We thank the reviewer for their constructive comments. We have addressed all of them and made changes to the manuscript, indicated by italics.

1. As the results shown in Figure 2, compared with the observations, all of the four models tend to slightly overestimate the lower MDA8 O3 concentrations and to underpredict the higher ones. The four models have very small mean bias (MB, around 1ppbv) when predicting the top30 MDA8 O3 concentrations (shown in Table S3), but they all have higher MB with the average ~10 ppbv underestimation on the 4th high MDA8 O3 (shown in Table 2). As shown in Figure 3, more than 90% predicted O3 concentrations are lower than the observations, which is consistent to the underestimations on the higher MDA8 O3 shown in Figure 2. It indicates that the relationships between model inputs and predicted ozone are different at different ozone levels even addressing the highest 30 MDA8 O3 concentrations. I wonder whether lower MB and RMSE for predicting the 4th high MDA8 O3 would be expected with the empirical models developed using much higher MDA8 O3 (for example, using the data on the top 15 MDA8 O3 days).

   We thank the reviewer for pointing this out. We trained all these four models with the top 15 MDA8 ozone and made the 4th highest MDA8 ozone predictions using these four models, and the results shown below. The model performances for the top 15 MDA8 ozone and the 4th highest ozone predictions are improved, especially the traditional regression models (GAM and MARS). The reason we did not use the top 15 MDA8 ozone days to build the models in this study is that we wanted to include more potential meteorological factors that have an impact on peak ozone formation and make the models more robust with a relatively large amount of data to avoid type II errors. The RH at 850 mb and ENSO index were insignificant when we used the top 15 MDA8 ozone days to build the models. We put this table in the supplement information (Table. S5).

Table S5. Summary of statistical results of the top 15 MDA8 ozone concentrations and the 4th highest ozone predictions using four methods at Crestline site.

Method	Top 15 MDA8 ozone days			4th highest MDA8 ozone		
	Mean Bias (ppbV)	R^2	RMSE (ppbV)	Mean Bias (ppbV)	R^2	RMSE (ppbV)
GAM	0.02	0.90	8.30	-3.94	0.98	5.64
MARS	-0.27	0.89	8.55	-4.84	0.97	6.76
RF^1	-0.40	0.85	10.2	-6.09	0.97	8.12
RF^2	-0.24	0.85	10.1	-5.89	0.96	8.39
SVR^1	-1.22	0.86	9.92	-4.31	0.93	7.37
SVR^2	-1.16	0.88	9.19	-4.60	0.90	9.73

The subscript 1 and 2 in this table: RF/ SVR model with the same variables as GAM model and RF/ SVR model with the optimal combination of the indicators.

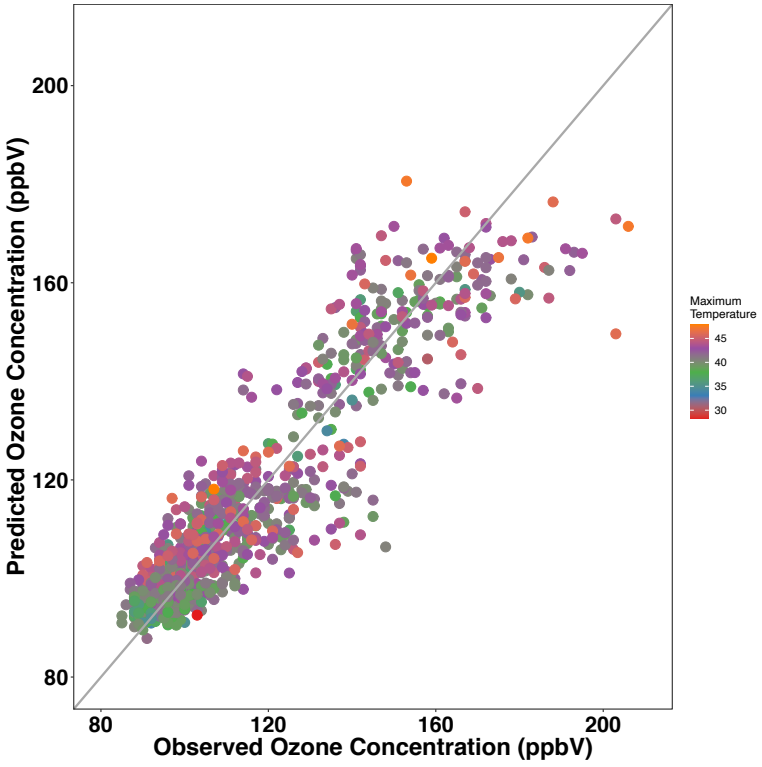
2. As discussed in the Section 3.3 (Limitations), the precursors' emissions in SoCAB and local meteorological variables have been included in the development of the four models. The structure of the built model equations in the manuscript would be applicable for those regions where top MDA8 O3 concentrations are mainly affected by local emissions. However, for the regions where the top MDA8 O3 are significantly influenced by cross-regional O3 transport, more variables might be considered in developing the predicting models (such as the precursors' emissions in surrounding regions).

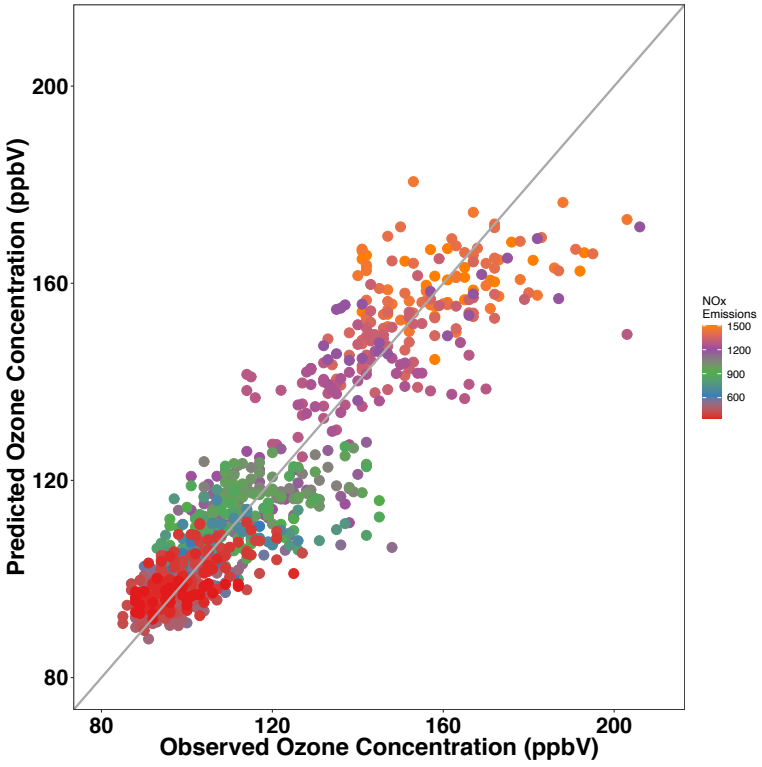
   Thank you for this suggestion. The average wind speed, wind direction, and the precursors' emissions from upwind areas can be used in the models to explain the air pollutants from upwind areas that influence the smog formation in the downwind areas. In this study, the upwind area of the Crestline site is Los Angeles. The precursors' emissions in Los Angeles are included in the annual NOx and ROG emissions in the South Coast Air Basin. Modified section 3.3 by revising and adding the following sentence between line 490 and line 493.

   "Given that Crestline is downwind of Los Angeles, which then is bordered by the Pacific Ocean, the models using SoCAB emissions capture the upwind conditions. In other regions, such models could be expanded to include both local emissions and upwind states' emissions."

3. In the study, the precursors' emissions have been proved to be the most significant factors impacting the peak O3 levels in SoCAB, and maximum temperature is of relatively high importance among all the meteorological variables. The annual NOx and VOCs emission amounts and maximum temperature from 1990 to 2019 are suggested to be illustrated together with the corresponding 4th high MDA8 O3 (or the top30 MDA8 O3 concentrations) in the Supplementary Information.

   We thank the reviewer for this suggestion. We added the figures below in the supplement information (Fig. S6). The peak ozone concentrations are usually related to a relatively high daily maximum temperature and high precursors' emissions, although some peak ozone days show the opposite relationship. The function of ozone formation is nonlinear with the ambient NOx and VOC in the presence of sunlight. The maximum temperature is highly related to sunlight, so the relationship between maximum temperature and peak ozone levels is nonlinear.


Predicted Ozone Concentration (ppbV)
200
160
120
80
80
120
160
200
Observed Ozone Concentration (ppbV)
Maximum Temperature
45
40
35
30


Predicted Ozone Concentration (ppbV)
200
160
120
80
80
120
160
200
Observed Ozone Concentration (ppbV)
NOx Emissions
1500
1200
900
600

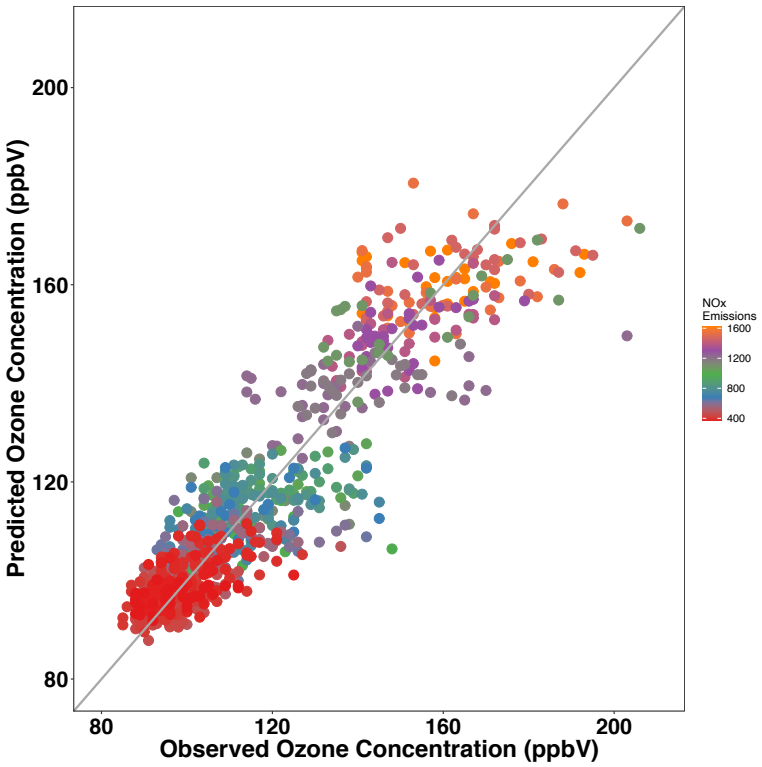


Figure S6. Observed and predicted top 30 MDA8 ozone concentrations with the corresponding annual NOx and VOC emissions and maximum temperature from 1990 to 2019 at Crestline site (the color of the points shows the value of maximum temperature, annual NOx and VOC emissions).