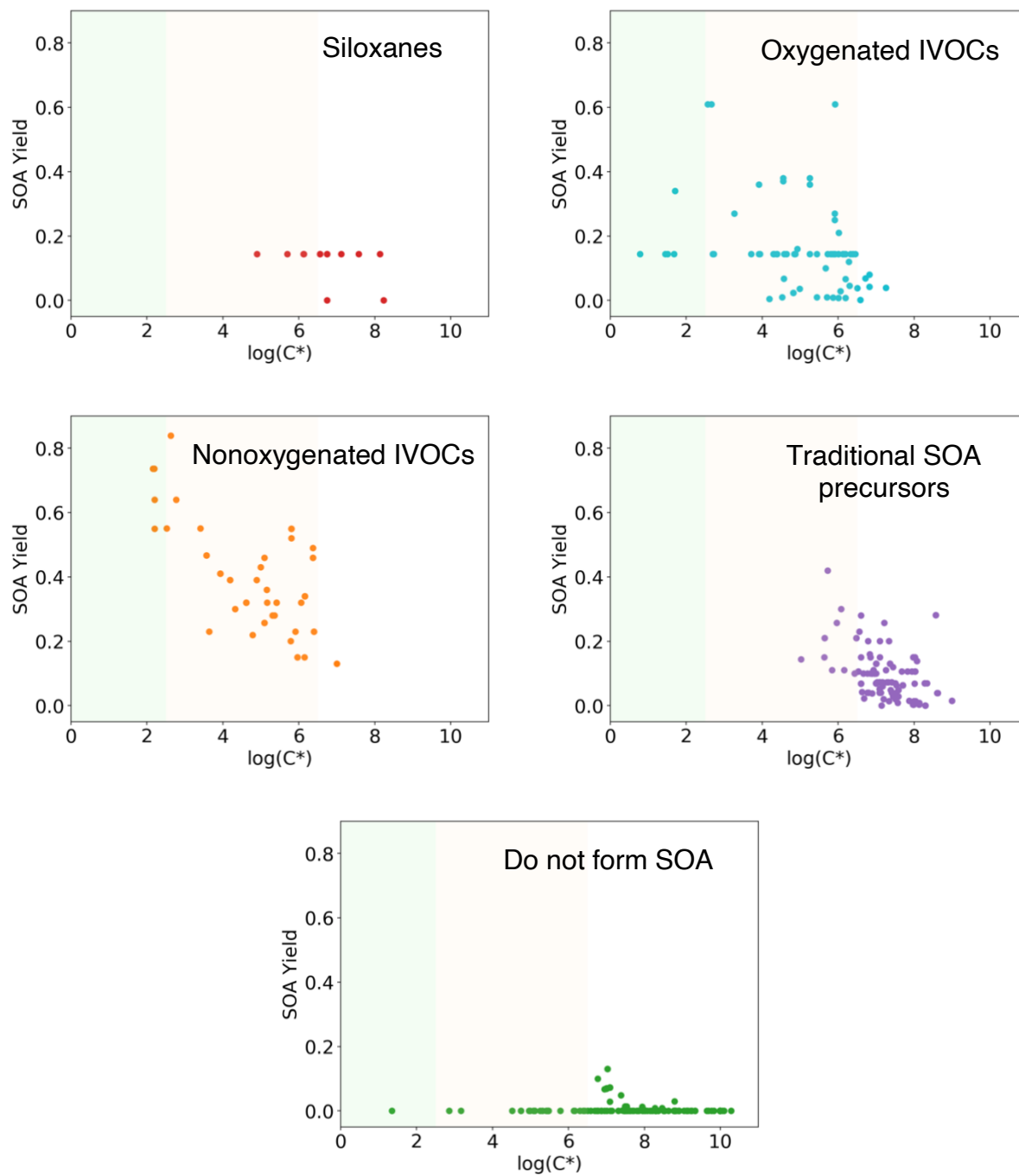



Figure S2. Average hourly concentrations of gas-phase IVOCs (oxygenated + nonoxygenated) predicted by the model for the zero VCP (green) and CMAQv5.3.2 (blue) cases. Horizontal lines depict campaign-average values for hydrocarbon-like IVOCs (6.3 µg m^{-3}) and oxygenated + hydrocarbon-like IVOCs (10.5 µg m^{-3}) from (Zhao et al., 2014).

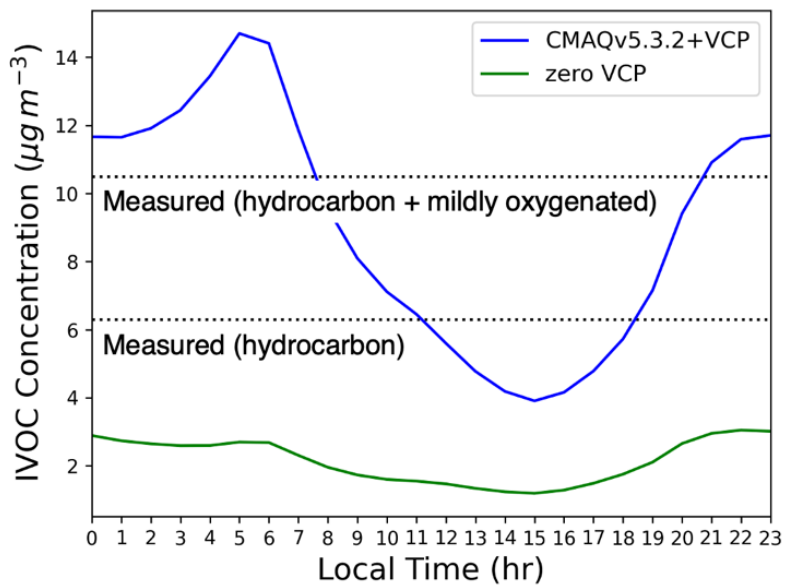



Figure S3. Average hourly concentrations observed and simulated by all three modeling cases May 15-June 15. Box and whiskers show all hourly concentrations observed at the Pasadena CalNex site. a) Background-corrected PM$_1$ SOA. A constant background value was removed from all observed concentrations according to the method in (Hayes et al., 2015). The background value of each simulation was determined by averaging the lower 50% of hourly concentrations from 00:00 LT to 04:00 LT and subtracting that from each curve. b) Formaldehyde (HCHO). Background values were not removed. c) Ozone (O$_3$). Background values were not removed. d) Carbon monoxide (CO). Background values were not removed. Box and whiskers were removed because they obscured the y-axis scale.

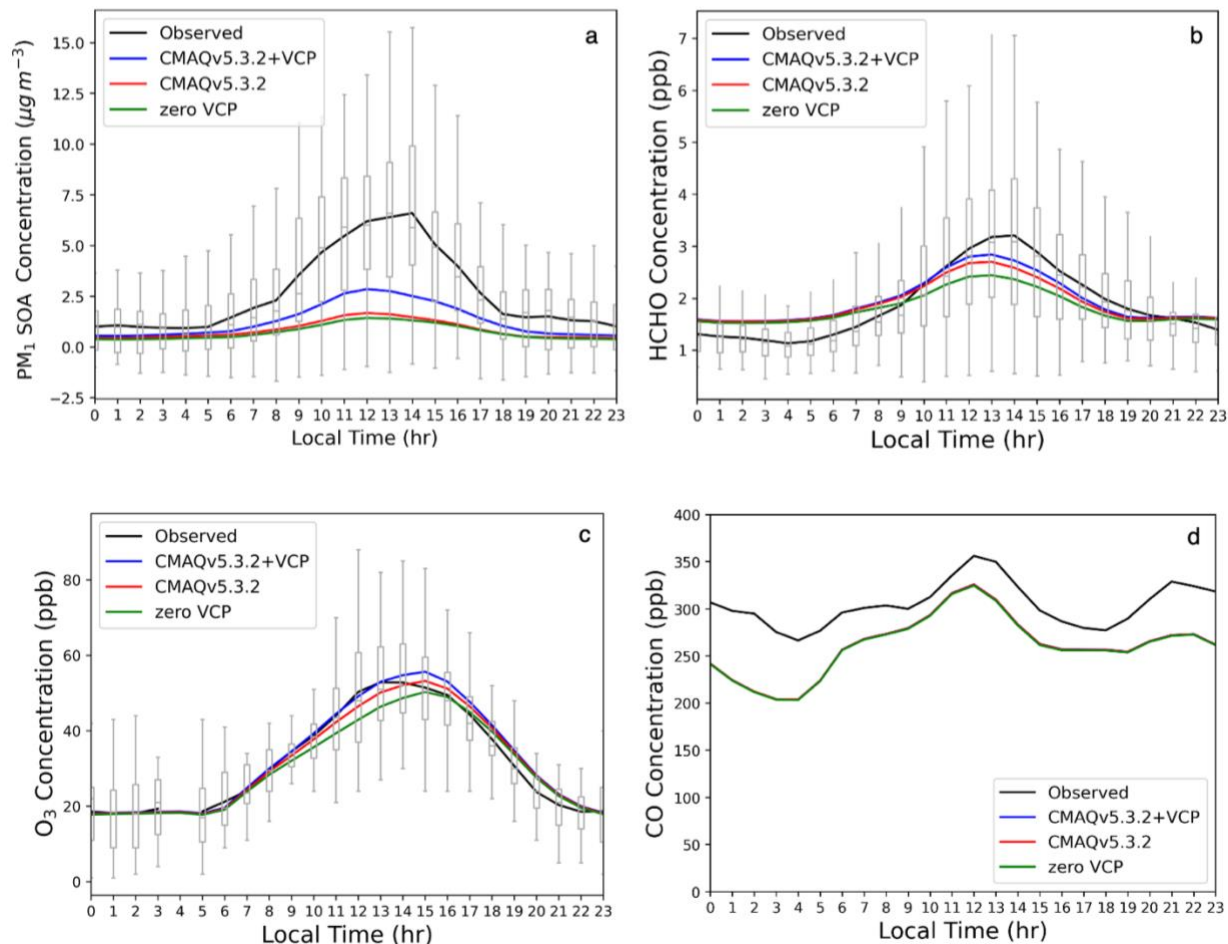



Figure S4. Modeled concentrations predicted by CMAQv5.3.2 case (red) and CMAQv5.3.2+VCP case (blue) vs. observations from the CalNex Pasadena ground site. a) Hourly PM$_1$ SOA. b) Hourly formaldehyde (HCHO). c) MDA8 O$_3$. Background values were not removed from any panels.

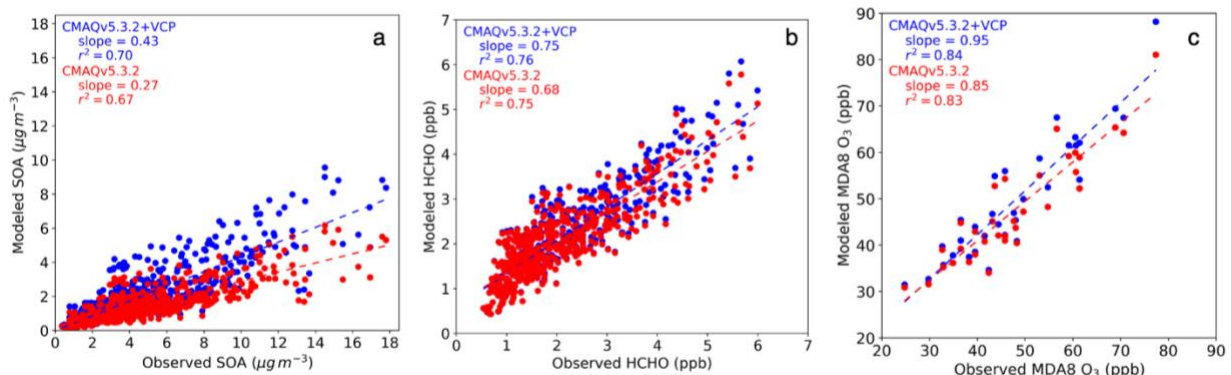



Figure S5. Modeled vs. observed MDA8 O$_3$ concentration for 178 routine monitoring sites from the AQS monitoring network in California for the zero VCP case (a), CMAQv5.3.2 case (b), and CMAQv5.3.2+VCP case (c).

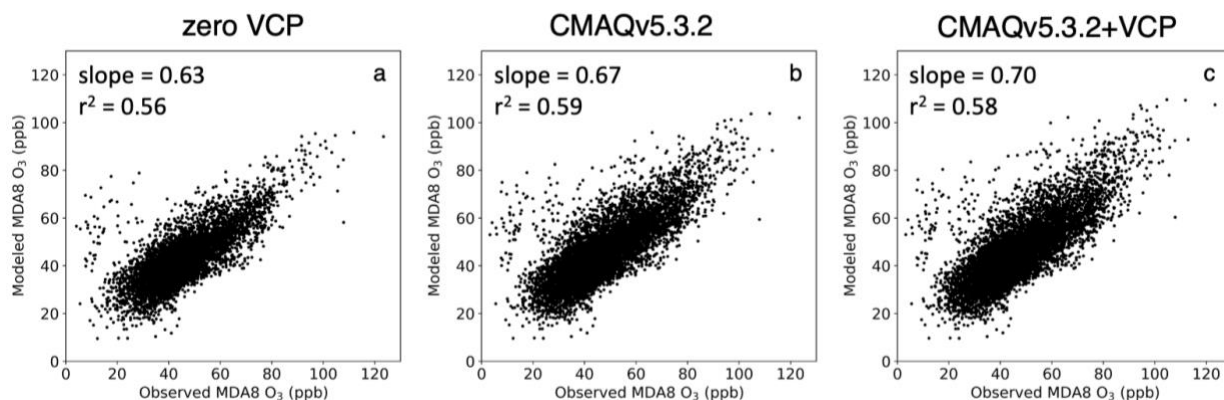



Figure S6. Bias (modeled - observed) of hourly concentrations vs. hourly modeled temperature for the CMAQv5.3.2 case (red) and CMAQv5.3.2+VCP case (blue). a) PM_1 SOA bias ($\mu g\ m^{-3}$). b) PM_1 POA bias ($\mu g\ m^{-3}$). c) Formaldehyde (HCHO) bias (ppb). d) CO bias (ppb).

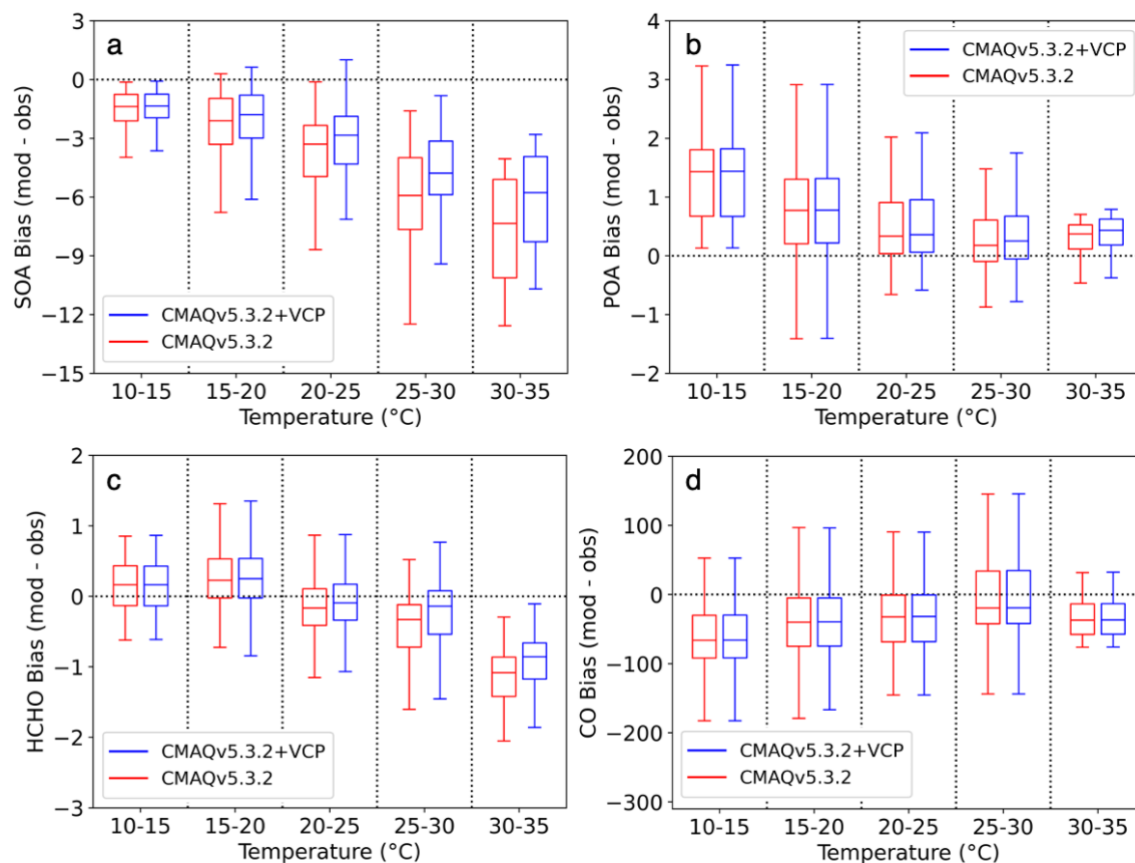



Figure S7. PM$_1$ SOA (a) and PM$_1$ POA (b) vs. temperature for zero VCP case (green), CMAQv5.3.2+VCP case (blue), and CalNex observations (black). Background values were not removed from any concentrations.

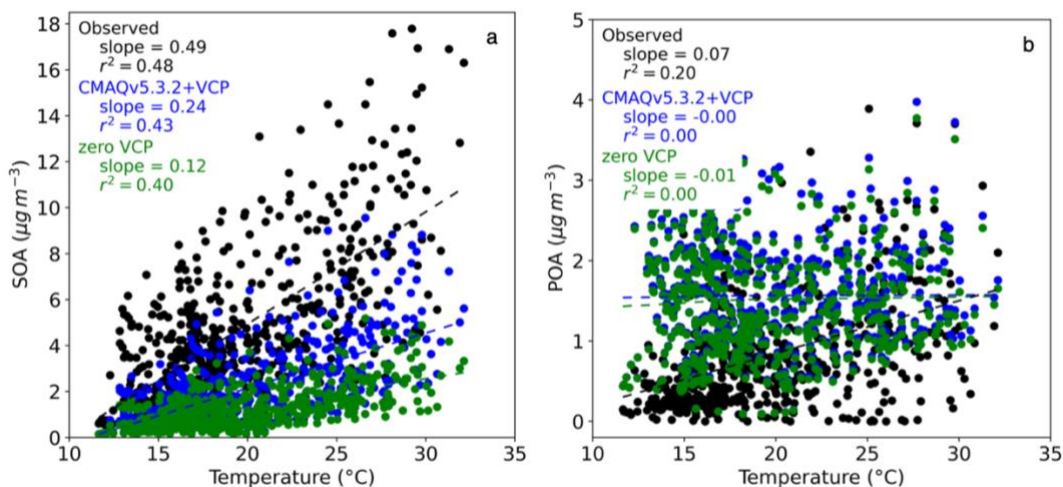
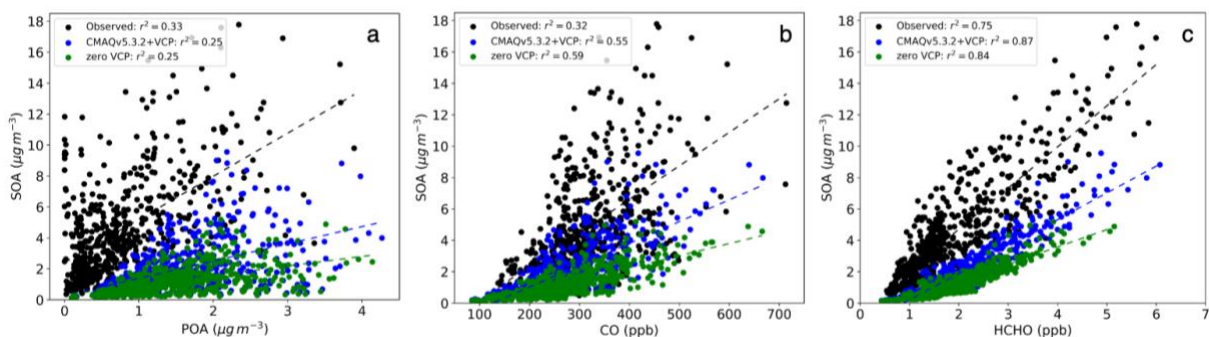



Figure S8. PM$_1$ SOA vs. PM$_1$ POA (a), PM$_1$ SOA vs. CO (b), PM$_1$ SOA vs. HCHO (c) for zero VCP case (green), CMAQv5.3.2+VCP case (blue), and CalNex observations (black). Background values were not removed from any concentrations.

References

Carter, W. P. L. (2010). Development of the SAPRC-07 chemical mechanism. *Atmospheric Environment*, *44*(40), 5324–5335. https://doi.org/10.1016/j.atmosenv.2010.01.026

Hayes, P. L., Carlton, A. G., Baker, K. R., Ahmadov, R., Washenfelder, R. A., Alvarez, S., Rappenglück, B., Gilman, J. B., Kuster, W. C., de Gouw, J. A., Zotter, P., Prévôt, A. S. H., Szidat, S., Kleindienst, T. E., Offenberg, J. H., Ma, P. K., & Jimenez, J. L. (2015). Modeling the formation and aging of secondary organic aerosols in Los Angeles during CalNex 2010. *Atmospheric Chemistry and Physics*, *15*(10), 5773–5801. https://doi.org/10.5194/acp-15-5773-2015

Lu, Q., Murphy, B. N., Qin, M., Adams, P. J., Zhao, Y., Pye, H. O. T., Efstathiou, C., Allen, C., & Robinson, A. L. (2020). Simulation of organic aerosol formation during the CalNex

study: Updated mobile emissions and secondary organic aerosol parameterization for intermediate-volatility organic compounds. *Atmospheric Chemistry and Physics*, *20*(7), 4313–4332. https://doi.org/10.5194/acp-20-4313-2020

US EPA, O. (2019, July 16). *SPECIATE 5.1 and 5.0 Addendum and Final Report* [Other Policies and Guidance]. US EPA. https://www.epa.gov/air-emissions-modeling/speciate-51-and-50-addendum-and-final-report

US EPA Office of Research and Development. (2020). *CMAQ*. Zenodo. https://doi.org/10.5281/zenodo.4081737

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. (2014). Intermediate-Volatility Organic Compounds: A Large Source of Secondary Organic Aerosol. *Environmental Science & Technology*, *48*(23), 13743–13750. https://doi.org/10.1021/es5035188