Supplement of Atmos. Meas. Tech., 12, 129–146, 2019
https://doi.org/10.5194/amt-12-129-2019-supplement

Atmospheric Measurement Techniques
Open Access

Supplement of

Evaluation of version 3.0B of the BEHR OMI NO_2 product

Joshua L. Laughner et al.

Correspondence to: Ronald C. Cohen (rccohen@berkeley.edu)

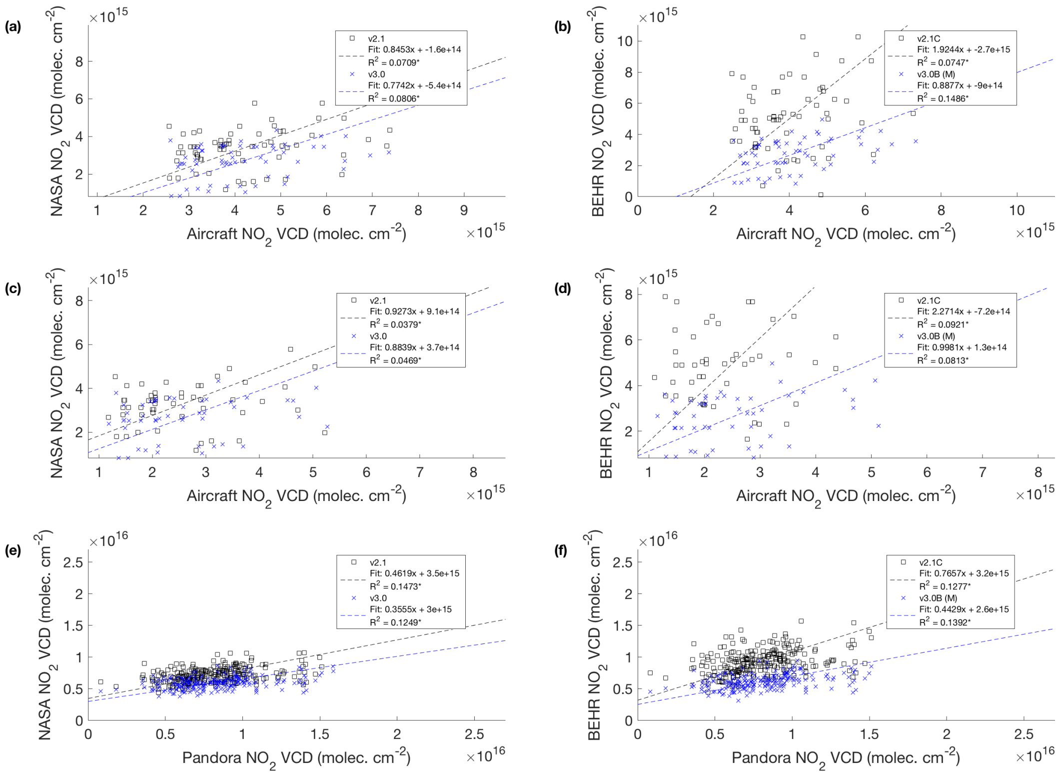

Figure S1: Scatter plots comparing **(a,c,e)** NASA Standard Product and **(b,d,f)** the BEHR product VCDs to **(a,b)** aircraft profiles extended with GEOS-Chem, **(c,d)** aircraft profiles extended by extrapolation, **(e,f)** Pandora columns measured during the DISCOVER-AQ Maryland campaign. An asterisk (*) after the R^2 value in the legend indicates the slope is statistically different from 0 at the 95% confidence level.

S1 VCD comparison detail

BEHR v3.0 intercepts and R^2 values are generally similar to or better than NASA SP v3.0, though the discrepancy in intercepts is greater when comparing against Pandora data alone. In theory, comparing against aircraft data, the intercepts would indicate a bias in the stratospheric separation or total column (for Pandora comparisons, it can only be in the total column). In practice, it is not fully orthogonal to errors in the AMF. However, the stratospheric separation and total column will still be a significant component to the intercept, so it is reasonable that the BEHR and NASA intercepts are similar, as both use the same stratospheric separation and total columns.

		Extended with GEOS-Chem			Extended by extrapolation		
Campaign	Product	Slope	Intercept	R^2	Slope	Intercept	R^2
DISCOVER-MD	BEHR v3.0B (D)	N/A	N/A	N/A	N/A	N/A	N/A
	BEHR v3.0B (M)	0.888	-8.96×10^{14}	0.149	0.998	1.28×10^{14}	0.0813
	BEHR v2.1C	1.92	-2.69×10^{15}	0.0747	2.27	-7.24×10^{14}	0.0921
	SP v3.0	0.774	-5.35×10^{14}	0.0806	0.884	3.67×10^{14}	0.0469
DISCOVER-CA	BEHR v3.0B (D)	0.627	1.12×10^{15}	0.567	0.716	9.6×10^{14}	0.587
	BEHR v3.0B (M)	0.665	9.91×10^{14}	0.521	0.775	7.98×10^{14}	0.554
	BEHR v2.1C	0.944	8.29×10^{14}	0.554	1.14	5.25×10^{14}	0.625
	SP v3.0	0.512	9.66×10^{14}	0.291	0.513	9.93×10^{14}	0.311
DISCOVER-TX	BEHR v3.0B (D)	0.483	2.66×10^{14}	0.295	0.512	4.54×10^{14}	0.275
	BEHR v3.0B (M)	0.407	3.9×10^{14}	0.321	0.425	5.54×10^{14}	0.238
	BEHR v2.1C	0.867	1.03×10^{15}	0.386	0.836	1.58×10^{15}	0.359
	SP v3.0	0.299	7.15×10^{14}	0.352	0.317	7.9×10^{14}	0.355
DISCOVER-CO ($V > 0$)	BEHR v3.0B (D)	1.03	-5.82×10^{14}	0.0279	1.01	1.26×10^{14}	0.0234
	BEHR v3.0B (M)	0.915	-5.27×10^{14}	0.0244	0.957	8.24×10^{13}	0.062
	BEHR v2.1C	1.84	-1.32×10^{15}	0.018	1.84	-1.65×10^{14}	0.0186
	SP v3.0	0.705	-4.04×10^{14}	0.0553	0.65	1.69×10^{14}	0.0423
SENEX	BEHR v3.0B (D)	2.25	-2.59×10^{15}	0.154	1.72	-1.2×10^{15}	0.0549
	BEHR v3.0B (M)	0.953	-4.6×10^{14}	0.265	0.879	-8.89×10^{13}	0.153
	BEHR v2.1C	1.43	-2.25×10^{14}	0.164	1.48	-2.44×10^{14}	0.111
	SP v3.0	1.07	-5.84×10^{14}	0.274	0.835	2.35×10^{13}	0.107
SEAC4RS	BEHR v3.0B (D)	0.888	-1.51×10^{14}	0.0122	0.695	5.89×10^{14}	0.0473
	BEHR v3.0B (M)	1.22	-4.56×10^{14}	0.228	0.971	5.81×10^{14}	0.376
	BEHR v2.1C	2.63	-1.55×10^{15}	0.597	2.53	-7.82×10^{13}	0.422
	SP v3.0	1.01	-3.97×10^{14}	0.236	0.823	4.31×10^{14}	0.154

Table S1: Slopes, intercepts, and R^2 values for RMA regression of satellite VCDs against in situ calculated VCDs. Outliers are removed before calculating these parameters; negative VCDs are retained unless noted.

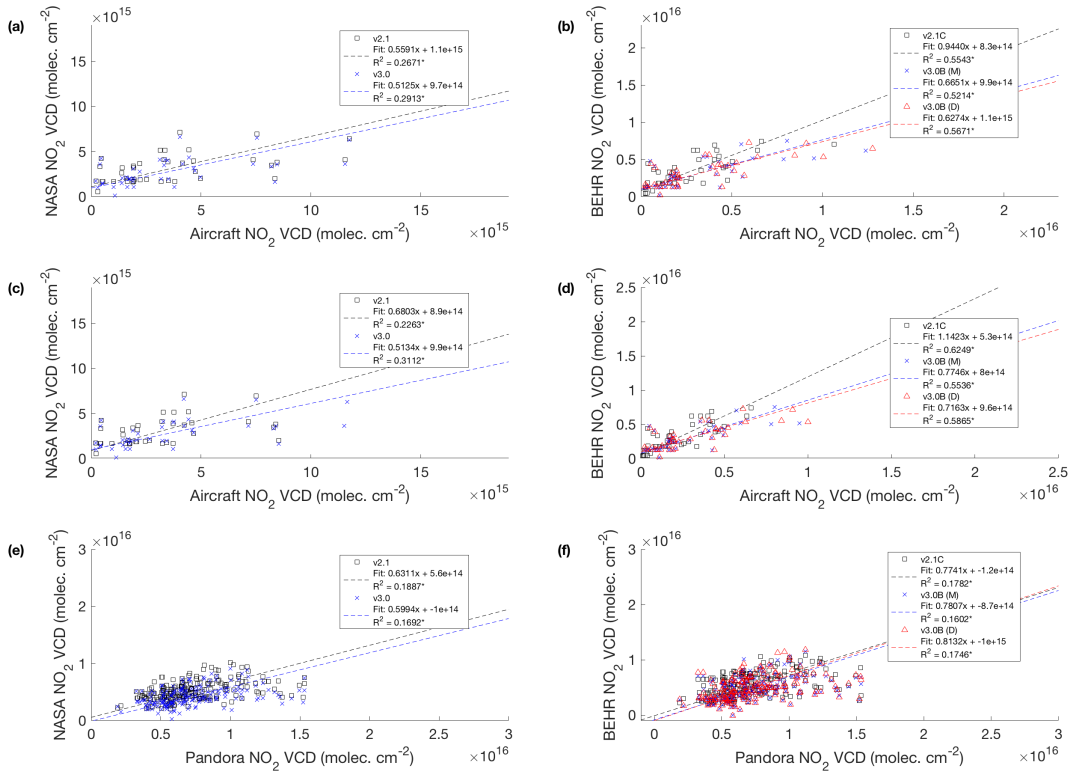

Figure S2: Scatter plots comparing **(a,c,e)** NASA Standard Product and **(b,d,f)** the BEHR product VCDs to **(a,b)** aircraft profiles extended with GEOS-Chem, **(c,d)** aircraft profiles extended by extrapolation, **(e,f)** Pandora columns measured during the DISCOVER-AQ California campaign. An asterisk (*) after the R^2 value in the legend indicates the slope is statistically different from 0 at the 95% confidence level.

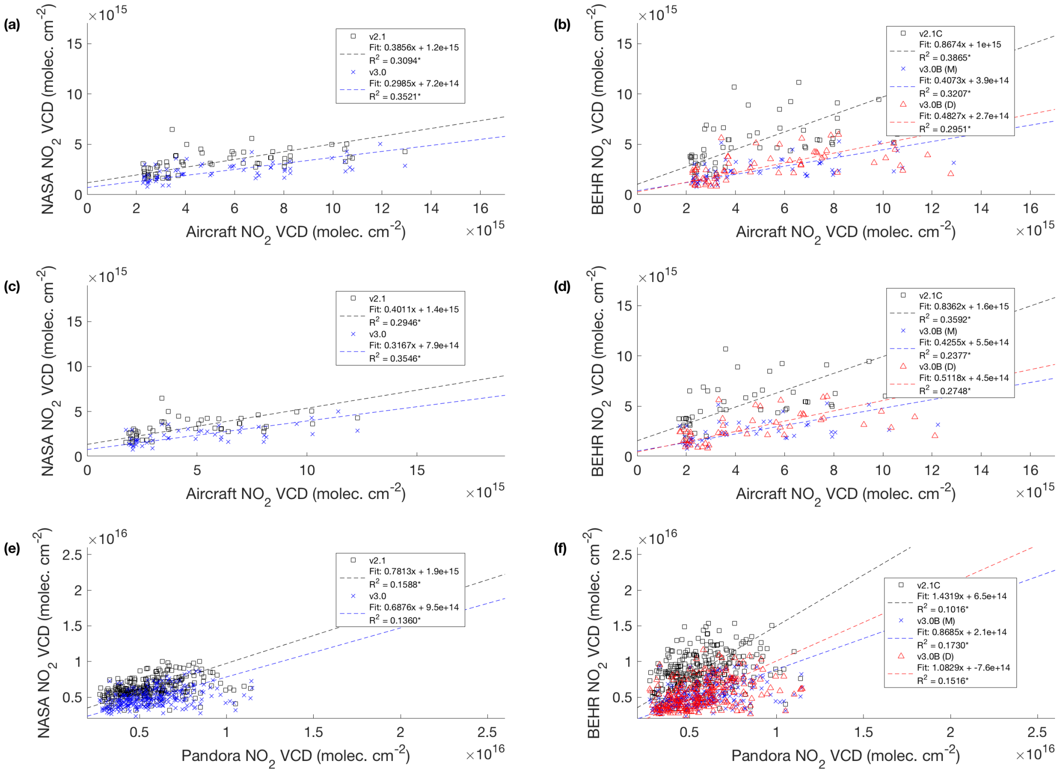

Figure S3: Scatter plots comparing **(a,c,e)** NASA Standard Product and **(b,d,f)** the BEHR product VCDs to **(a,b)** aircraft profiles extended with GEOS-Chem, **(c,d)** aircraft profiles extended by extrapolation, **(e,f)** Pandora columns measured during the DISCOVER-AQ Texas campaign. An asterisk (*) after the R^2 value in the legend indicates the slope is statistically different from 0 at the 95% confidence level.

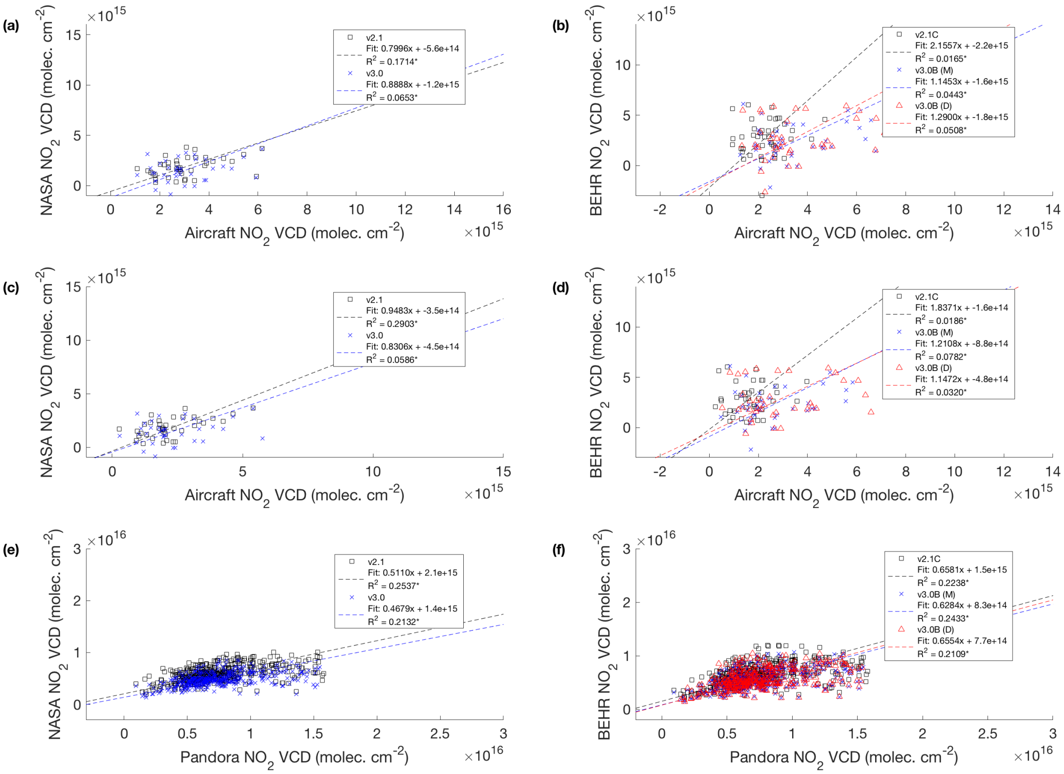

Figure S4: Scatter plots comparing **(a,c,e)** NASA Standard Product and **(b,d,f)** the BEHR product VCDs to **(a,b)** aircraft profiles extended with GEOS-Chem, **(c,d)** aircraft profiles extended by extrapolation, **(e,f)** Pandora columns measured during the DISCOVER-AQ Colorado campaign. Negative VCDs are not removed, in contrast to Table S3. An asterisk (*) after the R^2 value in the legend indicates the slope is statistically different from 0 at the 95% confidence level.

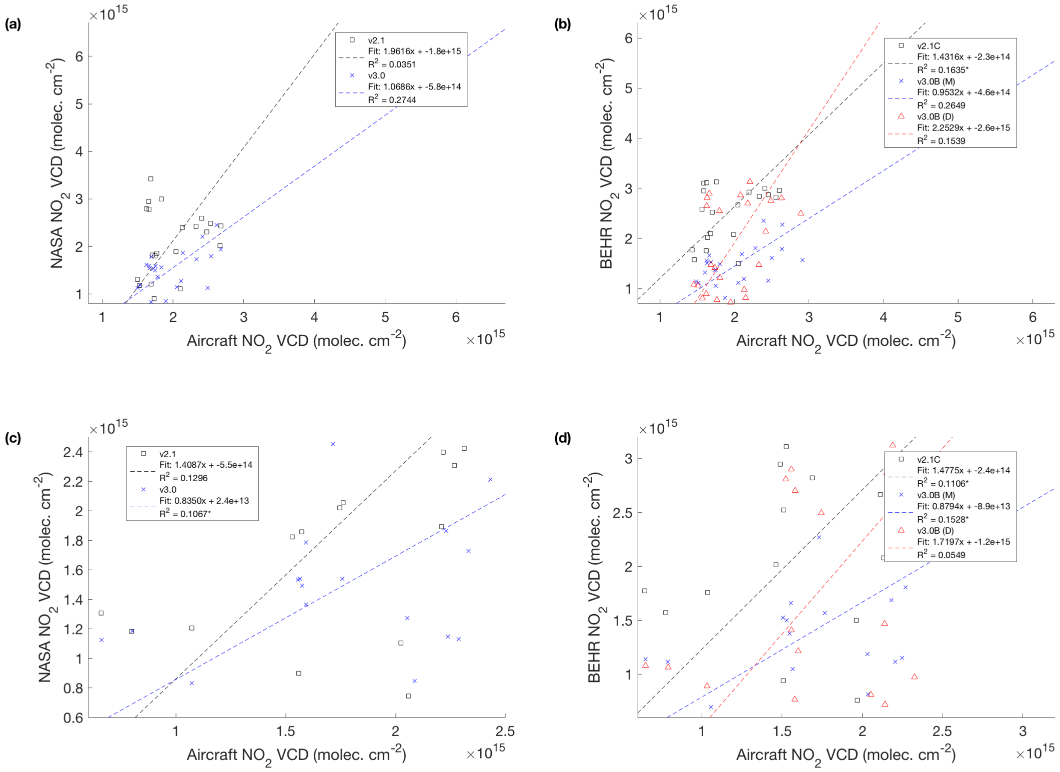

Figure S5: Scatter plots comparing **(a,c)** NASA Standard Product and **(b,d)** the BEHR product VCDs to **(a,b)** aircraft profiles extended with GEOS-Chem and **(c,d)** aircraft profiles extended by extrapolation measured during the SENEX campaign. An asterisk (*) after the R^2 value in the legend indicates the slope is statistically different from 0 at the 95% confidence level.

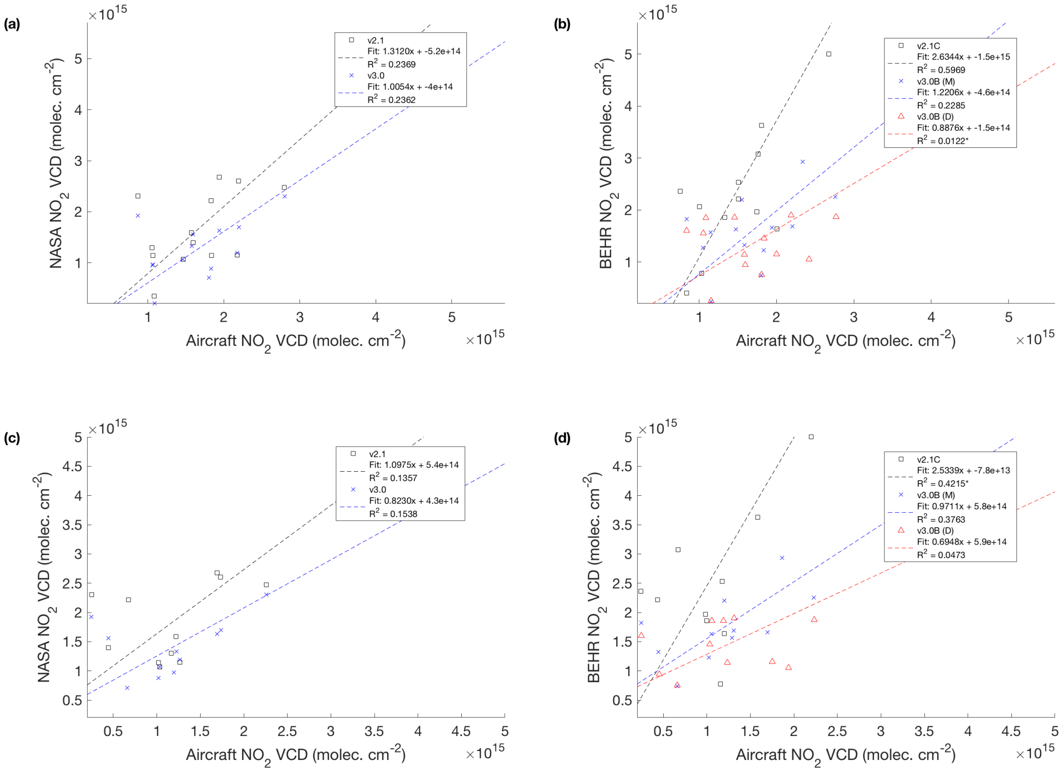

Figure S6: Scatter plots comparing **(a,c)** NASA Standard Product and **(b,d)** the BEHR product VCDs to **(a,b)** aircraft profiles extended with GEOS-Chem and **(c,d)** aircraft profiles extended by extrapolation measured during the SEAC4RS campaign. An asterisk (*) after the R^2 value in the legend indicates the slope is statistically different from 0 at the 95% confidence level.

Campaign	Product	Slope	Intercept	R^2
DISCOVER-MD	BEHR v3.0B (D)	N/A	N/A	N/A
	BEHR v3.0B (M)	0.443	2.55×10^{15}	0.139
	BEHR v2.1C	0.766	3.2×10^{15}	0.128
	SP v3.0	0.355	3.02×10^{15}	0.125
DISCOVER-CA	BEHR v3.0B (D)	0.813	-9.99×10^{14}	0.175
	BEHR v3.0B (M)	0.781	-8.66×10^{14}	0.16
	BEHR v2.1C	0.774	-1.22×10^{14}	0.178
	SP v3.0	0.599	-1.01×10^{14}	0.169
DISCOVER-TX	BEHR v3.0B (D)	1.08	-7.57×10^{14}	0.152
	BEHR v3.0B (M)	0.868	2.05×10^{14}	0.173
	BEHR v2.1C	1.43	6.54×10^{14}	0.102
	SP v3.0	0.688	9.54×10^{14}	0.136
DISCOVER-CO (V > 0)	BEHR v3.0B (D)	0.655	7.71×10^{14}	0.211
	BEHR v3.0B (M)	0.628	8.26×10^{14}	0.243
	BEHR v2.1C	0.658	1.5×10^{15}	0.224
	SP v3.0	0.468	1.38×10^{15}	0.213

Table S2: Slopes, intercepts, and R^2 values for RMA regression of satellite VCDs against Pandora VCDs. Outliers are removed before calculating these parameters.

		Matched Data			All Data		
Campaign	Product	Slope	Intercept	R2	Slope	Intercept	R2
DISCOVER-MD	BEHR v3.0B (D)	N/A	N/A	N/A	N/A	N/A	N/A
	BEHR v3.0B (M)	0.804	-4.16×10^{14}	0.407	0.637	8.46×10^{14}	0.361
	BEHR v2.1C	1.27	-2.92×10^{14}	0.292	0.874	2.23×10^{15}	0.304
	SP v3.0	0.788	-4.65×10^{14}	0.383	0.496	1.72×10^{15}	0.286
DISCOVER-CA	BEHR v3.0B (D)	0.493	1.3×10^{15}	0.525	0.677	1.27×10^{14}	0.321
	BEHR v3.0B (M)	0.509	1.18×10^{15}	0.514	0.661	1.5×10^{14}	0.317
	BEHR v2.1C	0.568	1.46×10^{15}	0.469	0.682	6.54×10^{14}	0.321
	SP v3.0	0.407	1.31×10^{15}	0.408	0.539	4.11×10^{14}	0.282
DISCOVER-TX	BEHR v3.0B (D)	0.692	1.35×10^{14}	0.157	1.00	-7.39×10^{14}	0.208
	BEHR v3.0B (M)	0.598	3.22×10^{14}	0.217	0.871	-2.26×10^{14}	0.221
	BEHR v2.1C	1.06	1.07×10^{15}	0.19	1.33	7.16×10^{14}	0.173
	SP v3.0	0.527	5.25×10^{14}	0.12	0.736	2.45×10^{14}	0.136
DISCOVER-CO	BEHR v3.0B (D)	0.663	6.38×10^{14}	0.394	0.664	6.92×10^{14}	0.267
	BEHR v3.0B (M)	0.696	2.68×10^{14}	0.347	0.628	7.56×10^{14}	0.277
	BEHR v2.1C	0.741	1.09×10^{15}	0.422	0.679	1.33×10^{15}	0.326
	SP v3.0	0.531	4.32×10^{14}	0.312	0.498	1.05×10^{15}	0.224

Table S3: Slopes, intercepts, and R^2 values of BEHR vs. combined aircraft (extended with GEOS-Chem profiles) and Pandora VCDs. Matched slopes use only Pandora data approximately coincident with aircraft profiles to get similar sampling; all uses all valid Pandora data. Outliers and negative VCDs are removed before computing slopes.

S2 Impact of hypsometric surface pressure correction

While we did not carry out an explicit test of how each change to the BEHR algorithm between v2.1C and v3.0B affected the comparison vs. aircraft and Pandora data, we did investigate the effect of different methods of computing the surface pressure of the OMI pixels. The AMF calculation requires a priori knowledge of the average surface pressure of the each OMI pixel, as the location of the surface affects the shape of the scattering weights (e.g. a low reflectivity surface high up in the atmosphere will cause the scattering weights to decrease more rapidly with decreasing altitude than a surface lower down in the atmosphere).

In BEHR v3.0A and earlier versions, this surface pressure was computed by averaging surface elevation data from the GLOBE database (Hastings and Dunbar, 1999) within the OMI pixel, which is then converted to from elevation to pressure using a 7.4 km scale height. In v3.0B, surface pressure taken from the same WRF-Chem model that supplied the NO_2 profiles is adjusted using the same average GLOBE surface elevation in the method described by Zhou et al. (2009). The Zhou et al. (2009) method was originally intended to downscale very coarse ($\sim 3^\circ \times 2^\circ$) modeled surface pressure to OMI pixels using a high resolution terrain database, therefore the effect of using it with already high resolution modeled surface pressure has not been tested.

Laughner et al. (2018) showed that switching to the (Zhou et al., 2009) method increased BEHR NO_2 VCDs by $\sim$ 5 to 10% over the Rocky Mountains during the summer months. This is a small but systematic change, and so was investigated as a way to correct the low bias in BEHR NO_2 VCDs vs. aircraft and Pandora measurements during the DISCOVER-AQ Colorado campaign.

Figure S7 shows regressions of BEHR VCDs against aircraft + Pandora VCDs using both the scale height and hypsometric equation methods of computing surface pressure. The latter method improves the slope by $\sim$ 18%, and while there is a small increase in most BEHR VCDs, the reduction of the 4 largest aircraft VCDs has a larger effect on the slope. The aircraft VCDs change because, when computing a VCD from the aircraft profiles, we integrate from the OMI pixel surface pressure to its tropopause pressure, for consistency between the aircraft VCD calculation and BEHR AMF calculation.

As the two methods of calculating surface pressure do not significantly alter the BEHR VCDs in this comparison, we cannot say explicitly that the surface pressure calculated with the hypsometric equation improves the BEHR retrieval. However, using that surface pressure does lead to greater consistency between BEHR and aircraft VCDs when also applied as the lower limit for integrating the aircraft profiles.

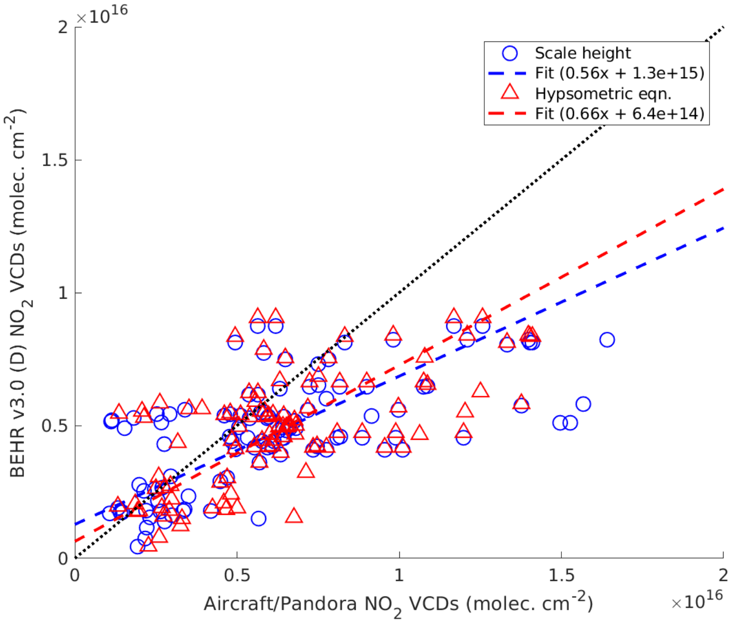

Figure S7: Regression of BEHR v3.0B (D) VCDs vs. aircraft + matched Pandora VCDs for the DISCOVER-CO campaign. The blue series and fit used surface pressured computed using a 7.4 km scale height to convert GLOBE elevations into pressures; the red series used the hypsometric equation as in Zhou et al. (2009).

S3 WRF Lightning - Individual Events

The analysis of a individual convective event taking place near the boundary of Alabama and Georgia on June 14 2012 in shown in Fig. S8. The spatial extent of flashes simulated by WRF-Chem is much broader than that measured by ENTLN, and outside of a few grid cells, the ENTLN flash counts are substantially less than the WRF-Chem simulation.

Outside of the southeast US, although the overall agreement in flash density improves, on smaller scales, we still see that flash density observed by ENTLN is concentrated in the convective core while the simulated flash density spreads over the convective area and fails to reproduce the gradient across the convective core (Fig. S9). The simulated flash density in the convection core is lower than observation despite the total flash counts are still comparable.

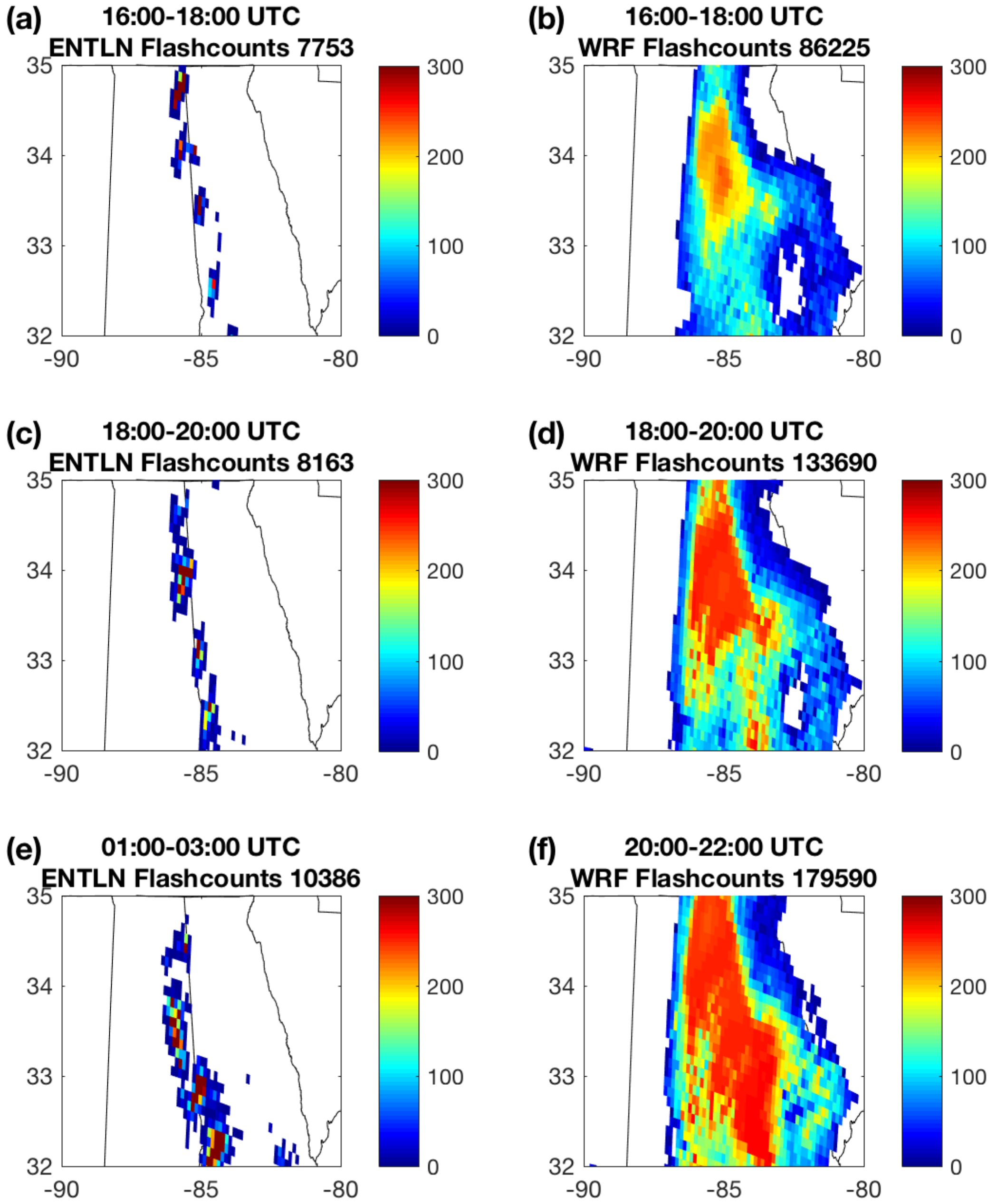

Figure S8: Time-evolved development of storm in the southeast US illustrated by lightning flashes observed by ENTLN **(a, c, e)** and simulated by WRF-Chem **(b, d, f)** on June 14 2012. The number of flashes occurring within the time range is denoted.

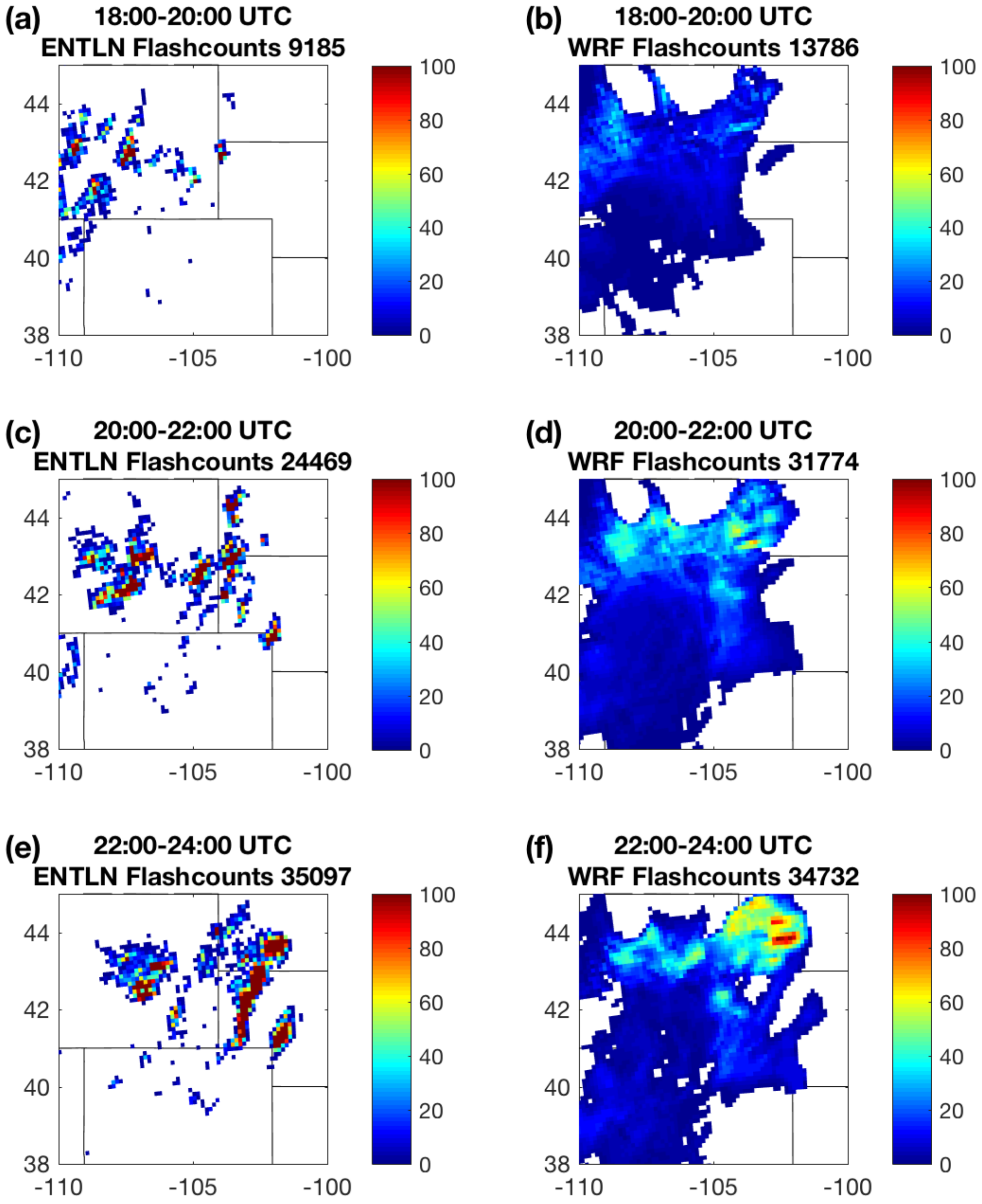

Figure S9: Time-evolved development of storm in the central US illustrated by lightning flashes observed by ENTLN **(a, c, e)** and simulated by WRF-Chem **(b, d, f)** on May 18 2012. The number of flashes occurring within the time range is denoted.

S4 Change in anthropogenic emissions in the a priori profiles, v2 to v3

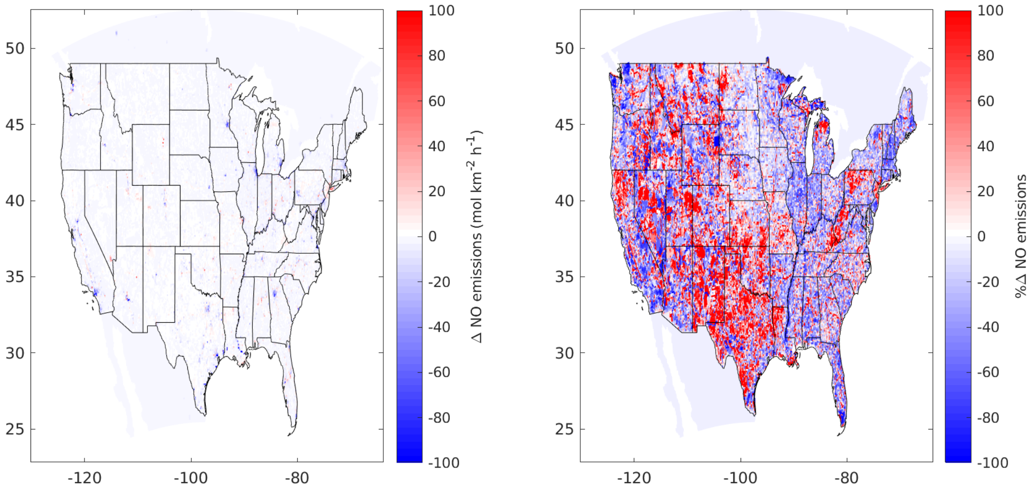

Figure S10: Absolute **(a)** and percent **(b)** change in WRF-Chem NO emissions between the EPA NEI 2005 inventory, unscaled, and the EPA NEI 2011 inventory, scaled to 2012 levels.

Figure S10 shows the absolute and percent difference between the NO emissions in the 2005 and 2011 (scaled to 2012 levels by multiplying by 0.94) EPA NEI inventories. This difference approximates the difference between the emissions used in the a priori profile simulations for BEHR v2.1C and v3.0B. (The emissions for v3.0B are exactly those used; the emissions for v2.1C were reproduced based on the description in Russell et al. 2011).

For BEHR v3.0B with daily profiles, the difference varies from year to year, but the difference will be constant over the domain, as the emissions are simply scaled by a scalar factor (Laughner et al., 2018). The difference shown in Fig. S10 will be $\sim 5\%$ and $\sim 10\%$ greater in 2013 and 2014, respectively, than in 2012. Urban areas see consistently large absolute and percent decreases. Rural areas can have large percent increases or decreases; however, the absolute difference is always very small.

S5 Surface reflectivity evaluation

In BEHR v3, we use the MODIS combined Band 3 MCD43D BRDF coefficients along with the Ross-Thick Li-Sparse kernels to compute a surface reflectivity. This is a computationally simpler approach than Vasilkov et al. (2017), who combined MODIS BRDF coefficients with the VLIDORT radiative transfer model to calculate a modified Lambertian Equivalent Reflectivity (m-LER) that assumes a uniform Lambertian surface under a scattering atmosphere. Here, we evaluate the difference resulting from using the MODIS BRDF directly.

We do so using the MODIS BRDF values using the SCIATRAN radiative transfer model (Rozanov et al., 2005). SCIATRAN is run in plane-parallel scattering mode, including polarization effect. The incident beam is assumed to be unpolarized (Stokes vector [1 0 0 0]). The aerosol profile uses the included WMO aerosol scenario, with 4 layers with upper boundaries of 2 km, 12 km, 30 km, and 100 km above the ground elevation, and aerosol types of continental, continental, background, and background, respectively.

We follow Vasilkov et al. (2017) to calculate modified-LERs (mLERs) from:

$$I(\lambda, \theta, \theta_0, \phi, P_s, R_{\mathrm{BRDF}}) = I_0(\lambda, \theta, \theta_0, \phi, P_s) + \frac{R_{\mathrm{LER}} \cdot T(\lambda, \theta, \theta_0, P_s)}{1 - R_{\mathrm{LER}} \cdot S_b(\lambda, P_s)} \tag{S1}$$

for R_{LER}, which is the mLER. The other variables are:

- I: the top-of-atmosphere intensity at wavelength λ for the given viewing zenith angle (VZA, θ), solar zenith angle (SZA, θ_0), relative azimuth angle (RAA, ϕ), surface pressure (P_s), and BRDF function (R_{BRDF})
- I_0: the top-of-atmosphere intensity for the same wavelength, geometry, and surface pressure as I, but with a 0-reflectivity (i.e. perfectly black) surface
- T: the intensity of light transmitted through the atmosphere; specifically, it represents the solar irradiance that reaches the surface, divided by π (to account for isotropic scattering from the assumed Lambertian surface, which reduces the intensity in a given solid angle), and multiplied by the transmittance of the atmosphere along the viewing direction.
- S_b: the spherical albedo of the atmosphere under the condition of illumination from below. This accounts for additional light incident on the surface due to downward scattering by the atmosphere of light already reflected from the surface.

As in Vasilkov et al. (2017), look up tables (LUTs) are created for I_0, T, and S_b. The I_0 LUT is created from the intensity output of SCIATRAN iterated over three wavelengths (450, 460, and 470 nm) and the same SZAs, VZAs, RAAs, and surface pressures used in the scattering weight LUT, with surface reflectivity set to 0. The LUTs for T and S_b are created by solving a system of linear equations obtained by rearranging Eq. (S1):

$$\begin{bmatrix} (I - I_0)_{R=0.05} & 1 \\ (I - I_0)_{R=0.1} & 1 \end{bmatrix} \begin{bmatrix} S_b \\ T \end{bmatrix} = \begin{bmatrix} \left(\frac{I-I_0}{R}\right)_{R=0.05} \\ \left(\frac{I-I_0}{R}\right)_{R=0.1} \end{bmatrix} \tag{S2}$$

where the subscripts $R = 0.05$ and $R = 0.1$ indicate that I was calculated with a Lambertian surface reflectivity of 0.05 and 0.1, respectively. These are computed for the same SZAs, VZAs, RAAs, surface pressures, and wavelengths as the I_0 table, although S_b and T are theoretically invariant with respect to some of those parameters. This holds in practice, except for S_b when both the SZA and VZA are very large.

The m-LER is then calculated at 85 sites throughout the continental United States for 189 geometries per site using MCD43D07, MCD43D08, and MCD43D09 coefficients from the first day of each month in 2005 by inputting those coefficients into SCIATRAN to calculate I in Eq. (S1) at 466 nm. Using the previously discussed LUTs for I_0, T, and S_b, we calculate the m-LER from Eq. (S1).

Finally, we calculate the BRDF albedo as in Laughner et al. (2018) for each geometry and month at each site, noting that the RAA definition for SCIATRAN is reversed from that for the BRDF kernels (i.e. $\phi_{\text{SCIA}} = 180 - \phi_{\text{MODIS}}$). We match each m-LER to the corresponding BRDF albedo for the comparison below.

We compare surface reflectances calculated directly from the MODIS coefficients and BRDF kernels with m-LERs calculated with the SCIATRAN radiative transfer model. Figure S11 shows the results for 85 sites (a combination of urban, power plant, and rural sites) with 189 geometry combinations for each site. Figure S11a shows only a 3% variation from a 1:1 line in the regression, and Fig. S11b and c shows a median difference of only 0.005 (8%), with the 75th percentile difference of 0.007 (14%). We retrieved 1 June 2012 with a 14% increase in surface reflectance and found, on average, only a $1.5 \pm 4\%$ (1σ) decrease in the NO_2 column. Since the overall effect of including the radiative transfer calculations on the retrieved columns is small, we choose to use the BRDF coefficients directly to account for the directional dependence of surface reflectance.

We do note that for surface reflectances < 0.3, larger differences in the surface reflectance obtained with radiative transfer calculations compared to the raw BRDF coefficients are associated with large solar zenith angles ($\sim 70°$). This indicates that the uncertainty in individual pixels due to the choice of surface reflectance will be greater during the winter months. However, when individual months are fit, the slope does not change significantly (range 1.011 ± 0.001 to 1.0395 ± 0.0005), indicating that the average uncertainty does not vary significantly with season. This is explored in more detail in Sect. S6 of the main paper.

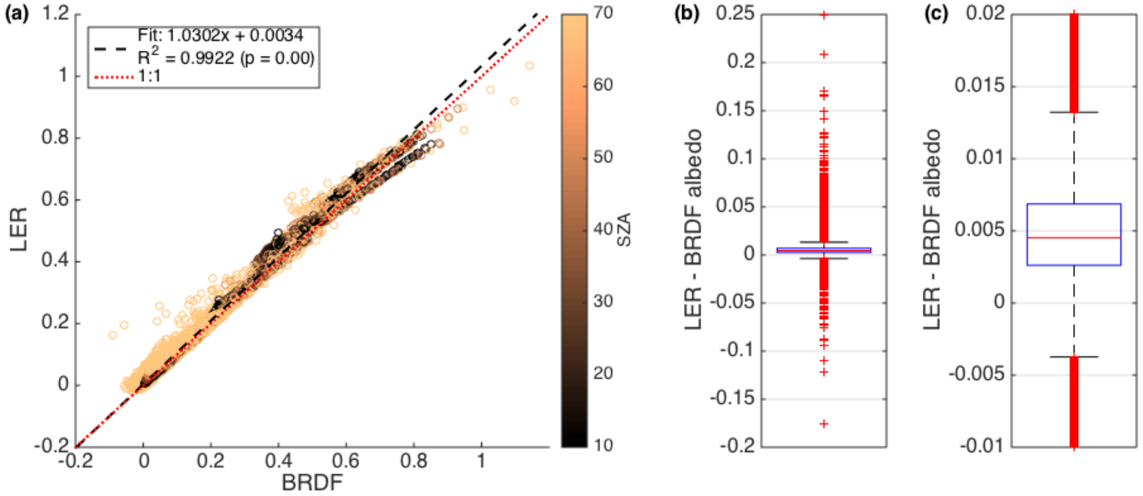

Figure S11: Comparison of a m-LER calculated with MODIS BRDF coefficients to the surface reflectance calculated directly from MODIS BRDF coefficients and kernels. **(a)** scatter plot of the m-LER on the y-axis and direct BRDF on the x-axis, colored by solar zenith angle; a reduced major axis regression is used to fit the data. **(b)** box plot of the difference between the two quantities. The red line is the median, the blue box the 25th and 75th percentiles, the black lines are the furthest non-outlier values, and the red crosses are outliers. **(c)** same as **(b)**, but zoomed in on the interquartile range.

S6 Uncertainty analysis

We determine the uncertainty in the AMF due to surface reflectance, surface pressure, tropopause pressure, cloud pressure, cloud radiance fraction, and profile shape numerically by perturbing each parameter in turn and re-retrieving the NO_2 VCDs with the perturbed values (Table S4). For each perturbation, we reretrieved all of 2012 with the varied parameter.

Surface reflectivity, surface pressure and tropopause pressure are varied by fixed percentages (surface and tropopause pressure are explicitly limited to the range 1020 to 60 hPa). The error in cloud pressure is given as a function of cloud pressure and fraction by Acarreta et al. (2004); we add and subtract the given error for each pixel. Acarreta et al. (2004) also indicates that the error in cloud fraction is < 0.05; to transform that to an error in cloud radiance fraction, we use:

$$\sigma_{f_r} = 0.05 \cdot \left.\frac{\partial f_r}{\partial f_g}\right|_{f_{g,\mathrm{pix}}} \tag{S3}$$

where f_r is the cloud radiance fraction and f_g the cloud fraction. We determine $\partial f_r / \partial f_g$ at $f_{g,\mathrm{pix}}$ (the cloud fraction of a specific pixel) by binning all f_r and f_g for the current OMI

Quantity	Perturbation	Reasoning
Surface reflectivity	±17%	Quadrature sum of 14% LER error and 10% from Schaaf et al. (2010)
Surface pressure	±1.5%	Comparing WRF and BEHR surface pressure
Tropopause presssure	Replace w/NASA tropopause	Alternate method
Cloud pressure	Variable	Fig. 3 of Acarreta et al. (2004)
Cloud radiance fraction	Cloud fraction ±0.05	Acarreta et al. (2004) with correlation of cloud frac. and cloud rad. frac.
Profiles	Quasi-Monte Carlo	Assume variability of model profiles is a reasonable metric

Table S4: Perturbation of input parameters to the AMF calculation used in the uncertainty analysis.

orbit in increments of 0.05 and using that relationship to numerically convert the error in cloud fraction to an error in cloud radiance fraction.

To determine the error due to profile uncertainty, we take advantage of the high spatial and temporal resolution of our WRF-Chem profiles, akin to Boersma et al. (2004). We run two sensitivity retrievals, first allowing the profile to be taken from any day of the same month as the satellite observation, and second allowing each pixel to shift by −0.2, 0, or +0.2 degrees in the longitudinal and latitudinal directions for the purpose of matching it with the corresponding NO_2 and temperature profiles. The first is a very conservative simulation of the possible error due to erroneous meteorological drivers (especially wind speed and direction); the second effectively simulates errors in emissions location, chemical kinetics, and transport by moving the pixel so that its profile reflects different aging time since emission.

Figure S12 shows the summed uncertainty for the four seasons as well as the individual contributions to the uncertainty. In all seasons, the a priori NO_2 profiles dominate the uncertainty. ProfileTime is the largest component in all seasons; the represents the uncertainty due primarily to errors in wind direction and speed, since it is calculated by randomly choosing profiles from a different day of the same month as the OMI data. It is unsurprising that this is the greatest contributor to uncertainty, since errors in meteorology may completely change the NO_2 profile of any given pixel, i.e. is it downwind of a source or not, as well as the impact of lightning in the SE US. This is a conservative upper bound, as we saw in Sect. 3.2, WRF captures the plume direction well $\sim 70\%$ of the time, whereas the uncertainty analysis essentially assumes that the WRF winds are uncorrelated with the real winds. Reducing the uncertainty by 70% as a rough correction would make it of similar magnitude to the other contributions and significantly reduce the total uncertainty.

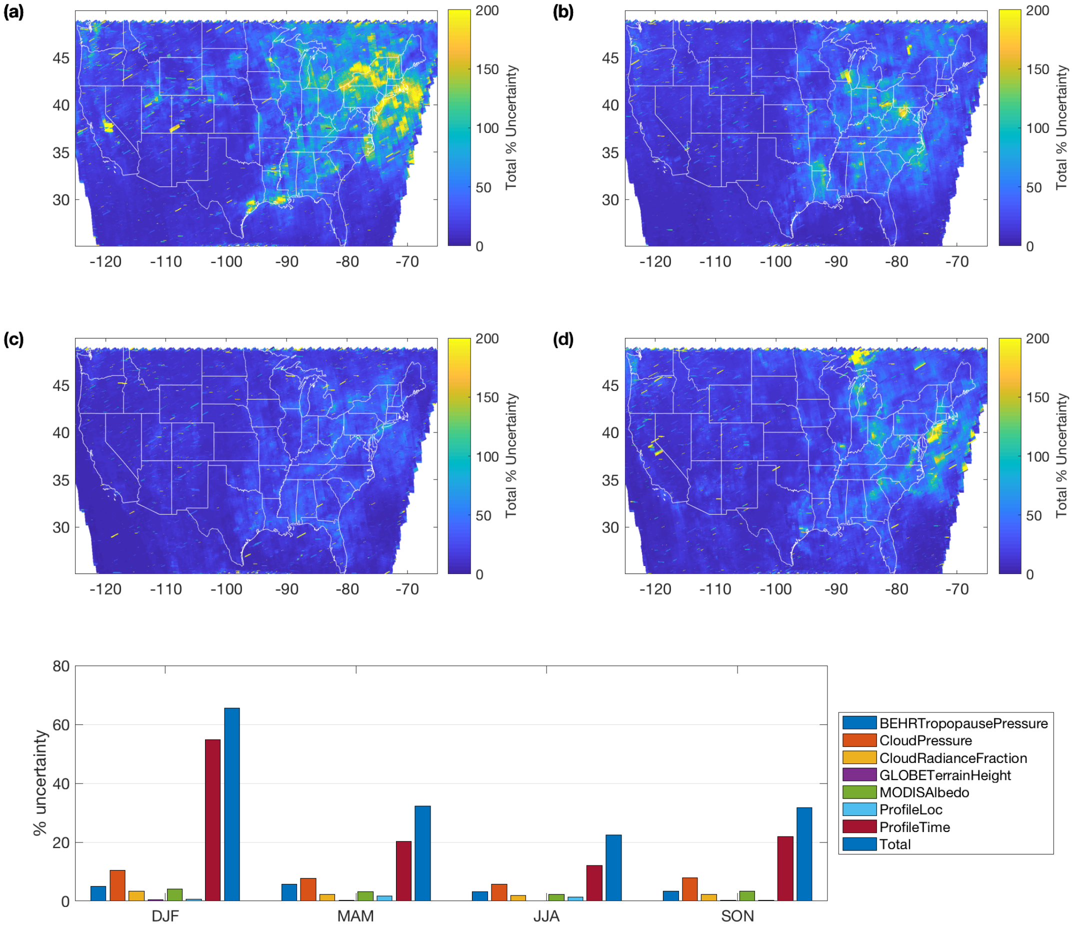

Figure S12: **(a–d)** Total percent uncertainty in tropospheric NO_2 VCDs for **(a)** Jan, Feb, Dec; **(b)** Mar.–May, **(c)** June–Aug., and **(d)** Sept.–Nov., 2012. **(e)** The domain average effect of each varied parameter and the domain average total uncertainty for the same four time periods.

The tropopause and cloud pressures are the next two largest contributors to uncertainty in most seasons. Of the non-profile contributors, the retrieval is most sensitive to cloud pressure. The retrieval sensitivity to the other four non-profile parameters is similar (~ 5 to -10%) in all seasons. The tropopause pressure is generally the second largest non-profile contributor to the uncertainty as one of the integration limits in the AMF calculation; the sensitivity of the NO_2 columns to it and the terrain height ($\%\Delta V_{\mathrm{NO_2}}/\%\Delta p_{\mathrm{trop}}$) are generally similar ($\sim 0.25\%/\%$, not shown), but the greater uncertainty in the tropopause pressure calculation causes it to have the greater impact on the retrieved VCDs.

Overall, the uncertainty due to the AMF calculation is $\sim 70\%$ in the winter, but much smaller ($\leq 30\%$) during the remainder of the year. The 30% uncertainty is similar to that calculated for polluted conditions in Boersma et al. (2004). This seems reasonable, as in winter, longer NO_x lifetime means that more pixels will have high levels of surface NO_2, and getting the wind direction wrong (i.e. what is tested with ProfileTime) will have effects over larger ares. In the summer the error in urban plumes is still important, but over smaller areas. The highest uncertainties are found in the northeast US, which has a significant number of urban areas. Our greater average uncertainty compared to Boersma et al. (2004) likely follows from the greater variability of our 12 km a priori profiles than the $5^\circ \times 3.75^\circ$ used in Boersma et al. (2004).

Data files containing the seasonal average uncertainties may be downloaded at `behr.cchem.berkeley.edu` for users who require spatially varying uncertainty information for their applications. It is also included in the data repository for this paper.

References

Acarreta, J. R., J. F. De Haan, and P. Stammes (2004). “Cloud pressure retrieval using the O2-O2 absorption band at 477 nm”. In: *J. Geophys. Res. Atmos.* 109.D5, p. D05204. ISSN: 2156-2202. DOI: `10.1029/2003JD003915`. URL: `http://dx.doi.org/10.1029/2003JD003915`.

Boersma, K. F., H. J. Eskes, and E. J. Brinksma (2004). “Error analysis for tropospheric NO2retrieval from space”. In: *J. Geophys. Res. Atmos.* 109.D4, n/a–n/a. DOI: `10.1029/2003jd003962`. URL: `https://doi.org/10.1029/2003jd003962`.

Hastings, D.A. and P.K. Dunbar (1999). *Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, Documentation, Volume 1.0. National Oceanic and Atmospheric Administration, National Geophysical Data Center, 325 Broadway, Boulder, Colorado 80303, U.S.A.* Key to Geophysical Records Documentation (KGRD) 34.

Laughner, J. L., Q. Zhu, and R. C. Cohen (2018). “The Berkeley High Resolution Tropospheric NO_2 Product”. In: *Earth System Science Data Discussions* 2018, pp. 1–33. DOI: `10.5194/essd-2018-66`. URL: `https://www.earth-syst-sci-data-discuss.net/essd-2018-66/`.

Rozanov, A. et al. (2005). “SCIATRAN 2.0 – A new radiative transfer model for geophysical applications in the 175–2400nm spectral region”. In: *Advances in Space Research* 36.5, pp. 1015–1019. DOI: `10.1016/j.asr.2005.03.012`. URL: `https://doi.org/10.1016/j.asr.2005.03.012`.

Russell, A. R. et al. (2011). “A high spatial resolution retrieval of NO_2 column densities from OMI: method and evaluation”. In: *Atmos. Chem. Phys.* 11.16, pp. 8543–8554. DOI: `10.5194/acp-11-8543-2011`. URL: `https://doi.org/10.5194/acp-11-8543-2011`.

Schaaf, Crystal Barker et al. (2010). “Aqua and Terra MODIS Albedo and Reflectance Anisotropy Products”. In: *Land Remote Sensing and Global Environmental Change.* Springer New York, pp. 549–561. DOI: `10.1007/978-1-4419-6749-7_24`. URL: `https://doi.org/10.1007/978-1-4419-6749-7_24`.

Vasilkov, Alexander et al. (2017). “Accounting for the effects of surface BRDF on satellite cloud and trace-gas retrievals: a new approach based on geometry-dependent Lambertian equivalent reflectivity applied to OMI algorithms”. In: *Atmospheric Measurement Techniques* 10.1, pp. 333–349. DOI: `10.5194/amt-10-333-2017`.

Zhou, Y. et al. (2009). “An improved tropospheric NO_2 retrieval for OMI observations in the vicinity of mountainous terrain”. In: *Atmos. Meas. Tech.* 2.2, pp. 401–416. DOI: `10.5194/amt-2-401-2009`. URL: `https://www.atmos-meas-tech.net/2/401/2009/`.