Supporting Information

Empirical Development of Ozone Design Value Isopleths: Applications to Los Angeles

[1]Yu Qian, [2]Lucas R.F. Henneman, [1]James A. Mulholland and [1*]Armistead G. Russell

[1]School of Civil and Environmental Engineering, Georgia Institute of Technology, GA 30332, USA

[2]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts 02115, United States

****Corresponding Author. Phone: 404-894-3079; Email: ted.russell@ce.gatech.edu***

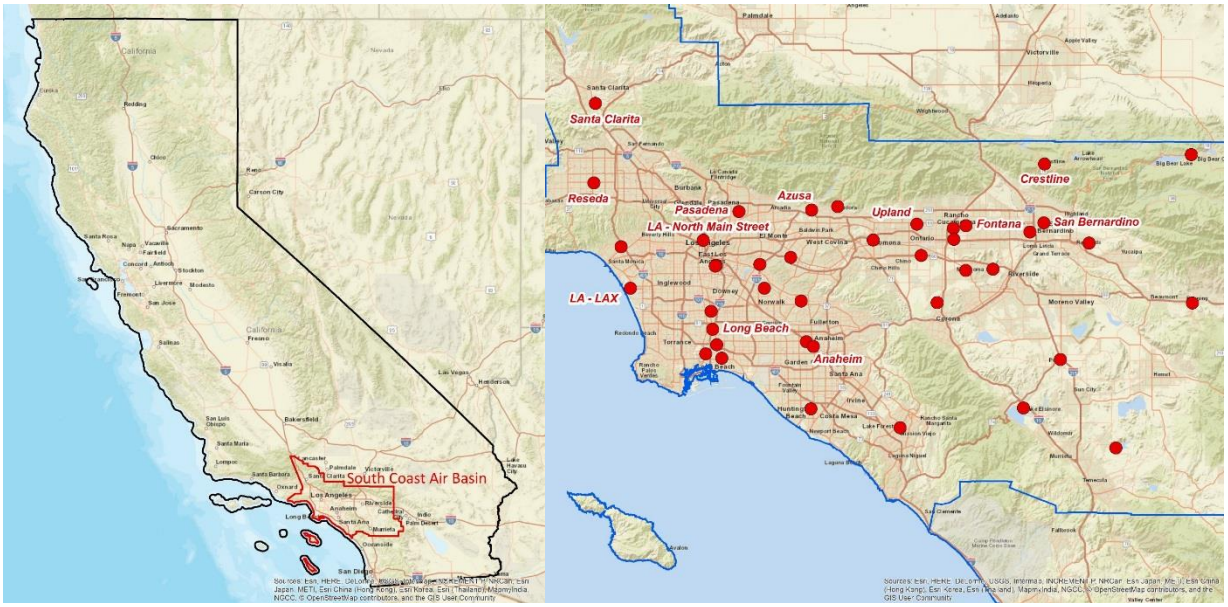

Fig. S1 Locations of California South Coast Air Basin (left) and 42 monitoring sites within SoCAB (right). SoCAB locates along the southwestern coast of California and is centered by Los Angeles. There are 42 monitoring sites in total in this air basin and most of them are along major highways.

Base Map Sources: Esri, DeLorme, HERE, USGS, Intermap, iPC, NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), MapmyIndia, Tomtom

Esri. "World Street Map" [basemap]. Scale Not Given. "World Street Map". March 13, 2019. https://www.arcgis.com/home/item.html?id=3b93337983e9436f8db950e38a8629af (accessed Apr 1, 2019).

Emission Data Source Description:

In this study, the estimated annually averaged emission data for NOx and VOC were obtained from California Almanac of Emissions and Air Quality(Cox et al., 2013, 2009). Two editions of the Almanac (2009 and 2013) were used as reference for comparison, though only **data from 2009 edition** were used because it covers a complete set of estimated emissions throughout the study period, which is more consistent for the regression analysis. The analysis uses, only, anthropogenic emissions estimates and does not include biogenic sources, nor the effect of climate change on biogenic emissions.

Between those two Almanac editions, organic ozone precursors are treated using two different terms. Historically, CARB used reactive organic gases (ROG) as ozone precursors and changed to using VOC in the 2013 edition. Based on the CARB's definition(Schwehr and Propper, 2009), ROG is not identical to U.S. EPA's term "VOC", but the two are similar with few difference in the list of exempted compounds. We conducted a comparison between estimated VOC and ROG emissions, as well as between Almanac and SCAQMD inventory, which shows a high consistency with very limited difference (average ROG to VOC ratio is 1.15, R^2 = 0.98). (Fig. S2).

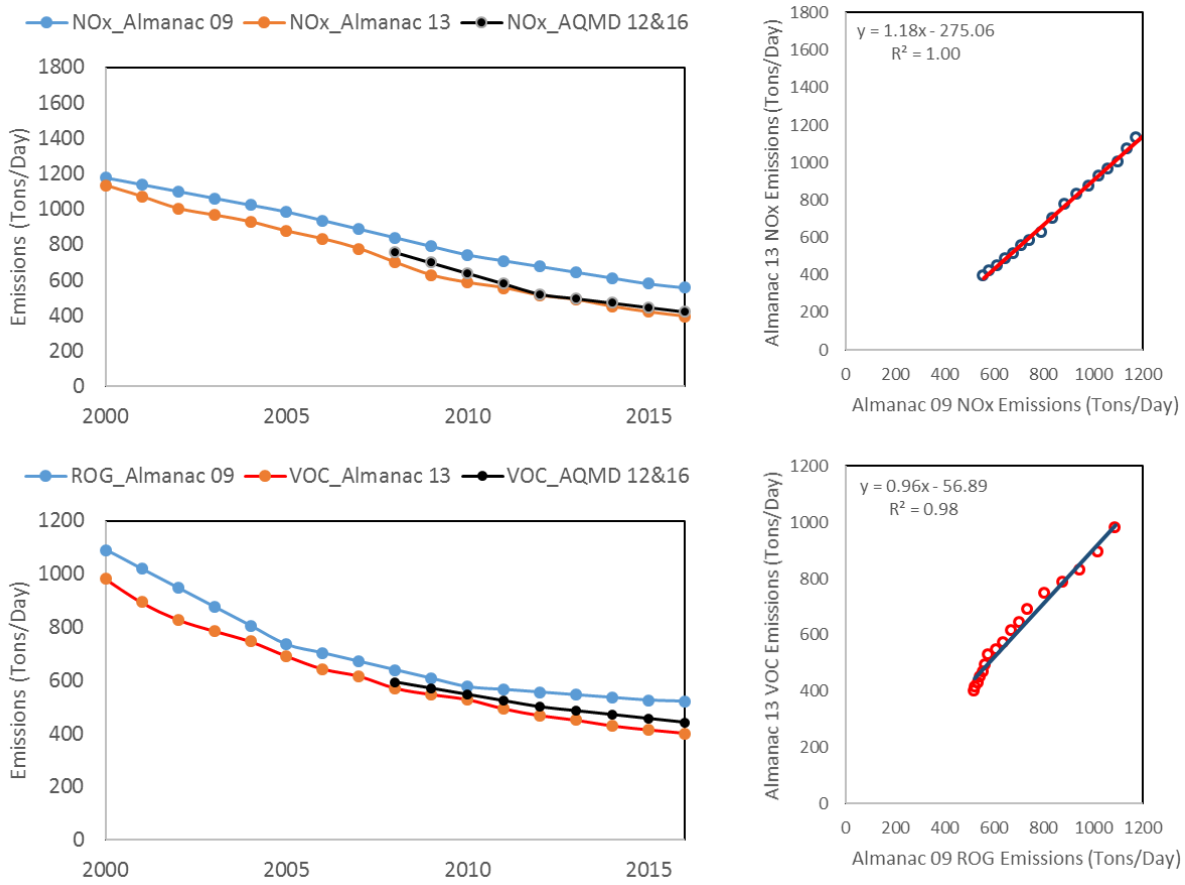

Fig. S2 Comparison between estimated ROG emissions and VOC emissions from Almanac 2009 edition and 2013 edition.

10-Fold Cross-Validation Analysis In addition to only examine the R^2 as an evaluation of the model performance, a cross-validation was also conducted to assess how the model is affected by the observation data set, how well when it is applied to an independent dataset. On the other word, it is to evaluate the predictive ability of the regression model we developed. This is conducted by holding 10% of the observation dataset as test set and only use the rest 90% of the observations as training set to build the model. Then we apply the model result to test set and evaluate the model performance. This process is repeated for 10 times to cover all of the observations.

Both models show good performance in this cross-validation test (Fig. S3). For most runs, the R^2s between observations from test set and estimations by the model are close to 1, with the slope also close to 1. There is no significant difference observed in the model performance between different runs. This result basically shows the regression modeling method we developed have robust productivity.

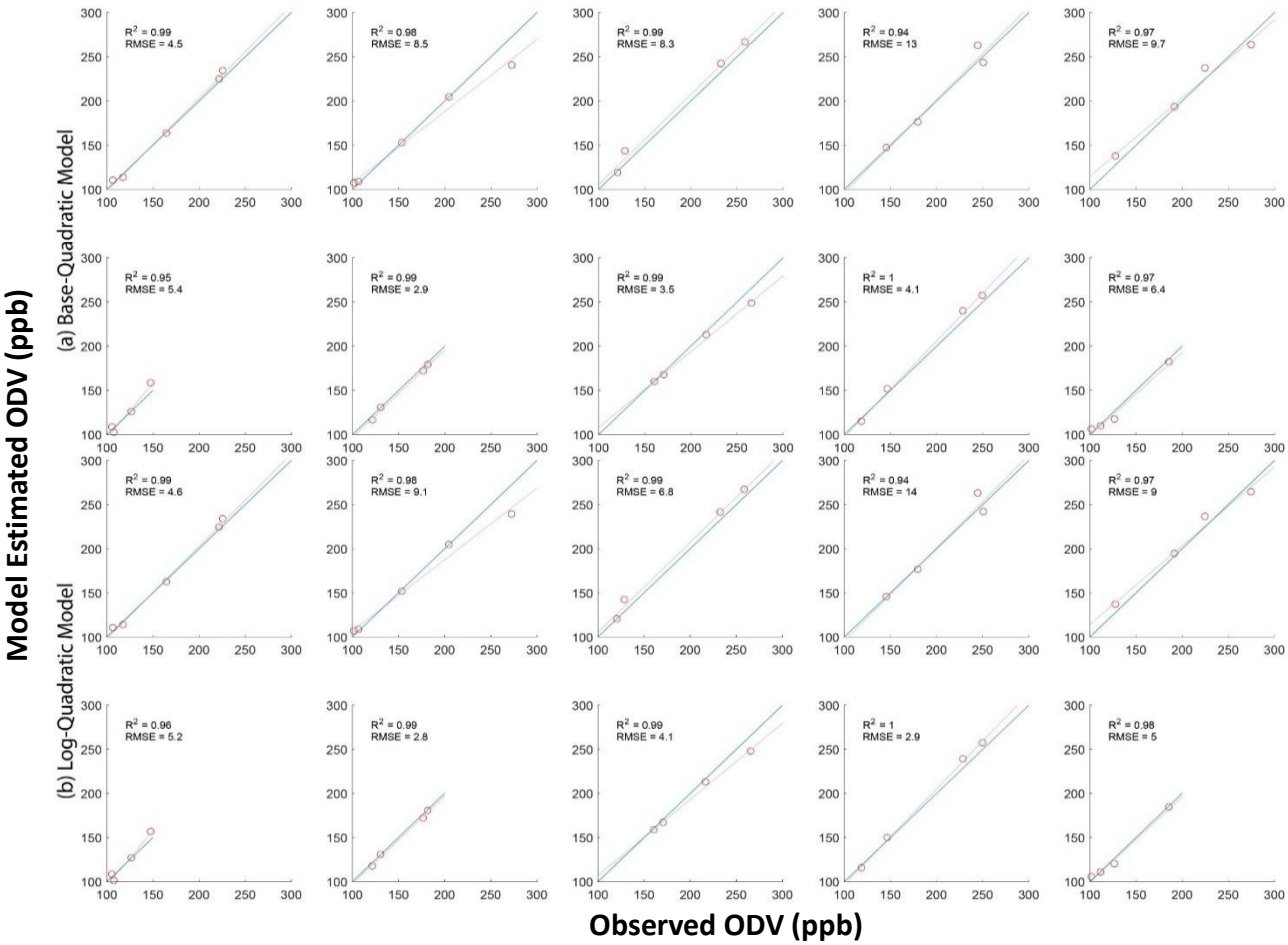

Fig. S3 Scatter plots of 10% cross-validation result for base-quadratic model (a) and log-quadratic model (b) respectively. 10% of the data was holdout, the model was trained by the leftover 90%, and the plots show the comparison between observation and prediction of the held out data (10 sets for each model). For 20 test sets from both models, everyone shows high correlation with slope close to one. The average R^2 is 0.98 and RMSE is 6. This evaluation strengthened the good performance and strong prediction ability of those models.

Derivation of Empirical Model Uncertainty:

Formal analysis was conducted based on the regression models(Helwig, 2017). The general formation of the regression equation can be expressed as:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{e} \quad \text{(EQN. S1)}$$

where,

$$\mathbf{y} = (\mathrm{y}_1, \dots, \mathrm{y}_\mathrm{n})' \in \mathbb{R}^\mathrm{n} \text{ is the n } \times 1 \text{ response vector (ODVs)}$$

$$\mathbf{X} = \left[1_\mathrm{n}, \mathbf{x}_1, \dots, \mathbf{x}_\mathrm{p}\right] \in \mathbb{R}^{\mathrm{n}\times(\mathrm{p}+1)} \text{ is the n } \times (\mathrm{p}+1) \text{ design matrix (Emission Variables)}$$

$$\mathbf{b} = \left(\mathrm{b}_0, b_1 \dots, \mathrm{b}_\mathrm{p}\right)' \in \mathbb{R}^{\mathrm{p}+1} \text{ is the } (\mathrm{p}+1) \times 1 \text{ vector of coefficients}$$

$$\mathbf{e} = (\mathrm{e}_1, \dots, \mathrm{e}_\mathrm{n})' \in \mathbb{R}^\mathrm{n} \text{ is the n } \times 1 \text{ error vector}$$

n is the number of data points

p is the number of independent variables

The ordinary least squares (OLS) solution has the form:

$$\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad \text{(EQN. S2)}$$

where $\hat{\mathbf{b}}$ is the estimated vector of coefficients based on observation $\mathbf{X}$ and $\mathbf{y}$.

The fitted values can be calculated as:

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{b}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad \text{(EQN. S3)}$$

where $\hat{\mathbf{y}}$ is the vector of estimated response variable (ODV).

The estimated error variance is $\hat{\sigma}^2$:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{(n-p-1)} \quad \text{(EQN. S4)}$$

which is an unbiased estimate of error variance σ^2.

When we use this model to predict a new observation $\mathbf{x_h} = \left[1, \mathrm{x_{h1}}, \dots, \mathrm{x_{hp}}\right]$, the fitted value is $\hat{y}_h = \boldsymbol{x}_h\hat{\boldsymbol{b}}$, where $\mathbf{x_h}$ is the vector of emission variables at a new observation point, $\hat{y}_h$ is the estimated value of the response variable (i.e. ODV) at the new observation point.

Using this equation, there are two types of uncertainties involved: first is related to the location of the distribution of $\hat{y}_h$ for $\mathbf{x_h}$ (captured by $\sigma^2_{\hat{y}_h}$); and second is the variability within the distribution of y (captured by σ^2).

$$\sigma^2 = \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{(n-p-1)} \quad \text{(EQN. S5)}$$

Variance of $\hat{y}_h$ is given by: $\sigma^2_{\hat{y}_h} = \mathrm{Var}(\mathbf{x}_h\hat{\mathbf{b}}) = \mathbf{x}_h\mathrm{Var}(\hat{\mathbf{b}})\mathbf{x}_\mathbf{h}' = \sigma^2\mathbf{x}_h(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{x}_\mathbf{h}'$ (EQN. S6)

The overall variance of the fitted value becomes: $\sigma^2_{y_h} = \sigma^2_{\hat{y}_h} + \sigma^2$ (EQN. S7)

Under the test $H_0 : E(y_h) = y_h^* \; \boldsymbol{vs.} \; H_1 : E(y_h) \neq y_h^*$,

100(1-α) % Prediction Interval for E(y_h): $\hat{y}_h \pm t_{n-p-1}^{(\frac{\alpha}{2})}\sigma_{y_h}$ (EQN. S8)

Where α is the confidence interval and t is the t-distribution value with probability is α/2 and degrees of freedom n-p-1.

The uncertainty of the prediction at any observation point $\mathbf{x}_h$ is then:

Base Quadratic Model: $t_{n-p-1}^{(\frac{\alpha}{2})}\sigma\sqrt{(\mathbf{x}_h(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{x}_\mathbf{h}' + \mathbf{1})}$ (EQN.S9)

Log Quadratic Model: $10^{\hat{y}_h + t_{n-p-1}^{(\frac{\alpha}{2})}\sigma\sqrt{(\mathbf{x}_h(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{x}_\mathbf{h}'+\mathbf{1})}} - 10^{\hat{y}_h - t_{n-p-1}^{(\frac{\alpha}{2})}\sigma\sqrt{(\mathbf{x}_h(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{x}_\mathbf{h}'+\mathbf{1})}}$ (EQN. S10)

Assessment of Reduced Coefficient Models:

In addition to the full model with 5 variables, we also developed the reduced form of those models based on Akaike information criterion (AIC) model selection method. This analysis is to investigate the necessity of those pre-designed variables and to evaluate how the model characteristics could change by trimming the model based on the variables statistical significance.

Reduced Base-Quadratic Regression Function:

$ODV (E_{NOx}, E_{VOC}) = 8.25 + 0.11*E_{NOx} + 0.11*E_{VOC} - 9*10^{-5} *(E_{NOx}^2)$ (EQN. S11)

Reduced Log-quadratic Regression Function:

$Log_{10} (ODV (E_{NOx}, E_{VOC})) = 1.75 + 0.0003*E_{NOx} + 0.0003*E_{VOC} - 2*10^{-7} *(E_{NOx}^2) - 2*10^{-8} *(E_{VOC}^2)$ (EQN. S12)

In both cases, the NOx-VOC interaction terms were removed. In general, the uncertainty of both models have been improved considerably. However, chemical modeling finds that the interaction terms can become important as emissions change beyond the range of observations, supporting the use of the full models. This can be also be supported by the resulting isopleth developed by reduced models (Fig. S4), which does not capture the typical isopleth form.

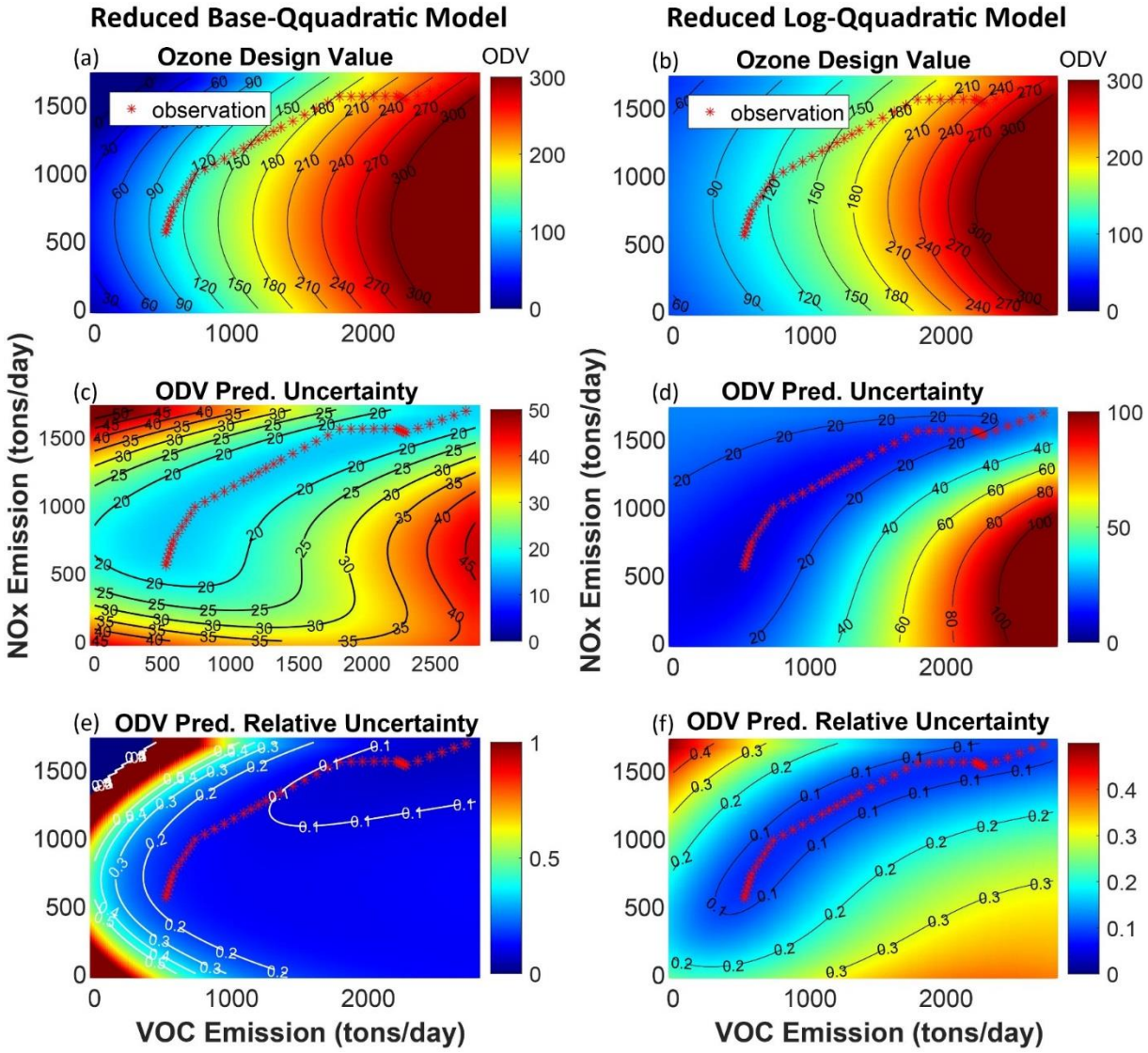

Fig. S4 ODV-Emissions Isopleths developed by empirically-derived non linear regression model with AIC model selection conducted and the uncertainty related to the models. Axes define the emissions space with varying levels of estimated NOx and VOC and corresponding modeled ODV indicated by color. Historical observations are noted with red asterisk. Left column shows the result of reduced base-quadratic model (a, c, e). Right column shows the result of reduced log-quadratic model (b, d, f). Top Row: emission space and Isopleths constructed based on different regression models(a, b). Middle Row: Prediction uncertainty of regression models at different emissions levels (c, d). Bottom Row: Relative Prediction uncertainty of regression models at different emissions levels (e, f).

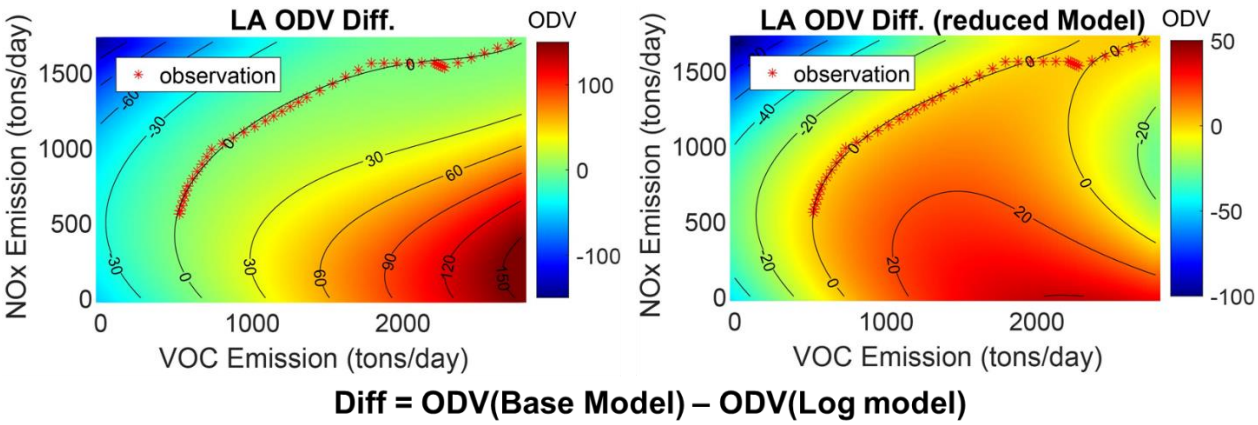

Fig. S5 The Comparison between base-quadratic models and log-quadratic models under different emission levels for both full models (left) and reduced models (right). The difference is an evaluation of model uncertainty. Lower uncertainty is observed near the observations and the differences increase when emissions are away from observations.

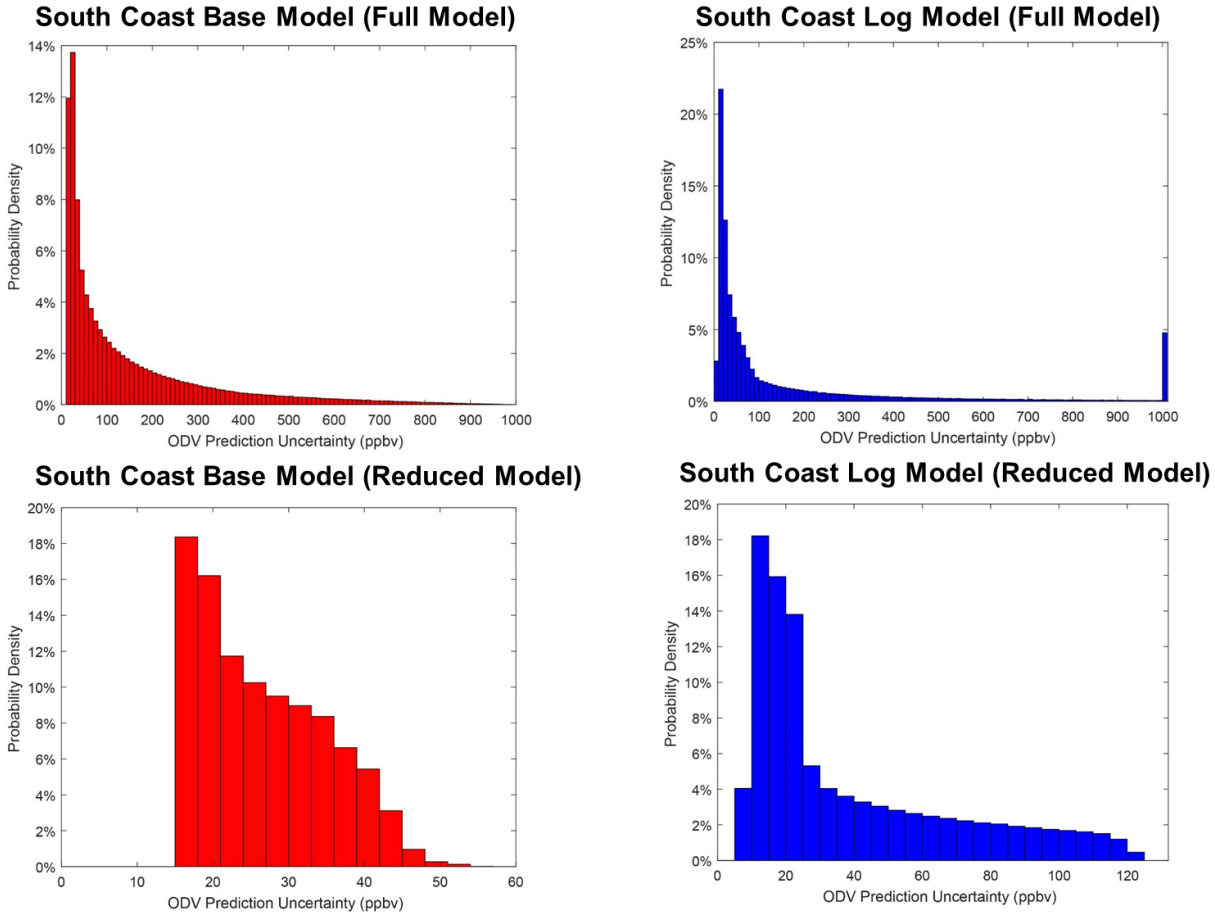

Fig. S6 The distributions of uncertainties for different regression models: Base-quadratic full model (top-left), log-quadratic full model (top-right), base-quadratic reduced model (bottom-left), and log-quadratic reduced model (bottom-right). Generally log models have relatively lower uncertainty than base models, and the model selection can significantly decrease prediction uncertainty. However, even though reduced models show lower uncertainty, the analysis based on the constructed isopleth indicating that the reduced form may not reflect the realistic relationship between emissions and ODV.

Assessment of how emissions uncertainties impact model results.

It is known that there are potentially major uncertainties in the emissions(Hanna et al., 2005; McDonald et al., 2018), and the question arises as to how much that can influence the approach taken here. Quantitative assessment of the uncertainties is, however, fraught with lack of knowledge of the magnitude of the uncertainties or the uncertainty structures. Here, the approach taken is to examine 1000 emission reduction trajectories, starting with the emissions at the beginning of the period (1975) and adding uncertainty to the reduction in emissions estimated each year. The approach is recursive such that the uncertainties compound over time. Descriptively, the modeled reduction between each year is the estimated reduction between those two years, plus a random uncertainty based on the emissions level (not the amount reduced).

Derivation of Empirical Model Uncertainty Based upon Independent Variables (Emissions): Simulation method was developed to evaluate the uncertainty of regression model caused by uncertainties in emissions trends.

1. Generate Simulated Emission Inventory Trajectories

Calculate the simulated emission ratio between each two years,

$$RNOx_{i+1,j} = \left(\frac{ENOx_{i+1}^{est}}{ENOx_{i}^{est}}\right) * (1 + 0.05 * Rand_{i,j}^{NOx}) \quad \text{(EQN. S13)}$$

$$RVOC_{i+1,j} = \left(\frac{EVOC_{i+1}^{est}}{EVOC_{i}^{est}}\right) * (1 + 0.05 * Rand_{i,j}^{VOC})) \quad \text{(EQN. S14)}$$

Where:

i is the index for year, from 1975 to 2015;

j is the index for run of simulations, which is from 1 to N, here we take N = 1000;

$ENOx_i^{est}$ is the estimated NOx emission for year i based on CARB's emission inventory;

$EVOC_i^{est}$ is the estimated VOC emission for year i based on CARB's emission inventory;

$RNOx_{i+1,j}$ is the simulated NOx emission ratio between year i+1 and year i, for run j;

$RVOC_{i+1,j}$ is the simulated VOC emission ratio between year i+1 and year i, for run j;

$Rand_{i,j}^{NOx}$ and $Rand_{i,j}^{VOC}$ are both randomly generated, normally distributed (μ = 0, σ = 1) numbers for year i and run j.

Based on this algorithm, for each run j, the emissions ratio between every two adjacent years for both NOx and VOC are simulated based on the actual estimated emission ratio with a random uncertainty of 5% level per year.

$$ENOx_{i+1,j} = ENOx_{i,j} * RNOx_{i+1,j} \quad \text{(EQN. S15)}$$

$$EVOC_{i+1,j} = EVOC_{i,j} * RVOC_{i+1,j} \quad \text{(EQN. S16)}$$

$ENOx_{i,j}$ is the simulated NOx emission for year i, run j; $ENOx_{1,j} = ENOx_1^{est}$;

$EVOC_{i,j}$ is the simulated VOC emission for year i, run j; $ENOx_{1,j} = ENOx_1^{est}$;

Table S1 Simulation result of independent variable (emissions) uncertainty analysis. Statistics of simulated emissions trajectories: including original estimated emissions based on inventory, mean and standard deviation of NOx and VOC emissions for each year over 1000 simulations. Here also listed ODV calculated by original models, comparing to mean and standard deviation of 1000 ODVs calculated by 1000 simulated models for each year.

	NOx Emission (tons/day)			VOC Emission (tons/day)			ODV (ppb)				
Year	Estimated	Simulated		Estimated	Simulated		Observed	Simulated (Based Model)		Simulated (Log Model)	
		Mean	Std.		Mean	Std.		Mean	Std.	Mean	Std.
1975	1691			2718			275	269	11.4	269	13.0
1976	1659	1658	82	2630	2635	126	259	261	10.2	262	11.3
1977	1626	1626	111	2542	2545	166	245	256	10.0	256	10.9
1978	1594	1596	135	2455	2458	195	250	251	10.4	251	11.3
1979	1562	1561	152	2367	2369	221	266	247	11.7	247	12.7
1980	1530	1531	169	2279	2275	245	273	241	12.4	240	13.2
1981	1536	1542	188	2266	2260	271	251	238	11.5	238	12.2
1982	1542	1544	204	2252	2245	291	233	235	10.8	234	11.3
1983	1549	1548	215	2238	2232	308	229	232	10.0	232	10.5
1984	1555	1557	232	2225	2216	323	225	230	9.6	230	10.0
1985	1561	1562	245	2211	2203	344	226	229	10.1	228	10.5
1986	1560	1559	258	2128	2120	344	222	222	10.2	222	10.5
1987	1560	1553	270	2044	2033	346	217	216	9.9	215	9.9
1988	1559	1557	279	1961	1948	345	205	209	9.7	208	9.6
1989	1558	1555	288	1877	1865	344	192	202	9.7	201	9.4
1990	1558	1557	301	1793	1786	342	186	196	9.9	195	9.4
1991	1513	1510	302	1708	1697	340	182	189	9.2	188	8.6
1992	1467	1465	297	1622	1607	332	180	182	8.8	182	8.1
1993	1422	1417	296	1536	1521	324	177	176	8.0	176	7.3
1994	1377	1371	292	1450	1437	314	171	170	7.3	170	6.5
1995	1332	1326	287	1365	1354	303	165	164	6.7	164	5.8
1996	1301	1297	288	1310	1297	299	161	160	6.7	159	5.8
1997	1270	1262	280	1255	1242	294	148	156	6.6	155	5.6
1998	1239	1232	281	1200	1187	285	154	152	6.3	152	5.2
1999	1208	1203	280	1145	1133	279	147	149	5.7	148	4.7
2000	1177	1173	279	1090	1077	269	146	145	5.6	144	4.5
2001	1138	1135	274	1019	1007	254	129	140	5.2	139	4.1
2002	1100	1096	273	948	936	238	128	135	4.8	135	3.7
2003	1062	1055	270	877	865	224	131	131	4.4	131	3.4
2004	1023	1016	265	806	798	211	127	127	3.8	127	2.8
2005	985	976	259	735	728	197	127	122	3.8	123	2.7
2006	937	926	251	703	698	192	121	120	3.5	121	2.6
2007	888	877	244	671	668	187	122	118	3.3	119	2.4
2008	839	827	234	640	635	181	119	116	3.4	116	2.4
2009	791	780	225	608	605	175	118	113	3.1	114	2.2
2010	742	732	213	576	573	168	112	111	3.0	111	2.0
2011	710	701	206	566	563	170	107	110	2.8	110	1.9
2012	677	669	199	556	551	169	106	108	2.7	108	1.9
2013	645	637	193	546	541	168	107	107	2.6	107	1.8
2014	612	607	188	536	532	168	102	106	2.4	106	1.6
2015	580	575	181	526	522	167	102	105	2.9	105	1.8
2016	557	553	176	522	517	168	108	104	3.7	104	2.3

For each run, simulated emissions trend for both NOx and VOC are calculated based on the simulated emission ratio, and the first years emissions are set to be equal to the estimated first year emissions. After all runs, we will have N time series of both NOx and VOC emissions for 1975 to 2016.

The choice of 1975 base emissions as the starting point does not mean that we believe that the 1975 estimate is the most accurate, but the interest is in the emissions trajectories. A similar result is found if taking any other year as the starting year (e. g, one can start with 2016 and go backwards using the same recursive approach with similar results).

2. Fitting the Regression Model with the Simulated Emission Trends

Once the 1000 simulated emission inventory estimates were generated for each year, we applied the same regression method to get a regression model for each run j for either base-quadratic model or log-quadratic model. The resulting mean coefficients are shown in Table S1. In addition, the ozone isopleth resulting from the uncertainty analysis can also be plotted by taking the mean of ODVs calculated by the 1000 different regression models at the different emission levels (Fig. S7). Similarly, the a heat map of the ozone standard deviation is also be plotted by taking the standard deviation of ODVs calculated by the 1000 different regression models for the range of emissions (Fig. S7).

3. Result

This approach assumes that the trajectory of emission reductions is more certain than the absolute emissions. With the accumulative uncertainty in the emission ratios at 5% level, the variation (uncertainty) simulated at different emissions levels is significant (Table S1, Fig. S7). In general, with considerable uncertainty simulated in the emissions (independent variables), the average result of 1000 emissions uncertainty simulations is very similar to the original models (Fig. S7, Table S2). The signs of averaged regression coefficients are consistent with those calculated based on the original model and the values of those coefficients are in the similar level (Table S2). On average, after involving considerable uncertainty in emissions, the simulated models did not significantly decrease the model performance (R^2 only decreases 0.02). We found a small uncertainty near the observations, growing as one moves away from the historical observations, which is consistent with the formal uncertainty analysis for independent variables. Only limited uncertainty is observed over the potential future emissions trajectory. Also, the difference between simulated base-quadratic model and log-quadratic model is very limited, which shows the stability of this approach under uncertainties involved in emissions.

Table S2 Simulation results of independent variable (emissions) uncertainty analysi (statistics of regression coefficients: mean and standard deviation) and comparison with original model results.

	Coefficients	Intercept	α_NOx	α_VOC	α_NOx*VOC	α_NOx^2	α_VOC^2	R^2
Base	Original	1.03	0.1347	0.1101	3.52E-05	-1.19E-04	-1.10E-05	0.98
	Simulation - Mean	62.78	0.0414	0.0440	4.42E-05	-3.78E-05	-1.07E-05	0.96
	Simulation - Std	54.89	0.1525	0.0995	1.87E-04	1.74E-04	5.41E-05	0.030
Log	Original	1.72	0.0004	0.0002	1.80E-07	-3.52E-07	-6.90E-08	0.99
	Simulation - Mean	1.84	0.0002	0.0002	8.19E-08	-1.09E-07	-3.81E-08	0.97
	Simulation - Std	0.13	0.0003	0.0002	3.71E-07	3.54E-07	1.07E-07	0.019

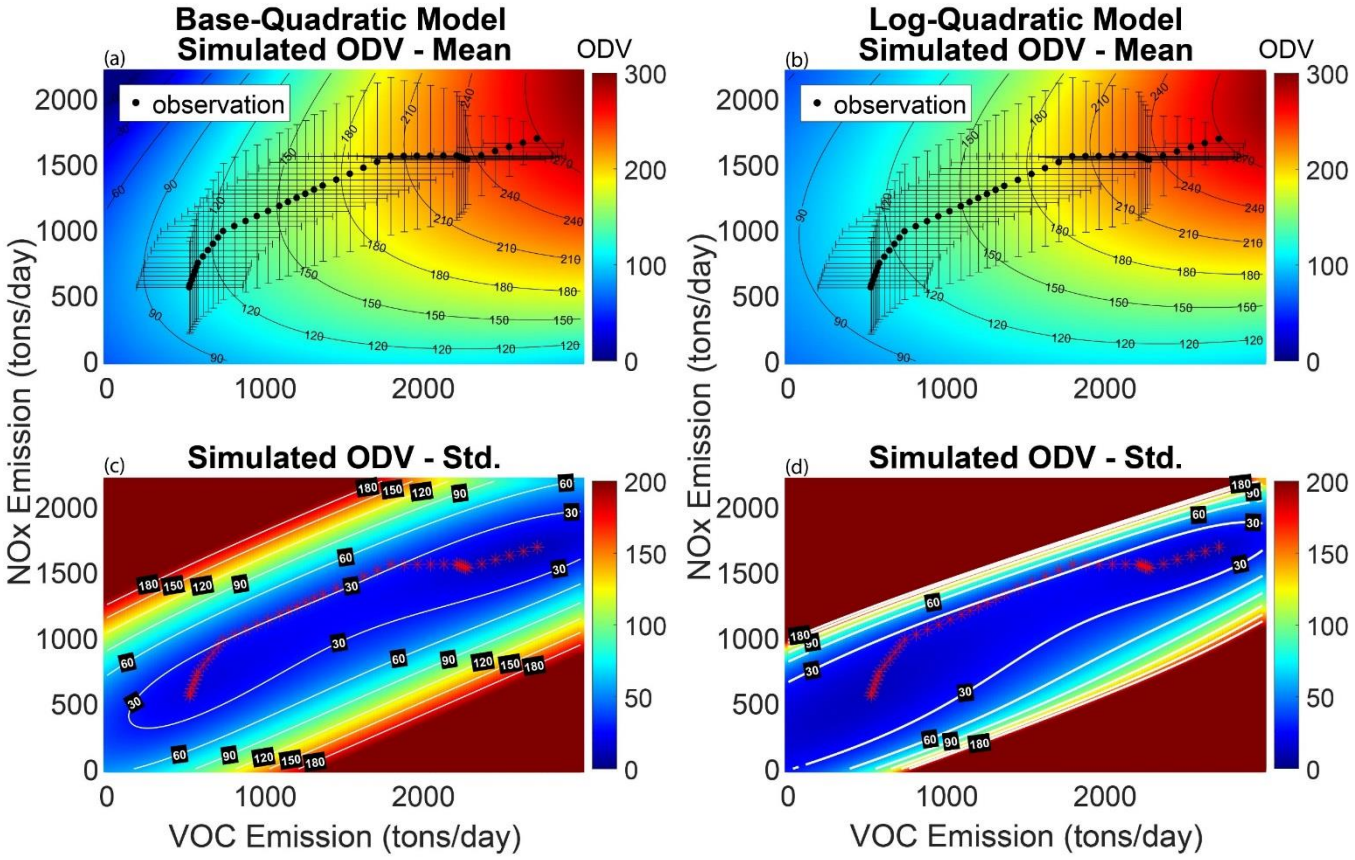

Fig. S7 Simulation result of independent variable (emissions) uncertainty analysis. (a, b) ODV-Emissions Isopleths developed by mean of ODVs modeled by 1000 simulated regression models, at different emissions levels. Cross error bars indicate the double standard deviation of 1000 simulated emissions at each year's estimated emissions level. (c, d) The uncertainty of modeled ODVs related to the uncertainty of independent variables. Axes define the emissions space with varying levels of estimated NOx and VOC and corresponding modeled ODV indicated by color. Historical observations are noted with black dots/red asterisks. Left column shows the result of full base-quadratic model (a, c). Right column shows the result of full log-quadratic model (b, d). Top Row: emission space and Isopleths constructed based on different regression models(a, b). Bottom Row: Standard deviation of ODVs modeled by 1000 simulated regression models, at different emissions levels (c, d).

Comparison between the trends of CARB estimated emissions inventory with observed concentrations for CO, NOx, and ROG.

One question that arises is how well the estimated emissions trends agree with observations in the emissions. For example, a recent article by McDonald et al.,(McDonald et al., 2018) suggested that there may be significant uncertainties in the VOC emissions estimated due to the presence of volatile consumer products (VCPs). Prior studies, (e.g., Pollack et al.(Pollack et al., 2013), Warneke et al.(Warneke et al., 2012) found good agreement between observations and emission inventories. Here, we also obtained observed concentrations of individual organics observed in Central Los Angeles (LA North Main Street), and compared them with the estimated emissions. Like the other studies, similar trends are found in the observations and emissions (Fig. S8), except for a discontinuity in 2012, after which a similar decreasing trend is found. *We have been unable to ascertain the reason for the large increase that one year.* Comparisons between observed NOx concentrations and emissions, and CO concentrations and emissions, also show similar agreement(Kim et al., 2016) (Fig. S8). In addition, the reduction rate of emissions and observations for the past two decades are in good agreement for those species (Table S3). These results do not prove that the absolute emissions estimates are correct, but provide support that the emissions trends are being captured.

Table S3 Summary of emissions and concentrations reductions from 1994 to 2016, including NOx, CO, and ROG emissions and concentrations. The trends found in the CARB inventory are in good agreement with the ambient observations

Reductions from 1994 to 2016 (%)			
NOx Emission	60%	NO2	56%
CO Emission	71%	CO	81%
ROG Emission	64%	Propane	52%
Acetylene	58%	Toluene	86%
n-Butane	59%	Ethylbenzene	65%
Benzene	82%	o-Xylene	73%

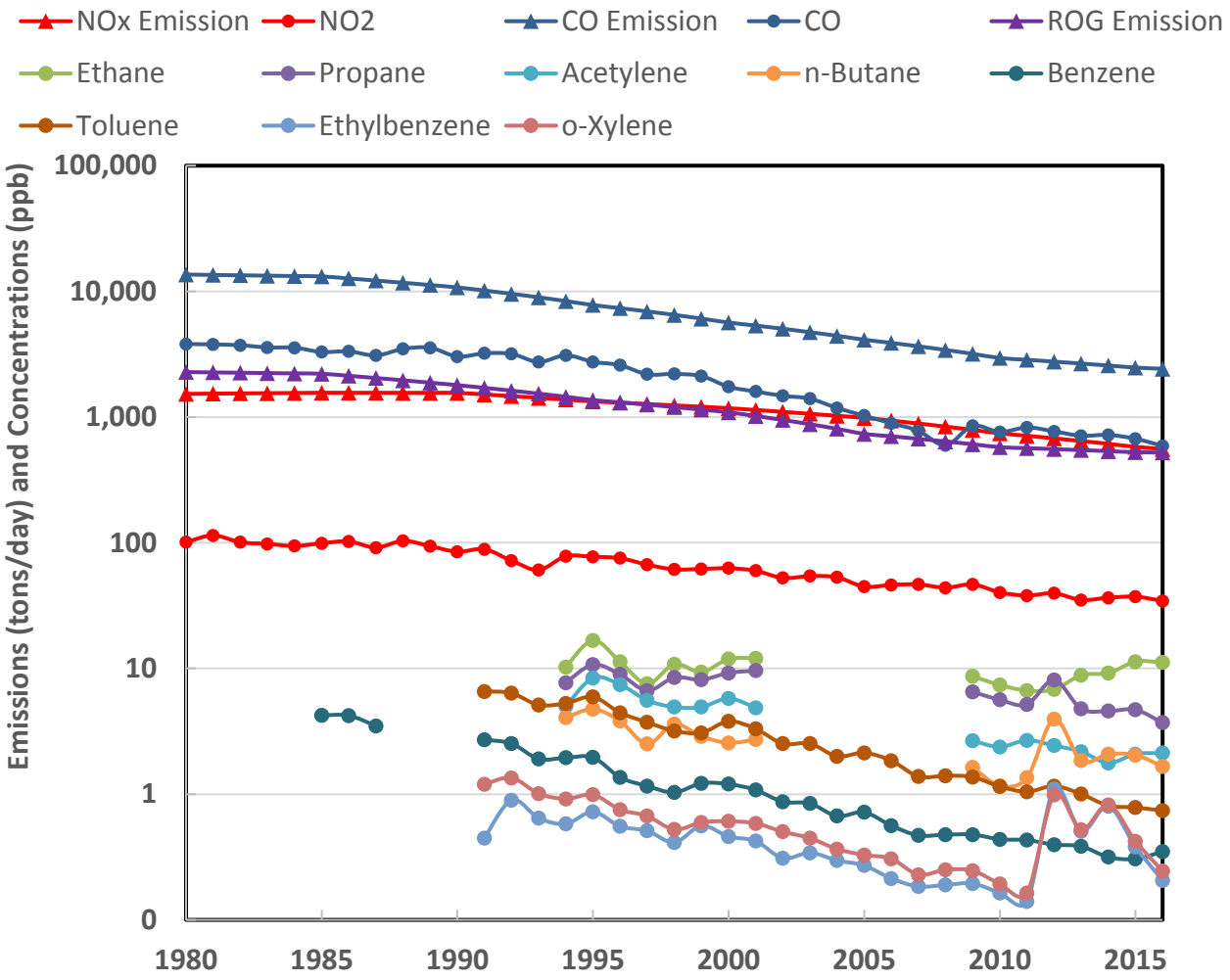

Fig. S8 Annual average trends of emissions and concentrations, in log scale, including NOx, CO, and ROG emissions and concentrations. Observation data is from LA North Main Street monitoring site (obtained from the AQS system). The trends found in the CARB inventory are in good agreement with the ambient observations, though there is a very large increase in some species in 2013, followed by reductions there after. We have been unable to ascertain the reason for the large increase in 2012.

Reference

Cox, P., Delao, A., Komorniczak, A., 2013. The California Almanac of Emissions and Air Quality - 2013 Edition.

Cox, P., Delao, A., Komorniczak, A., Weller, R., 2009. The California Almanac of Emissions and Air Quality - 2009 Edition.

Hanna, S.R., Russell, A.G., Wilkinson, J.G., Vukovich, J., Hansen, D.A., 2005. Monte Carlo estimation of uncertainties in BEIS3 emission outputs and their effects on uncertainties in chemical transport model predictions. J. Geophys. Res. Atmos. 110. https://doi.org/10.1029/2004JD004986

Helwig, N.E., 2017. Multivariate Linear Regression [WWW Document]. URL http://users.stat.umn.edu/~helwig/notes/mvlr-Notes.pdf

Kim, S.-W., McDonald, B.C., Baidar, S., Brown, S.S., Dube, B., Ferrare, R.A., Frost, G.J., Harley, R.A., Holloway, J.S., Lee, H.-J., McKeen, S.A., Neuman, J.A., Nowak, J.B., Oetjen, H., Ortega, I., Pollack, I.B., Roberts, J.M., Ryerson, T.B., Scarino, A.J., Senff, C.J., Thalman, R., Trainer, M., Volkamer, R., Wagner, N., Washenfelder, R.A., Waxman, E., Young, C.J., 2016. Modeling the weekly cycle of NOx and CO emissions and their impacts on O3 in the Los Angeles-South Coast Air Basin during the CalNex 2010 field campaign. J. Geophys. Res. Atmos. 121, 1340–1360. https://doi.org/10.1002/2015JD024292

McDonald, B.C., de Gouw, J.A., Gilman, J.B., Jathar, S.H., Akherati, A., Cappa, C.D., Jimenez, J.L., Lee-Taylor, J., Hayes, P.L., McKeen, S.A., Cui, Y.Y., Kim, S.-W., Gentner, D.R., Isaacman-VanWertz, G., Goldstein, A.H., Harley, R.A., Frost, G.J., Roberts, J.M., Ryerson, T.B., Trainer, M., 2018. Volatile chemical products emerging as largest petrochemical source of urban organic emissions. Science (80-.). 359, 760–764. https://doi.org/10.1126/science.aaq0524

Pollack, I.B., Ryerson, T.B., Trainer, M., Neuman, J.A., Roberts, J.M., Parrish, D.D., 2013. Trends in ozone, its precursors, and related secondary oxidation products in Los Angeles, California: A synthesis of measurements from 1960 to 2010. J. Geophys. Res. Atmos. 118, 5893–5911. https://doi.org/10.1002/jgrd.50472

Schwehr, B., Propper, R., 2009. Definitions of VOC and ROG [WWW Document]. URL https://www.arb.ca.gov/ei/speciate/voc_rog_dfn_1_09.pdf

Warneke, C., Gouw, J.A. de, Holloway, J.S., Peischl, J., Ryerson, T.B., Atlas, E., Blake, D., Trainer, M., Parrish, D.D., 2012. Multiyear trends in volatile organic compounds in Los Angeles, California: Five decades of decreasing emissions. J. Geophys. Res. Atmos. 117. https://doi.org/10.1029/2012JD017899