S1 Influence of boundary layer vs. free troposphere

Figure S1 illustrates the possible effects on the AMF of day-to-day changes in the a priori upper troposphere NO_2 profile. From Sect. 2.2, Eq. (2) and Eq. (3) show that it is the relative contribution of each altitude to the a priori profile that determines the AMF. Therefore, if an increase in near-surface NO_2 is balanced by an increase in upper tropospheric NO_2, there may be no or very little net change to the AMF. This is the case to the left of the city in Fig. S1. However, the reverse can also occur, where changes in the upper tropospheric a priori profile accentuate the effect of changes near the surface. This occurs to the right of the city in Fig. S1.

Here we consider the magnitude of the effect day-to-day changes in the free tropospheric NO_2 profile have on the retrieval. Figure S2 compares the difference in pseudo-retrieval AMFs among the three a priori types used: the daily profiles over the full extent of the troposphere, "hybrid" daily profiles which include day-to-day variation in boundary layer but use a monthly average above 750 hPa, and the monthly average profiles. These statistics are derived over the entire pseudo-retrieval domain for the entire time period. Comparing either the full or hybrid daily profile to the monthly average profile yields very similar statistics. Both exhibit a median change near $+2.5\%$ ($+2.8\%$ hybrid, $+2.5\%$ full) and show similar upper and lower quartiles. When using the full daily profile, the range of the most extreme values is slightly greater, occurring when the changes in the boundary layer and free troposphere act in the same direction on the AMF.

The third column shows the difference between the hybrid and full daily profile AMFs. As expected, the mean and median differences are 0, with upper and lower quartile values of $+2.20\%$ and -2.88%. This is $\sim 35\%$ of the interquartile range of the difference between AMFs resulting from either the full or hybrid daily profiles and the monthly average profile. Day-to-day changes in the free tropospheric a priori profile are smaller in magnitude than those in the boundary layer, but usually occur over a much greater vertical extent. This, combined with the greater scattering weights at these altitudes, explains why the effect on the AMF is as large as it is, although day-to-day changes in the boundary layer still dominate the effect using daily profiles has on the AMFs, It should be noted that these a priori are derived from a model without lightning NO emission; therefore this should be considered a lower bound for the effect of the free tropospheric profile. The presence of lightning NO_x would likely lead to larger increases in AMF due to the higher upper tropospheric sensitivity.

The implications of this response are varied. For applications (such as data assimilation) where reducing the uncertainty in a single day's observations is critical, this result indicates that accurate modeling of upper tropospheric NO_2 is important. However, over the 91 days of this study, the effect of including a daily free tropospheric a priori profile averages out to 0. Winds in the free troposphere are not correlated with surface winds (Endlick et al., 1969). In methods sorting observations by wind speed or direction (Valin et al., 2013; Lu et al., 2015), day-to-day variations in the AMF due to changes in the free troposphere will therefore be random in character. Over long periods of averaging, the impact due to these variations will have

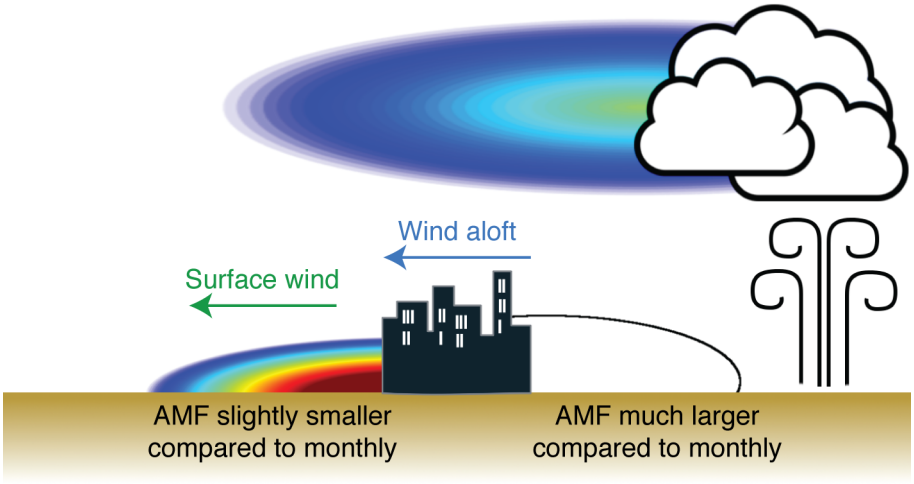

Figure S1. An illustration of how day-to-day variations in the upper troposphere impact the calculation of the AMF. The colored gradients represent the day-to-day NO_2 plumes, the black outline represents the monthly average plume. To the left of the city, the increase in near-surface NO_2 compared to the monthly average would result in a much smaller AMF; however the addition of NO_2 in the upper troposphere balances this, keeping the difference in AMF smaller. To the right, both the lack of near-surface NO_2 and the introduction of NO_2 in the upper troposphere result in a much greater AMF than the monthly average.

	Percent of days with Δ VCD $\geq 1.0 \times 10^{15}$ molec. cm^{-2}	Percent of days with Δ VCD $\geq 0.6 \times 10^{15}$ molec. cm^{-2}	Percent of days with Δ VCD $\geq 20\%+$ 0.6×10^{15} molec. cm^{-2}
Atlanta	43%	75%	25%
Birmingham	59%	83%	48%
Montgomery	27%	66%	25%

Table S1. Expansion of Table 2 with different values for the single measurement uncertainty; percent of days with at least one clear pixel that have a change in retrieved tropospheric VCDs above the specified uncertainty.

no net impact. Only changes associated with the surface winds will lead to systematic changes in the results obtained by these methods.

S2 Expansion on number of days with change in VCD greater than uncertainty

In the main paper, we used 1×10^{15} molec. cm^{-2} as the uncertainty of a single pixel VCD based on the global mean uncertainty reported in Bucsela et al. (2013). Here we will consider two other criteria.

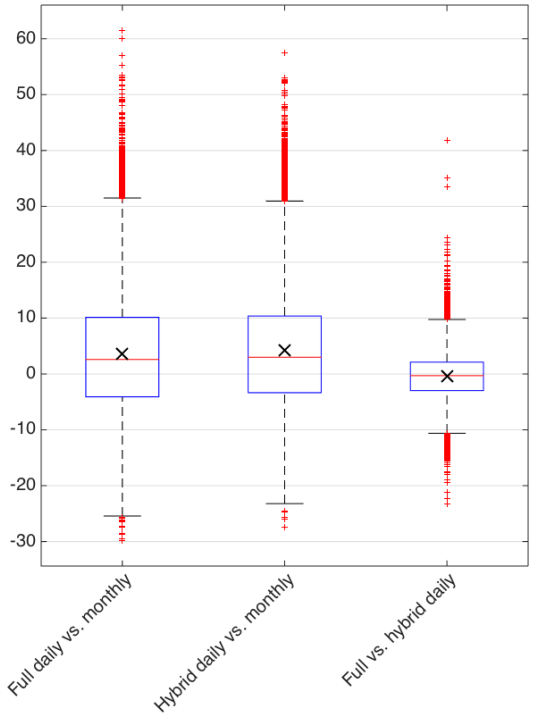

Figure S2. Percent changes in the AMF for the pseudo-retrieval over the full time period (1 June to 30 Aug) among the three a priori used in the pseudo-retrieval. For "new vs. base," the percent change is calculated as $(\text{new} - \text{base})/\text{base} \times 100\%$. The red line is the median, the box edges are the 25th and 75th percentiles, the whiskers cover the remainder of the data not considered outliers, and the red pluses are outliers. A point is considered an outlier if it is more than 1.5 times the interquartile range from the closer quartile. The black X marks the average.

Table S1 introduces two additional values for the uncertainty. The lowest value of 0.6×10^{15} molec. cm^{-2} is the quadrature sum of estimated uncertainty due to DOAS fitting (0.4×10^{15} molec. cm^{-2}) and stratospheric subtraction (0.45×10^{15} molec. cm^{-2}) in Boersma et al. (2004). We compare against this because it is the sum of errors outside the calculation of the AMF, which we are revising here. Naturally, this significantly increases the number of days for which the inclusion of a daily profile changes the retrieved VCD by an amount greater than this uncertainty. This is an overly generous assumption of uncertainty because errors in the AMF are not solely due to the a priori profile.

The other new criterion requires the difference be greater than the quadrature sum of 20% uncertainty in the AMF estimated by (Bucsela et al., 2013) and the minimum threshold uncertainty imposed by the stratospheric separation and DOAS fitting of 0.6×10^{15} molec. cm^{-3}. The percent of days with changes above this threshold decreases the most for Atlanta, Nevertheless, still up to half of days with any valid observations will exhibit significant changes in the retrieved VCD when using daily profiles.

We retain $\geq 1 \times 10^{15}$ molec. cm^{-2} for the main discussion as it is simplest reasonable criterion.

S3 Selection of constraints on and initial values of EMG fitting

The constraints listed in Table 1 were found to be necessary to ensure that a physically realistic fit was obtained. The following upper and lower limits for each of the parameters were imposed:

- $a \in [0, \infty)$: a corresponds to the total NO_2 burden and therefore must be a positive value.
- $x_0 \in [1.6, \infty)$: x_0 is the distance traveled by the plume in one lifetime. Therefore it must be positive; however, if it becomes too small, the fit fails (returning a flat line). At very slow wind speeds of 1 m s^{-1}, a short lifetime of 1 h would translate to an x_0 of 3.6 km, we choose a minimum value for x_0 shorter than this to allow for the possibility of very short lifetimes. 1.6 km was specifically chosen as one-third the distance between adjacent data points in the line densities on the oversampled $0.05^\circ \times 0.05^\circ$ ($\approx 5 \times 5$ km) grid. This means that three lifetimes would pass and $> 95\%$ of the NO_x enhancement due to a source would be removed; shorter lifetimes than this are unlikely to be resolved at this grid resolution.
- $\mu_x \in [\min(x), \max(x)]$: μ_x is the apparent position of the emission center relative to the geographic city center. It must be within the domain of distances chosen, otherwise the domain was chosen poorly.
- $\sigma_x \in [2.5, x_{\max NO_2} - \min(x)]$: σ_x is the Gaussian smoothing parameter, serving the same role as the σ parameter in any Gaussian function. 2.5 km was chosen as the lower bound under the assumption that the smallest observable Gaussian requires at least 3 data points (the base on either side and the maximum). As $\sigma = 4 \times$ (full width at base) this implies that the smallest observable signal is $\frac{1}{2}$ the oversampled grid resolution: if only 3 points defined the Gaussian, one-half the full width at the base ($= 2\sigma$) would fit within one grid cell. The upper limit simply specifies that σ_x cannot be larger than the distance from the upwind edge of the domain to the x-coordinate of maximum line density, i.e. that the Gaussian build-up on the upwind side is fully captured in the domain.
- $B \in [0, \max(NO_2)]$: B is the background line density. It must be positive, and should not be larger than the maximum observed line density.

Two additional constraints were imposed using the ability of `fmincon` to accept linear and nonlinear relationships between the fitting parameters:

- $x_0 + \mu_x \leq \max(x)$: Simply, one lifetime must pass between the apparent emissions center and the downwind edge of the domain. If not, the domain was chosen poorly.
- $\exp\left(\frac{\mu_x}{x_0} + \frac{\sigma_x^2}{2x_0^2} - \frac{x}{x_0}\right) \leq 20$: This is a second numerical constraint to prevent the case where the exponential goes to infinity and the error function complement goes to 0, thus creating a return value of NaN from the fitting function. (The first numerical constraint is the replacement of NaNs with infinity discussed in the main text.)

For the best-guess initial values:

- $a = \int_{x_{\min}}^{x_{\max}} \mathrm{NO_2}(x)\, dx$: Since a corresponds to the total burden of $\mathrm{NO_2}$ present, we use the integral over the domain as the initial guess.

- $x_0 = 54$: A best guess of 54 km follows from an average summer lifetime of 3 h (Lu et al., 2015) and a wind speed of 5 m $\mathrm{s^{-1}}$.

- $\mu_x = x_{\max(\mathrm{NO2})}$: Logically, one would expect the apparent emission source to be somewhat near the maximum concentration, the fitting procedure can then identify if it is slightly displaced.

- $\sigma_x = \mathrm{FWHM}/2.355$: This uses the relationship between the full width at half max and the standard deviation of a Gaussian. The half maximum line density is computed as:

$$\mathrm{HM} = \frac{\max(\mathrm{NO_2}) - \mathrm{NO_2}(x_{\min})}{2} \tag{S1}$$

  The FWHM is then:

$$\mathrm{FWHM} = |x_{\mathrm{HM}} - x_{\max(\mathrm{NO2})}| \tag{S2}$$

  where x_{HM} is the x-coordinate where the half-max exists, found by interpolating to the half maximum line density.

- $B = \min(\mathrm{NO_2})$: It is a natural guess that the background is simply the lowest observed line density.

S4 Computation of uncertainty in EMG parameters (a, x_0, μ_x, σ_x, B, E, and τ)

To compute the uncertainty in both the fitting parameters and the values of E and τ derived from them, we base our calculation off those in Beirle et al. (2011) and Lu et al. (2015). Lu et al. (2015) use the values in Table S2. The value of 25% for the uncertainty in the VCDs used in Lu et al. (2015) is below the lower bound of uncertainty given in Boersma et al. (2004), likely reflecting improvements such as the temperature correction of the $\mathrm{NO_2}$ cross section used in determining the scattering weights. Lu et al. (2015) do not reduce the uncertainty by $\sqrt{n}$, where n is the number of observations. We do reduce this uncertainty by $\sqrt{n}$, as the errors contributing to it should be random in nature. We determine n for a given fit as the smallest number of observations used in a single grid cell of the rotated column densities (§2.5). This is a conservative estimate, as the number of observations used in computing a single line density will be several times that because multiple pixels span the across wind distance. Since we oversample these pixels, each grid cell is not necessarily an independent measurement, and so we choose to treat this value conservatively.

Our computation of uncertainty also differs from Lu et al. (2015) in that we include the uncertainty for the VCDs in the calculation of uncertainty for τ. As shown in this paper, the choice of a priori profiles can introduce a spatial bias into the exponential decay related to effective lifetime, therefore the inclusion of the VCD uncertainty in the lifetime uncertainty is logical.

To compute the uncertainty due to the fitting process itself, we first need the standard deviations of the fitting parameters. We begin by computing the variance-covariance matrix as (Bard, 1974; Dovì et al., 1991):

Source	Uncertainty	Citation
VCD	$25\%/\sqrt{n}$*	Lu et al. (2015)
Across wind integration dist.	10%	Beirle et al. (2011)
Choice of wind fields	10%	Beirle et al. (2011)
Fitting uncertainty	Computed	§S5
NO_x:NO_2 ratio	10%	Beirle et al. (2011)

Table S2. Values of uncertainty for the various steps of the EMG fitting process. *Note: Lu et al. (2015) do not indicate that the uncertainty in VCDs is reduced by $\sqrt{n}$, i.e. the number of observations; we are doing so here.

$$Cov([a, x_0, \mu_x, \sigma_x, B]) \approx \mathrm{diag}(s^2\,\mathbf{H}^{-1});\ \text{where} \tag{S3}$$

$$s^2 = \frac{1}{n_{\mathrm{fit}} - 5} \sum_x \left[\mathrm{NO_2}(x) - F(x|a, x_0, \mu_x, \sigma_x, B)\right]^2 \tag{S4}$$

that is, the variance-covariance matrix of the fitting parameters is approximately equal to the inverse Hessian matrix for the fitting function (Eq. 9) evaluated at the optimum values of the fitting parameters and scaled by s^2, the sum of squared residuals for the optimum fit divided by the number of degrees of freedom. The variance for each parameter is, as usual, the corresponding element on the diagonal of the variance-covariance matrix.

In Eq. S3, we use the Hessian matrix returned by the unconstrained Matlab minimization function `fminunc`. The documentation for `fmincon` states that the Hessian matrix returned can be inaccurate; therefore we pass the fitting function (Eq. 9) to `fminunc` and initialize it at the optimum values of the fitting parameters. `fminunc` is set to only evaluate the fitting function once; this returns the unconstrained Hessian at the constrained optimum found by `fmincon`.

As in Beirle et al. (2011), we represent the uncertainty in the fitting parameters due to the fitting process as 95% confidence intervals. We compute these for each individual parameter with:

$$c_{95} = \frac{t\sigma}{\sqrt{n_{\mathrm{fit}}}} \tag{S5}$$

where t is Student's t-value and is computed in Matlab as `tinv(0.975, n_fit - 5)`. (0.975 is used as the p value because `tinv` returns one-tailed t values, so 0.975 returns the 0.95 two-tailed value.) n_{fit} is the number of points used to obtain the fit, and $n_{\mathrm{fit}} - 5$ degrees of freedom are used for determinine the value of t because 5 degrees of freedom have been fixed in the fitting process, one each for the five parameters fit. The standard deviation, σ, is taken as the square root of the corresponding diagonal element of the variance-covariance matrix from Eq. (S3). This treatment assumes that each of the fitting parameters is independent. Both the covariance matrix and an analysis of the change in optimum values for four of the parameters if one is fixed shows that the parameters are not fully independent, but using individual confidence intervals for each parameter lends itself to a more intuitive understanding of the uncertainty than would attempting to derive a five-dimensional confidence region for the five fitting parameters.

To combine this uncertainty with the other sources identified in Beirle et al. (2011) (VCD, across wind distance, and wind field), the uncertainties are then simply added in quadrature:

$$u_p = p \cdot \sqrt{\left(\frac{c_{95,p}}{p}\right)^2 + (\%u_{\mathrm{VCD}})^2 + (\%u_b)^2 + (\%u_{\mathrm{wf}})^2} \tag{S6}$$

where p is the value of the fitting parameter, u_p the total uncertainty for the parameter, $c_{95,p}$ the confidence interval for that parameter, and $\%u_{\mathrm{VCD}}$, $\%u_b$, and $\%u_{\mathrm{wf}}$ are the percent uncertainties in VCDs, across wind integration direction, and choice of wind fields given in Table S2 as decimal values (i.e. $\%u_b = 0.1$).

The uncertainty for lifetime (τ) and emissions (E) are then propagated from the fitting parameter uncertainties. For lifetime:

$$\tau = \frac{x_0}{w} \tag{S7}$$

$$\begin{aligned} u_\tau^2 &= \left(\frac{\partial \tau}{\partial x_0} u_{x_0}\right)^2 + \left(\frac{\partial \tau}{\partial w} u_w\right)^2 \\ &= \left(\frac{1}{w} u_{x_0}\right)^2 + \left(\frac{-x_0}{w^2} u_w\right)^2 \end{aligned} \tag{S8}$$

where w is the average wind speed. We compute u_w as the 95% confidence interval of the distribution of winds that fall within the bin used; that is, if we consider all days with wind speed > 3 m s^{-1}, then this is the 95% confidence interval of the distribution of wind speeds > 3 m s^{-1}. This helps account for the uncertainty in lifetime introduced because different wind speeds lead to different $\mathrm{NO_x}$ lifetimes (Valin et al., 2013).

For emissions:

$$E = \frac{raw}{x_0} = \frac{ra}{\tau} \tag{S9}$$

$$\begin{aligned} u_E^2 &= \left(\frac{\partial E}{\partial r} u_r\right)^2 + \left(\frac{\partial E}{\partial a} u_a\right)^2 + \left(\frac{\partial E}{\partial \tau} u_\tau\right)^2 \\ &= \left(\frac{a}{\tau} u_r\right)^2 + \left(\frac{r}{\tau} u_a\right)^2 + \left(\frac{-ra}{\tau^2} u_\tau\right)^2 \end{aligned} \tag{S10}$$

where r is the $\mathrm{NO_x}$:$\mathrm{NO_2}$ ratio of 1.32 from (Beirle et al., 2011).

S5 Validation of EMG fitting

To determine if the EMG fitting process is returning good fits to the data, we consider several criteria. First, we examine each fit to the data; all the best fits reproduce the shape of the data very well. Second, we consider the R value of the fits. Beirle et al. (2011) required an $R > 0.9$ for a fit to be considered acceptable. Out of 30 fits with an across-wind distance of 1° (2 wind bins for Atlanta and 3 for Birmingham times 3 a priori), only 2 had $R < 0.9$, and even then were > 0.89. The algorithm is finding good fits in nearly every case.

We also consider whether the `fmincon` algorithm is finding the global minimum of the fitting function or becoming trapped in a local minimum. As described in Sect. 2.5, we carry out 9 optimizations from random starting points, in addition to that from the best guess initial values, to sample different parts of the parameter space. We tested increasing the number of random starting points to 99 and found no change in the optimal fitting parameter values. In a separate experiment, we found only one case in which a local minimum, rather than the global minimum, was returned, but repeated fitting with 9 random start point optimizations returned the global minimum. Therefore we find that optimizing from 9 random starting points plus one best guess is a good balance between accuracy and computational efficiency.

S6 Model lifetime calculation

To compare the EMG derived $\mathrm{NO_x}$ lifetime against that from the WRF-Chem model, two loss processes were considered:

$$\mathrm{NO_2 + OH} \quad \rightarrow \mathrm{HNO_3} \tag{R1}$$

$$\mathrm{NO + RO_2} \quad \rightarrow \alpha\mathrm{RONO_2} \tag{R2}$$

where α is the $\mathrm{RONO_2}$ branching ratio. This leads to a total lifetime:

$$\tau_{\mathrm{total}} = \left(\frac{1}{\tau_{\mathrm{HNO_3}}} + \frac{1}{\tau_{\mathrm{RONO_2}}} \right)^{-1} \tag{S11}$$

The lifetime with respect to $\mathrm{HNO_3}$ is simply:

$$\tau_{\mathrm{HNO_3}} = \frac{1}{k_{R1}[\mathrm{OH}]} \tag{S12}$$

while the lifetime with respect to $\mathrm{RONO_2}$ is:

$$\tau_{\mathrm{RONO_2}} = \left(\sum_i \alpha_i k_{R2,i} [\mathrm{RO_2}]_i \right)^{-1} \tag{S13}$$

The concentrations of the $\mathrm{RO_2}$ species are not stored in the model output, so we assume steady state (Murphy et al., 2006):

$$\frac{d[\mathrm{RO_2}]}{dt} = 0 = -k_{R2}[\mathrm{RO_2}][\mathrm{NO}] + \sum_i k_{\mathrm{RH_i+OH}}[\mathrm{RH}]_i[\mathrm{OH}] \tag{S14}$$

$$\Rightarrow [\mathrm{RO_2}]_{SS,i} = \left(\sum_i k_{\mathrm{RH_i+OH}}[\mathrm{RH}]_i[\mathrm{OH}] \right) / \left(k_{R2}[\mathrm{NO}] \right) \tag{S15}$$

The overall lifetime is then:

$$\tau_{\mathrm{total}} = \left(k_{R1}[\mathrm{OH}] + \sum_i \alpha_i k_{R2,i} [\mathrm{RO_2}]_{SS,i} \right)^{-1} \tag{S16}$$

References

Bard, Y.: Interpretation of the Estimates, chap. VII, Academic Press Inc., 1974.

Beirle, S., Boersma, K., Platt, U., Lawrence, M., and Wagner, T.: "Megacity Emissions and Lifetimes of Nitrogen Oxides Probed from Space", Science, 333, 1737–1739, 2011.

Boersma, K., Eskes, H., and Brinksma, E.: "Error analysis for tropospheric NO_2 retrieval from space, J. Geophys. Res., 106, D04 311, doi:10.1029/2003JD003962, 2004.

Bucsela, E., Krotkov, N., Celarier, E., Lamsal, L., Swartz, W., Bhartia, P., Boersma, K., Veefkind, J., Gleason, J., and Pickering, K.: "A new tropospheric and stratospheric NO_2 retrieval algorithm for nadir-viewing satellite instruments: applications to OMI, Atmos. Meas. Tech., 6, 2607–2626, doi:10.5194/amt-6-2607-2013, 2013.

Dovì, V., Paladino, O., and Reverberi, A.: Some remarks on the Use of the Inverse Hessian Matrix of the Likelihood Function in the Estimation of Statistical Properties of Parameters, Appl. Math. Lett., 4, 87–90, 1991.

Endlick, R., Singleton, R., and Kaufman, J.: Spectral Analysis of Detailed Vertical Wind Speed Profiles, J. Atmos. Sci., 26, 1030–1041, 1969.

Lu, Z., Streets, D., de Foy, B., Lamsal, L., Duncan, B., and Xing, J.: "Emissions of nitrogen oxides from US urban areas: estimation from Ozone Monitoring Instrument retrievals for 2005–2014", Atmos. Chem. Phys., 15, 10 367–10 383, doi:10.5194/acp-15-10367-2015, 2015.

Murphy, J. G., Day, D. A., Cleary, P. A., Wooldridge, P. J., Millet, D. B., Goldstein, A. H., and Cohen, R. C.: The weekend effect within and downwind of Sacramento: Part 2. Observational evidence for chemical and dynamical contributions, Atmos. Chem. Phys. Discuss., 6, 11 971–12 019, doi:10.5194/acpd-6-11971-2006, 2006.

Valin, L., Russell, A., and Cohen, R.: "Variations of OH radical in an urban plume inferred from NO_2 column measurements", Geophys. Res. Lett., 40, 1856–1860, doi:10.1002/grl.50267, 2013.