Supplemental Information

1 Missing data interpolation

Overall, the SOAS dataset is very complete with very few gaps in observations; however; in order to constrain the model, all gaps must be filled. If the period of missing data is less than 2 hours, a cubic interpolation of the entire time series is used to replace the missing points. If the data gap is larger, the missing points are replaced in one of the following ways:

(1) For species with no clear diurnal cycle, the measurement average is used. This primarily applies to anthropogenic VOCs with low concentrations (i.e. xylenes).

(2) For species that exhibit clear, consistent diurnal cycle but may have entire days with missing data, the diurnal average is used. This primarily applies to OVOCs (i.e. IEPOX).

(3) For isoprene, missing data is filled using a standard least squares regression of temperature and measured mixing ratios for all measurement points (r^2=0.68). Similarly, standard least squares regression of isoprene and MVK+MACR is used to fill missing MVK+MACR data (r^2=0.55).

(4) For NO, NO_2, and O_3, missing data were filled using measurements from an instrument in a nearby trailer. For HNO_3, a standard least squares regression between ground and tower observations is used to fill data gaps (r^2=0.51).

For all figures and analysis, model outputs are not included for time periods during with gaps in OH, OH reactivity, or isoprene measurements. Results are also not included for 28 June, 9 July, and 10 July are also excluded as PTR-MS and GC isoprene measurements could not be reconciled on these days.

2 Sensitivity to dilution rate

Ideally, a time-dependent dilution constant would be applied that represents mixing in of the residual layer, strong boundary layer growth throughout the morning, a maximum boundary

layer height in the afternoon boundary layer height, and little vertical mixing at night. The entrainment rate into the boundary layer is given by:

$$\text{Entrainment Rate} = \frac{v}{BLH}([X]_{FT} - [X]_{BL}) \quad (1)$$

Where *BLH* is the boundary layer height, $[X]_{BLH}$ and $[X]_{FT}$ refer to the concentration of a given species in the boundary layer and the free troposphere, and *v* is the entrainment velocity. The entrainment rate constant (k_e) is *v/BLH*. As *v* is equivalent to the change of BLH with time, we arrive at:

$$k_e = \frac{1}{BLH} d[BLH]/dt \quad (2)$$

Integrating yields:

$$k_e = \ln(\frac{BLH_{t1}}{BLH_{t2}})/dt \quad (3)$$

We model OVOC concentrations and OH reactivity for three different dilution scenarios: (1) Applying the calculated entrainment rate k_e to all days, with no other dilution, and ignoring dilution where k_e<0. The calculated k_e from BLH measurements is very sensitive to measurement noise. Therefore, we calculate k_e by taking a smoothed version of the diurnal average BLH measurement acquired by ceilometer. (2) Applying a dilution constant that scales according to the ratio of observed BLH and maximum BLH, and (3) Using a constant dilution rate of either 2 $day^{-1,}$ 4 day^{-1}, and 40 day^{-1}. Calculated dilution rates are shown in Figure S1. As in the manuscript, measured VOCs are constrained when calculating OH reactivity. Results are shown in Figures S2 and S3.

As OH reactivity is dominated by measured species, k_{dil} has minimal impact on the calculated OH reactivity. In the relationship between total OH reactivity and reactivity from isoprene, the model slope and intercept are both slightly dependent on the dilution rate (Table S1). However, under all model scenarios, the slope and intercept are slightly underestimated. All scenarios are consistent with the primary conclusions: (1) the contribution to total OH reactivity from

unmeasured, unconstrained OVOCs is small (2) there is a small but significant discrepancy in the relationship between observed and modeled total reactivity and reactivity from isoprene alone.

In contrast to OH reactivity, model OVOCs are largely a function of dilution rate. The rate of 4 day^{-1} is empirically determined based on the resultant agreement with several measured species including HCHO, glyoxal, glycolaldehyde, and PAN. Further support of this is the good agreement between measured and model IEPOX when ISOPOOH is constrained using this rate constant (Figure S4), and the agreement with the ISOPOOH, MVK+MACR, and ISOPN results from the mixed layer chemistry model presented by Su et al. (2015).

Table S1. Least squares linear fit for model OH reactivity as a function of the OH reactivity from isoprene alone under different assumed dilution rates.

Dilution Rate	**Slope**	**Intercept (s^{-1})**
Calculated entrainment rate	1.18	4.92
Scaled to BLH	1.22	6.08
2 day^{-1}	1.22	6.04
4 day^{-1}	1.22	5.36
40 day^{-1}	1.17	4.65
Measurement	1.44	6.43

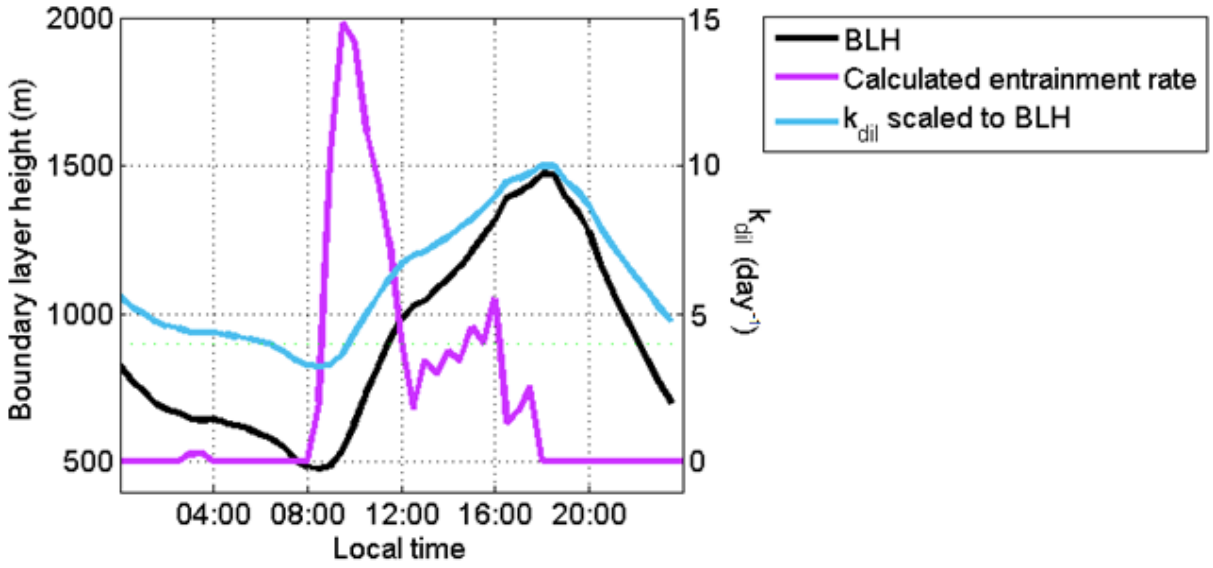

Figure S1. Diurnal average of boundary layer, smoothed over 2.5 hours, resultant dilution constant calculated according to Eq. 2, and dilution constant calculated from the ratio of BLH to maximum observed BLH.

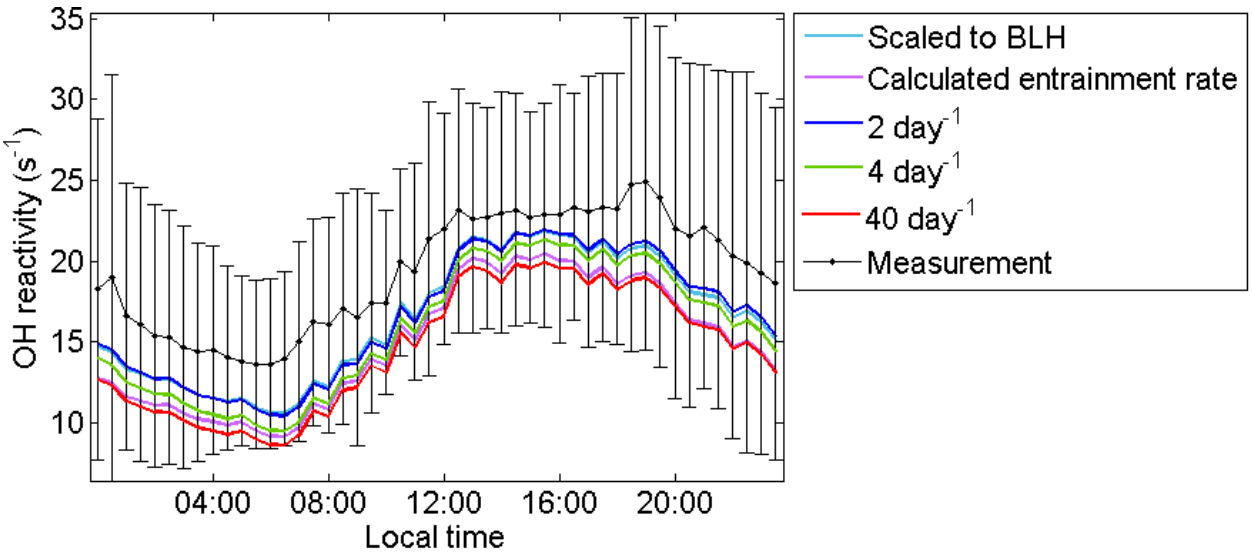

Figure S2. Sensitivity analysis for variability in the assumed dilution rate for OH reactivity. All measured species are constrained in this analysis. Error bars represent 1 σ diurnal variability in measurements. For clarity, diurnal variability is not shown in model results.

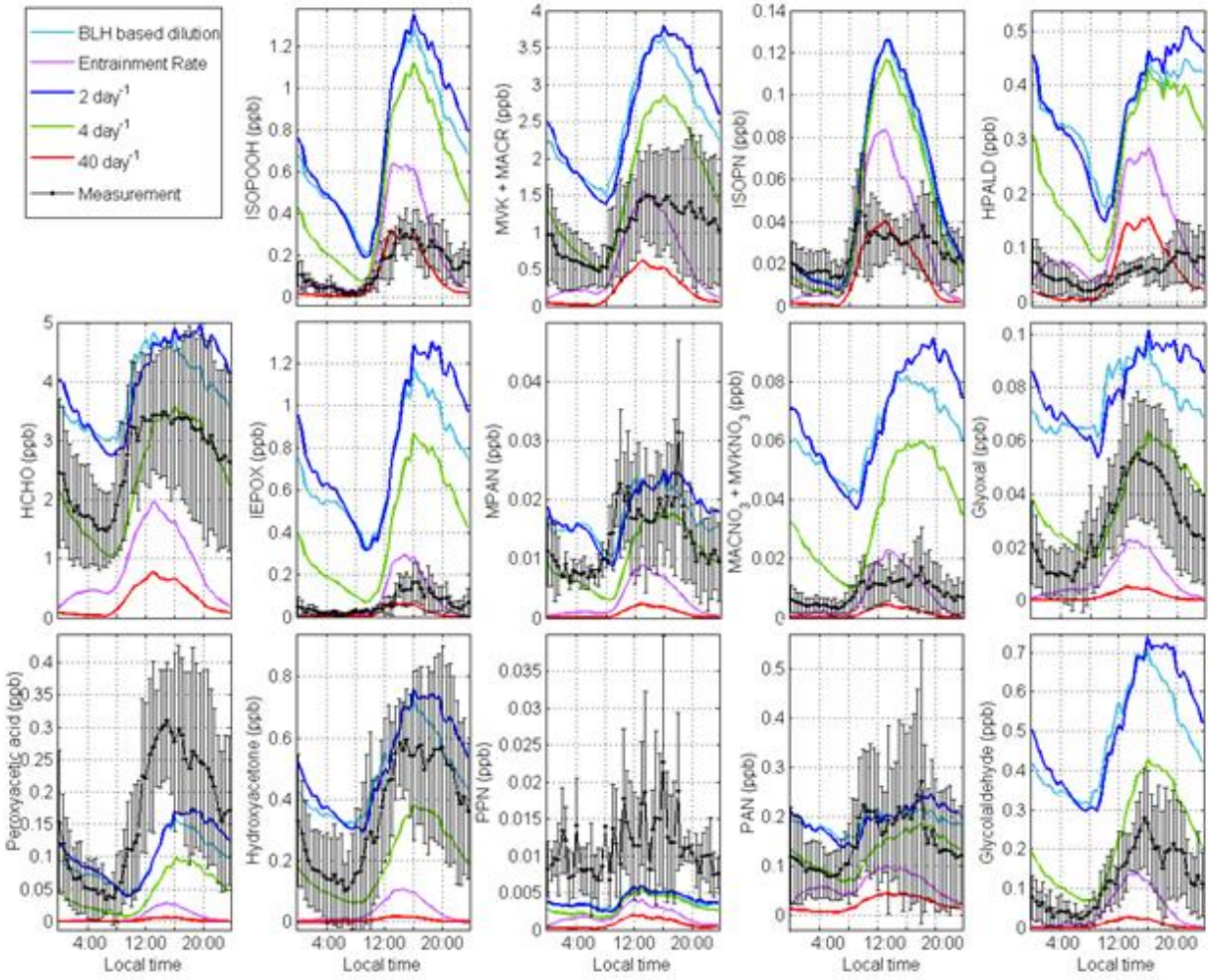

Figure S3. Sensitivity analysis for variability in the assumed dilution rate for OVOCs. Error bars represent 1 σ diurnal variability in measurements. For clarity, diurnal variability is not shown in model results. For each species, model results are not included for points where measurements are missing.

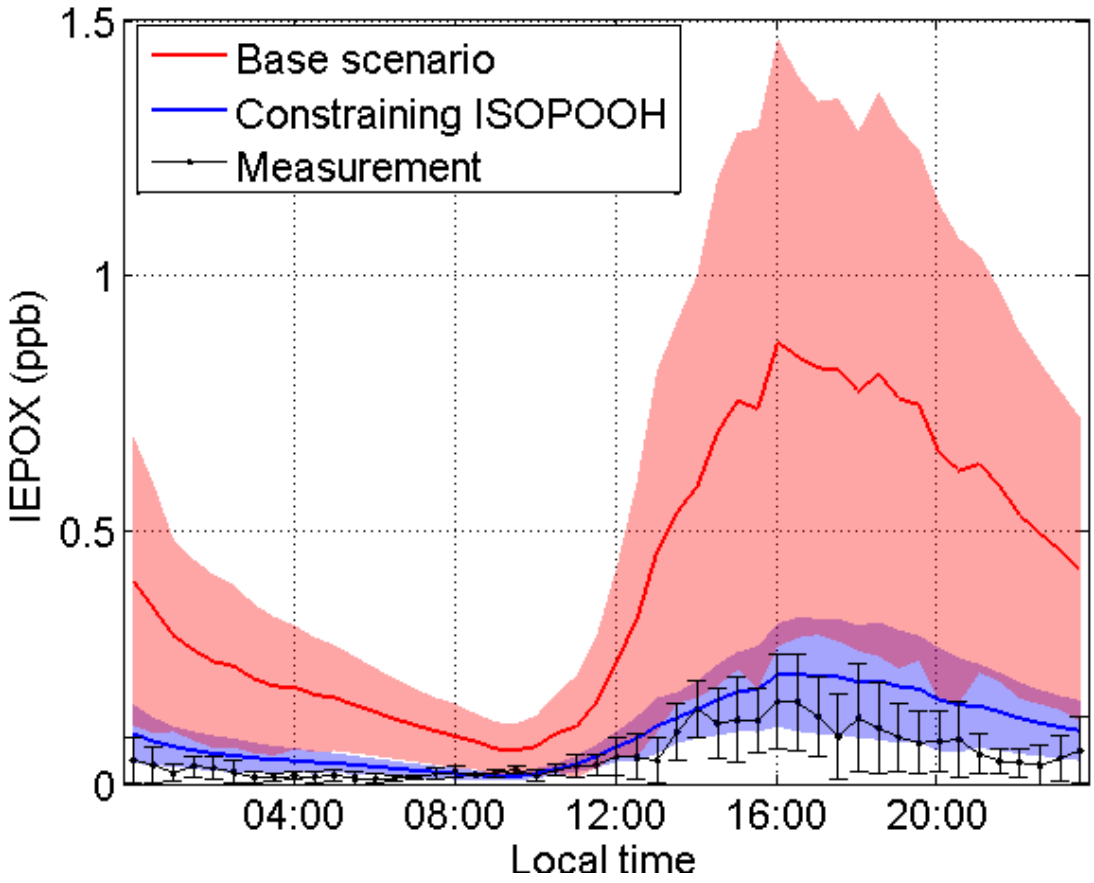

Figure S4. Comparison of measured and modeled concentrations of IEPOX with and without ISOPOOH constrained. Error bars and shaded area represent 1σ standard deviation of diurnal variability.